



HAL
open science

Exploration des mécanismes physiopathologiques des mucopolysaccharidoses et de la maladie de Fabry par approches "omiques" et modulation de l'autophagie.

Lénaïg Abily-Donval

► To cite this version:

Lénaïg Abily-Donval. Exploration des mécanismes physiopathologiques des mucopolysaccharidoses et de la maladie de Fabry par approches "omiques" et modulation de l'autophagie.. Médecine humaine et pathologie. Normandie Université, 2019. Français. NNT : 2019NORMR108 . tel-02498283

HAL Id: tel-02498283

<https://theses.hal.science/tel-02498283>

Submitted on 4 Mar 2020

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

Pour obtenir le diplôme de doctorat

Spécialité Sciences de la Vie et de la Santé

Préparée au sein de l'Université de Rouen

**EXPLORATION DES MECANISMES PHYSIOPATHOLOGIQUES DES
MUCOPOLYSACCHARIDOSES ET DE LA MALADIE DE FABRY PAR APPROCHES OMIQUES
ET MODULATION DE L'AUTOPHAGIE**

**Présentée et soutenue par
Lénaïg ABILY-DONVAL**

**Thèse soutenue publiquement le 19 novembre 2019
devant le jury composé de**

M. Jérôme AUSSEIL	Professeur des Universités – Praticien Hospitalier Université de Toulouse	Président du Jury
M. Cyril GOIZET	Professeur des Universités – Praticien Hospitalier Université de Bordeaux	Rapporteur
M. François MAILLOT	Professeur des Universités – Praticien Hospitalier Université de Tours	Rapporteur
Mme Annie LAQUERRIERE	Professeur des Universités – Praticien Hospitalier Université Rouen Normandie	Examineur
M. Bruno GONZALEZ	Directeur de recherche INSERM Université Rouen Normandie	Examineur
M. Stéphane MARRET	Professeur des Universités – Praticien Hospitalier Université Rouen Normandie	Codirecteur de thèse
Mme Soumeya BEKRI	Professeur des Universités – Praticien Hospitalier Université Rouen Normandie	Directrice de thèse

Thèse dirigée par

Pr Soumeya BEKRI, INSERM U1245 Université Rouen Normandie

Pr Stéphane MARRET, INSERM U1245 Université Rouen Normandie

*La partie la plus importante de notre art consiste à être en mesure
d'observer correctement.*

Laennec

Remerciements

Au Pr. Soumeya Bekri, ma directrice de thèse. Merci de m'avoir confié ce projet et de m'avoir encouragée jusqu'au bout de ce travail. Merci de m'avoir fait découvrir le monde passionnant des maladies héréditaires du métabolisme. Merci pour votre soutien malgré mon changement de projet professionnel...

Au Pr. Stéphane Marret, mon co-directeur de thèse. C'est grâce à vous que j'ai pu découvrir la recherche clinique et fondamentale et je vous en remercie. Merci de m'avoir encouragée jusqu'au bout de ce travail. Merci pour votre enseignement si précieux de la néonatalogie au lit du patient. Et enfin, merci pour votre soutien malgré mon changement de projet professionnel.

Au Pr. Jérôme Ausseil. Merci d'avoir accepté de juger ce travail et d'être président du jury. Soyez assuré de mon profond respect.

Au Pr Cyril Goizet. Merci d'avoir accepté de juger ce travail en étant rapporteur. Soyez assuré de ma profonde reconnaissance.

Au Pr François Maillot. Merci d'avoir accepté de juger ce travail en étant rapporteur. Soyez assuré de ma profonde reconnaissance.

Au Pr. Annie Laquerrière. Merci d'avoir accepté de juger ce travail en siégeant dans le jury. Merci pour vos enseignements au cours des staffs de néonatalogie et de maladies métaboliques.

Au Dr Bruno Gonzalez. Merci pour m'avoir permis de réaliser mon M2 et ma thèse au sein du laboratoire INSERM U1245. Merci d'accepter de juger ce travail. Une écoute attentive très appréciable quand j'en ai eu besoin. Je t'adresse ma sincère gratitude.

A Céline Lesueur sans qui tout cela n'aurait pas été possible... Merci pour ton aide technique mais également ton soutien et ton écoute au cours de ces années au laboratoire. Heureusement que j'ai pu bénéficier de ta présence pour finaliser ce travail. Et vive les légumes bio ! Je t'exprime ma sincère amitié.

Au Dr Abdellah Tebani. Merci beaucoup pour ta grande implication et ton aide précieuse dans ce travail. Merci de m'avoir initiée au monde de la métabolomique et merci pour la relecture attentive de ce manuscrit.

A tous les chercheurs qui ont collaboré à ce travail et en particulier un grand merci pour tous les échantillons biologiques utilisés pour cette étude.

Aux étudiants et aux post doc du laboratoire : Maryline, Nicolas, Cécile, Bérénice, Marion, Camille, Mélanie,... et à tous les autres que j'ai oublié de citer... Merci de m'avoir acceptée dans votre équipe de scientifiques. J'étais toujours accueillie avec le sourire et ça fait vraiment chaud au cœur. Je vous souhaite beaucoup de réussite dans votre vie professionnelle et personnelle.

Aux chercheurs du laboratoire. A Carole pour tes échanges et pour ton soutien. A Sylvie pour m'avoir accompagnée lors de mon M2 et pour m'avoir donné le goût de la rigueur. A Philippe pour les échanges scientifiques et pour l'étude HEMO PREMA (en espérant avoir des résultats intéressants dans quelques mois). A Isabelle et à Carine avec qui je n'ai pas eu l'occasion de travailler directement mais avec qui les échanges étaient toujours appréciables.

A Michèle avec qui j'ai passé de bons moments au laboratoire. Je te souhaite beaucoup de bonheur pour la retraite.

Aux néonatalogistes du CHU de Rouen qui m'ont laissé du temps pour pouvoir faire ces travaux de recherche et qui m'ont soutenue : Alexandra, Stéphanie, Stéphane, Aurélie, Sophie, Thierry, Didier, Caroline, Gaëlle PC. Une pensée particulière pour Stéphanie. Merci pour ton écoute attentive et ta compréhension...

Aux amies du post-internat qui m'ont encouragée et écoutée au cours de ces années difficiles : Gaëlle, Lise, Emmanuelle, Caroline, Dorothee, Anne, Marie. Je vous souhaite beaucoup de bonheur pour la suite aux quatre coins de la France...Et vive les cultures cellulaires !

A Isabelle Cordonnier et à Sabine Varin pour votre aide au cours de ce travail.

A Bénédicte, la super biochimiste ! Merci pour ton aide et ton soutien.

A l'équipe de technicien(nes) de Biochimie métabolique du CHU de Rouen. Merci pour votre accueil au sein du laboratoire et merci pour votre aide que ce soit en recherche ou en clinique.

A l'équipe des Maladies Héritaires du Métabolisme de l'hôpital Necker et en particulier au Pr. de Lonlay. Merci encore pour votre accueil au cours de mon année de médaille d'or et merci pour votre soutien.

A Nathalie Mestre pour ton soutien et ta gentillesse.

A Pascale Saugier Veber. Pour votre soutien au cours de ces années Rouennaises.

A mes patients et en particulier ceux qui m'ont beaucoup touchée ces dernières années... C'est grâce à vous que la recherche trouve son inspiration.

Et enfin, à Julien et à mes enfants chéris pour qui je n'ai pas été assez présente lors des derniers mois... Je vais me rattraper. Merci pour votre soutien et pour votre amour...

PRODUCTION SCIENTIFIQUE

Communications

1- JED 2019 : First untargeted metabolomic studies in Fabry disease. Lénaïg Abily-Donval, Nicolas Blah, Céline Lesueur, Stéphane Marret, Carlos Afonso, Abellah Tebani, Soumeya Bekri

2- JNRB 2018 : Effet de la modulation de l'autophagie dans la maladie de Fabry. Lénaïg Abily-Donval, Stéphanie Torre, Céline Lesueur, Marc Chevrier, Bruno Gonzalez, Stéphane Marret, Soumeya Bekri

3- World Symposium San Diego 2017 : Metabolic phenotyping of mucopolysaccharidoses using untargeted liquid chromatography ion mobility mass spectrometry-based strategy. Abdellah Tebani, Lénaïg Abily-Donval, Carlos Afonso, Stéphane Marret, Soumeya Bekri.

4- European Study Group on Lysosomal Diseases. Lyon 2017. Urinary metabolomics and data modeling unveil mucopolysaccharidosis type I metabolic impairments. Abdellah Tebani, Lénaïg Abily-Donval, Carlos Afonso, Stéphane Marret, Soumeya Bekri.

Publications

1- Analysis of Mucopolysaccharidosis Type VI through Integrative Functional Metabolomics. Int J Mol Sci. 2019 Jan 21;20(2). Tebani A, Abily-Donval L, Schmitz-Afonso I, Piraud M, Ausseil J, Zerimech F, Pilon C, Pereira T, Marret S, Afonso C, Bekri S.

2- Unveiling metabolic remodeling in mucopolysaccharidosis type III through integrative metabolomics and pathway analysis. J Transl Med. 2018 Sep 4;16(1) :248. Tebani A, Abily-Donval L, Schmitz-Afonso I, Héron B, Piraud M, Ausseil J, Zerimech F, Gonzalez B, Marret S, Afonso C, Bekri S.

3- Urinary metabolic phenotyping of mucopolysaccharidosis type I combining untargeted and targeted strategies with data modeling. Clin Chim Acta. 2017 Dec;475 :7-14. Tebani A, Schmitz-Afonso I, Abily-Donval L, Héron B, Piraud M, Ausseil J, Brassier A, De Lonlay P, Zerimech F, Vaz FM, Gonzalez BJ, Marret S, Afonso C, Bekri S.

4- The New Metabolic Window for Inborn Errors of Metabolism Investigations in the Post-Genomic Era. Int J Mol Sci. 2016 Jul 20;17(7). Tebani A, Abily-Donval L, Afonso C, Marret S, Bekri S.

5- Proteomics analysis in Fabry disease reveals predictive biological patterns. Tebani A, Mauhin W, Abily-Donval L, Nadjar Y, Lamari F, Laforêt P, Marret S, Lidove O, Bekri S.(article en cours de preparation)

6- Autophagy modulation enhances enzyme replacement therapy response in Fabry disease; L. Abily-Donval, F. Barbey, S. Torre, C. Lesueur, M. Chevrier, T. Pereira, C. Tardivel-Pilon, BJ. Gonzalez, S. Marret, A. Tebani, S. Bekri (article en cours de préparation)

Formations suivies (Ecole doctorale)

- Communiquer sur les sciences avec le public.
- Formation à la lecture rapide
- Valoriser ses résultats
- Spectrométrie de Masse (Pr Afonso)

ABREVIATIONS

ACPI	Atmospheric pressure chemical ionization
AMRF	Action myoclonus renal failure
AND	Acide désoxyribonucléique
AP	Adaptator protein
APPI	Atmospheric pressure photoionization
ARN	Acide ribonucléique
ATG	Autophagy related genes
Ca ⁺⁺	Ion calcium
CD-M6PR	Récepteur mannose 6 phosphate cation dépendant
CI-M6PR	Récepteur mannose 6 phosphate cation indépendant
CMH	Complexe majeur d'histocompatibilité
DAP	Death Associated Protein
DEPTOR	DEP domain containing m-TOR interacting protein
EEA	Early Endosom Antigen
EI	Ionisation électrique
ERAD	Endoplasmic Reticulum Associated protein Degradation
ESI	Electrospray ionisation
GAG	Glycosaminoglycane
GC-MS	Spectrométrie de masse avec chromatographie en phase gazeuse
GGA	Golgi localized gamma ear containing Arf-binding protein
GM	Ganglioside monosialic
HRMAS-RMN	High resolution magic-angle spinning
KO	Knock out
LAMP	Lysosome associated membrane protein
LC-MS	Spectrométrie de masse avec chromatographie en phase liquide
LIMP	Lysosome-integral membrane protein
M6P	Mannose-6-phosphate
M6PR	Mannose-6-phosphate récepteur
MALDI-MS	Matrix assisted laser desorption ionisation-mass spectrometry

MPS	Mucopolysaccharidose
MS	Spectrométrie de masse
mTOR	Mammalian or mechanistic target of rapamycin complex 1
OPLSDA	Orthogonal Partial Least Square Discriminant Analysis
PE	Phosphatidyléthanolamine
PLSDA	Partial Least Square Discriminant Analysis
RMN	Résonance magnétique nucléaire
SAP	Sphingolipid Activator Protein
SNARE	Soluble NSF attachment protein receptor
SNC	Système nerveux central
SPA	Sphingolipid Activator Protein
TA	Tension artérielle
TCSH	Transplantation cellules souches hématopoïétiques
TGN	Réseau transgolgien
TOF	Time of flight
TWIG	Travelling-wave ion guide
UVRAG	UV irradiation resistance associated gene
VIP	Variable importance
Vps	Vacuolar protein sorting
WES	Whole exome sequencing

TABLE DES MATIERES

REMERCIEMENTS	4
PRODUCTION SCIENTIFIQUE	7
ABREVIATIONS	8
LISTE DES TABLEAUX	13
LISTE DES FIGURES	13
RESUME	14
INTRODUCTION.....	16
I. Le système endosome lysosome	17
1.1 Le lysosome.....	17
1.1.1 Découverte du lysosome	17
1.1.2 Description du lysosome.....	18
1.2 L'endosome.....	23
1.2.1 Description de l'endosome	23
1.2.2 Maturation de l'endosome	24
II. L'autophagie	25
2.1 Définition	25
2.2 Les différentes voies autophagiques	25
2.2.1 Autophagie médiée par les chaperons (CMA)	25
2.2.2 Microautophagie.....	26
2.2.3 La macroautophagie	27
III. Les maladies lysosomales de surcharge	38
3.1 Définition	38
3.2 Présentation clinique générale	41
3.3 Thérapeutiques disponibles.....	41
3.3.1 Enzymothérapie	42
3.3.2 Transplantation de cellules souches hématopoïétiques	43
3.3.3 Réduction de substrat.....	43
3.3.4 Les approches variant-dépendantes (chaperons pharmacologiques).....	43
3.3.5 Thérapie génique	45
3.4 Les mucopolysaccharidoses.....	45
3.4.1 La mucopolysaccharidose de type I (maladie de Hurler, maladie de Scheie).....	47

3.4.2	Mucopolysaccharidose de type II ou maladie de Hunter	48
3.4.3	Mucopolysaccharidose de type III ou maladie de Sanfilippo	49
3.4.4	Mucopolysaccharidose de type IV ou maladie de Morquio	50
3.4.5	Mucopolysaccharidose de type VI ou Maladie de Maroteaux-Lamy	51
3.4.6	Mucopolysaccharidose de type VII ou maladie de Sly	53
3.5	La maladie de Fabry	53
3.5.1	Epidémiologie et génétique	53
3.5.2	Mécanismes physiopathologiques.....	55
3.5.3	Présentation clinique	58
3.5.4	Diagnostic de la maladie de Fabry	61
3.5.5	Prise en charge thérapeutique de la maladie de Fabry	63
IV.	La médecine de Précision	64
4.1	Définition	64
4.2	Les supports de l'information biologique	64
4.2.1	La génomique.....	65
4.2.2	La transcriptomique	66
4.2.3	La protéomique.....	69
4.2.4	La métabolomique	72
4.2.5	Traitement des données.....	82
4.3	Analyse des données.....	84
4.3.1	Analyse univariée	84
4.3.2	Analyse multivariée.....	84
4.3.3	Validation des modèles.....	86
4.3.4	Evaluation des variables discriminantes	87
	OBJECTIFS DE LA THESE	89
	TRAVAIL EXPERIMENTAL.....	92
I.	Présentation de l'étude METALYS	93
1.1	Objectifs	93
1.2	Analyse métabolomique des mucopolysaccharidoses par spectrométrie de masse haute résolution.....	93
1.2.1	Collecte des échantillons	93
1.2.2	Aspects éthiques de l'étude.....	94
1.2.3	Partenaires	94
1.2.4	Résultats et publications.....	94
1.2.5	Conclusion.....	95

1.3	Application des sciences « omiques » dans la maladie de Fabry : Projet METALYS - étude Fabomics.....	205
1.3.1	Objectifs de l'étude.....	205
1.3.2	Collecte des échantillons	205
1.3.3	Partenaires.....	205
1.3.4	Résultats.....	206
II.	Partie 2 : modulation de l'autophagie dans la maladie de Fabry	207
2.1	Objectifs de l'étude.....	207
2.2	Partenaires.....	208
2.3	Résultats.....	208
	DISCUSSION ET PERSPECTIVES.....	210
I.	Médecine de précision et maladies héréditaires du métabolisme	211
II.	Etude de l'adressage de l'enzyme au lysosome et modulation de l'autophagie dans la maladie de Fabry.....	213
III.	Conclusion.....	215
	BIBLIOGRAPHIE.....	216

LISTE DES TABLEAUX

Tableau 1 : Exemples de protéines de la membrane lysosomale et leur implication en physiopathologie	23
Tableau 2 : Dérégulation de l'autophagie et maladies lysosomales	38
Tableau 3 : Classification des maladies lysosomales selon la nature des molécules de surcharge	40
Tableau 4 : Enzyme déficiente, substrat accumulé et gène en cause dans les MPS	46

LISTE DES FIGURES

Figure 1 : Premières images en microscopie électronique du lysosome au sein de la cellule sur une coupe de foie. Reproduction de Novikoff (1956)	17
Figure 2 : Formation du résidu mannose-6-phosphate	19
Figure 3 : Trafic de l'enzyme au sein du système endosome lysosome	20
Figure 4 : Les différentes voies autophagiques : la macroautophagie, la microautophagie, l'autophagie médiée par les chaperons	25
Figure 5 : Les étapes du processus de macroautophagie	27
Figure 6 : Etape de maturation de la macroautophagie	30
Figure 7 : Régulation de la macro autophagie par le complexe mTORC1	33
Figure 8 : Régulation de l'autophagie par le complexe PI3K	34
Figure 9 : Principes généraux de traitement dans les maladies héréditaires du métabolisme	42
Figure 10 : Structure tridimensionnelle de l' α -Galactosidase A	54
Figure 11 : Angiokératomes chez un patient atteint de la maladie de Fabry	60
Figure 12 : Cornée verticillée chez un patient atteint de la maladie de Fabry	60
Figure 13 : Les différents supports de l'information biologique et les sciences « omiques » correspondantes qui permettent leur exploration	65
Figure 14 : Les différentes techniques utilisables en protéomique	70
Figure 15 : Analyse multi array MSD®	72
Figure 16 : Préparation des échantillons pour les études de métabolomique	75
Figure 17 : Représentation schématique d'un spectre de masse	77
Figure 18 : Principe de l'ionisation par électrospray	78
Figure 19 : Piège à ions injecteur et piège à ions électrostatique ORBITRAP®, Illustration Thermo Fisher scientific	81
Figure 20 : Représentation simplifiée d'un spectromètre de masse en tandem	82
Figure 21 : Exemple de Coordonnées factorielles (Score plot) et contributions factorielles (loading plot) obtenues à partir d'une matrice de données	85
Figure 22 : Représentation PLSDA montrant la séparation de deux groupes (groupe 1 et groupe 2)	86
Figure 23 : Répartition du nombre d'échantillons d'urines selon le type de mucopolysaccharidose	93

RESUME

Exploration des mécanismes physiopathologiques des mucopolysaccharidoses et de la maladie de fabry par approches « omiques » et modulation de l'autophagie

Résumé Les pathologies lysosomales sont des maladies liées au déficit quantitatif ou qualitatif d'une hydrolase ou d'un transporteur à l'origine d'une atteinte multiviscérale potentiellement sévère. Certaines de ces pathologies sont accessibles à des traitements mais ces thérapeutiques sont uniquement symptomatiques et ne guérissent pas les patients. Même si le phénomène de surcharge peut expliquer entre autres la symptomatologie observée, la physiopathologie de ces maladies est complexe et non précisément connue. Une meilleure connaissance de ces pathologies pourrait permettre d'améliorer leur prise en charge globale. L'objectif de ce travail était dans un premier temps d'appliquer des techniques « omiques » dans deux groupes de maladies : les mucopolysaccharidoses et la maladie de Fabry. Cette étude a permis la mise en place d'une méthodologie métabolomique non ciblée basée sur une stratégie analytique multidimensionnelle comportant la spectrométrie de masse à haute résolution couplée à la chromatographie liquide ultra-haute performance et la mobilité ionique. Dans les mucopolysaccharidoses, l'étude des voies métaboliques a mis en évidence des modifications dans le métabolisme de plusieurs acides aminés et du système oxydatif du glutathion. Dans la maladie de Fabry, des modifications ont été observées dans l'expression de l'interleukine 7 et du facteur de croissance FGF2. La deuxième partie du travail s'est intéressée à la modulation de l'autophagie dans la maladie de Fabry. Notre étude a montré une diminution du flux autophagique avec un retard d'adressage de l'enzyme au lysosome dans les cellules Fabry. L'inhibition de l'autophagie permet de diminuer l'accumulation du substrat accumulé (Gb3) et améliore l'efficacité de l'enzymothérapie substitutive. En conclusion ce travail a permis une meilleure compréhension des mécanismes physiopathologiques des pathologies lysosomales et a montré la complexité du fonctionnement du lysosome. Ces données permettent d'espérer l'amélioration des stratégies thérapeutiques et diagnostiques dans ces maladies.

Mots clés : pathologies lysosomales, sciences omiques, biomarqueurs, autophagie

Mucopolysaccharidosis and Fabry disease investigation using omics approaches and autophagy modulation

Summary : Lysosomal diseases caused by quantitative or qualitative hydrolase or transporter defect induce multiorgan features. Some specific symptomatic treatments are available but they do not cure patients. Pathophysiological bases of lysosomal disease are poorly understood and cannot be due to storage only. A better knowledge of these pathologies could improve their management. The first aim of this study was to apply “omics” strategies in mucopolysaccharidosis and in Fabry disease. This thesis allowed the implementation of an untargeted metabolomic methodology based on a multidimensional analytical strategy including high-resolution mass spectrometry coupled with ultra-high-performance liquid chromatography and ion mobility. Analysis of metabolic pathways showed a major remodeling of the amino acid metabolisms as well as oxidative stress via glutathione metabolism. In Fabry disease, changes were observed in expression of interleukin 7 and FGF2. The second study focused on modulation of autophagy in Fabry disease. In this work, we have shown a disruption of the autophagic process and a delay in enzyme targeting to the lysosome in Fabry disease. Autophagic inhibition reduced accumulation of accumulated substrate (Gb3) and improved the efficiency of enzyme replacement therapy. This work allowed a better knowledge of the physiopathological mechanisms implicated in lysosomal diseases and showed the complexity of lysosome. These data could ameliorate management of these disease and are associated with hope for patients.

Key words : lysosomal diseases, omics techniques, biomarkers, autophagy

INTRODUCTION

I. Le système endosome lysosome

1.1 Le lysosome

Le lysosome est une organelle intra cellulaire ayant un rôle de dégradation, de recyclage et de renouvellement des composants de la cellule. Il permet le maintien de l'homéostasie cellulaire en apportant à la cellule les monomères dont elle a besoin pour fonctionner et en dégradant des biomolécules ou des organelles non fonctionnelles. Le lysosome est impliqué dans une des grandes fonctions intra cellulaires : l'autophagie.

1.1.1 Découverte du lysosome

Entre 1961 et 1965, Christian de Duve étudia la répartition de l'activité phosphatasique acide sur des broyats d'hépatocytes. Certaines fractions de la cellule contenaient une très grande quantité de phosphatases acides. Le concept initial de lysosome était né et fut d'abord un concept biochimique (1).

En 1962, Novikoff et Essner décrivent les lysosomes en microscopie électronique sous forme de corpuscules de la cellule eucaryote limités par une membrane et ayant une forte activité phosphatasique acide (2) (**Figure 1**).

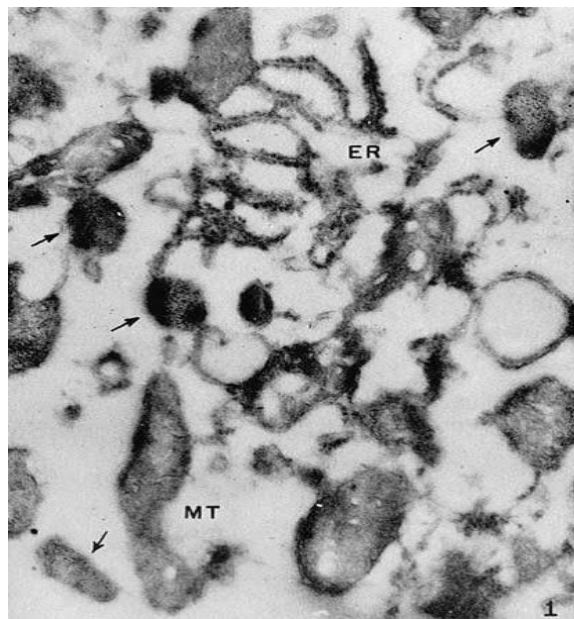


FIGURE 1 : PREMIERES IMAGES EN MICROSCOPIE ELECTRONIQUE DU LYSOSOME (FLECHES NOIRES) AU SEIN DE LA CELLULE SUR UNE COUPE DE FOIE. MT = MITOCHONDRIE, ER = RETICULUM ENDOPLASMIQUE. REPRODUCTION DE NOVIKOFF (1956).

1.1.2 Description du lysosome

1.1.2.1 A l'échelle cellulaire :

Les lysosomes sont des organelles intra cytoplasmiques sphériques ou ovalaires, de diamètre généralement compris entre 0,1 et 0,2 μm , limitées par une membrane trilamellaire. Ils sont présents dans toutes les cellules eucaryotes sauf dans les globules rouges. Leur contenu est riche en protons et notamment en calcium ionisé (Ca^{++}). Ils ont un pH compris entre 4,5 et 5,5 et différent des endosomes par leur degré d'acidification. Ils ont une densité élevée notamment en microscopie électronique.

Les lysosomes contiennent une cinquantaine d'hydrolases acides différentes qui participent à la dégradation des macromolécules et des organelles endommagées en monomères (3) ; et des protéines activatrices localisées dans la matrice. La membrane lysosomale contient quant à elle plus d'une vingtaine de protéines différentes : protéines de transport, canaux ioniques, pompes à protons...

Les lysosomes peuvent être marqués par le lysotracker, un agent acidotrope. Le lysotracker peut également marquer les endosomes tardifs et les autophagosomes car leur contenu est acide. Le lysosome ne possède pas de récepteurs mannose-6-phosphate ce qui le différencie des endosomes.

1.1.2.2 A l'échelle moléculaire :

1.1.2.2.1 Contenu interne du lysosome

Les hydrolases lysosomales

Modalités de fonctionnement des hydrolases lysosomales

Ces enzymes ont une activité optimale à un pH situé entre 4,5 et 5,5. Elles peuvent dégrader de nombreuses macromolécules : protéines, lipides, acides nucléiques, glucides,... La dénomination des hydrolases a été conçue selon la nature biochimique des substrats à dégrader : protéines - protéases, lipides - lipases, acides nucléiques - nucléases, glucides - glycosidases, ...

Ces enzymes nécessitent parfois pour fonctionner la présence de cofacteurs et de protéines activatrices. Ainsi, l'activité des enzymes impliquées dans la dégradation des sphingolipides, nécessite la présence de « protéines activatrices des sphingolipides » (SAP, *Sphingolipid Activator Protein*) ou saposines. Les quatre saposines, A, B, C et D, codées par le même gène la prosaposine, ne sont pas spécifiques d'une enzyme et sont les cofacteurs de plusieurs enzymes. Le GM2 *activator protein* est un activateur spécifique de la β -hexosaminidase impliquée dans la maladie de Tay-Sachs (4).

Synthèse et adressage des hydrolases lysosomales

Les hydrolases lysosomales sont synthétisées au niveau du réticulum endoplasmique (5) puis sont adressées à l'appareil de Golgi où elles subissent des modifications post-traductionnelles, notamment des réactions de glycosylation et de phosphorylation. La phosphorylation d'un résidu mannose au niveau d'une chaîne glucidique permet l'adressage de l'enzyme au lysosome. Cette phosphorylation est catalysée par deux enzymes. La N-acétylglucosamine-1-phosphotransférase assure le transfert d'un résidu N-acétylglucosamine-1-phosphate sur l'hydroxyle du carbone 6 d'un résidu mannose. La deuxième enzyme, la N-acétylglucosamine-1-phosphodiester α -N-acétylglucosaminidase, assure le clivage du résidu N-acétylglucosamine ce qui démasque le mannose-6-phosphate (M6P) (**Figure 2**).

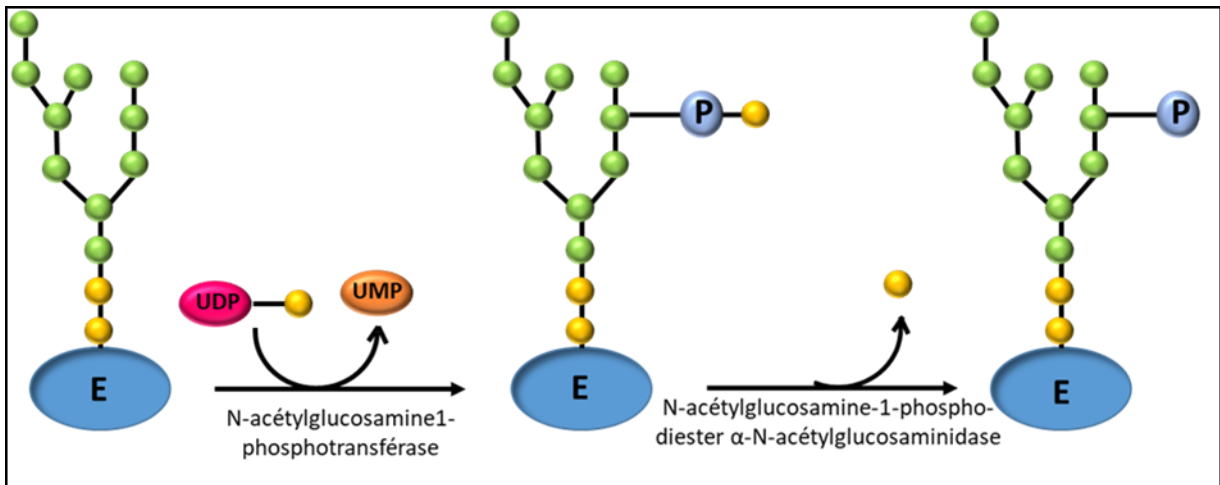


FIGURE 2 : FORMATION DU RESIDU MANNOSE-6-PHOSPHATE. APRES LA GLYCOSYLATION, UNE MAJORITE DES ENZYMES LYSOSOMALES EST PHOSPHORYLEE SUR UN RESIDU MANNOSE D'UNE CHAINE GLUCIDIQUE. LA N-ACÉTYLGLUCOSAMINE-1-PHOSPHOTRANSFERASE ASSURE LE TRANSFERT D'UN RESIDU N-ACÉTYLGLUCOSAMINE-1-PHOSPHATE PUIS LA N-ACÉTYLGLUCOSAMINE-1-PHOSPHO-DIESTER α -N-ACÉTYLGLUCOSAMINIDASE CLIVE LE RESIDU N-ACÉTYLGLUCOSAMINE CE QUI DEMASQUE LE MANNOSE-6-PHOSPHATE.

Le résidu mannose-6-phosphate permet à l'enzyme d'être adressée à l'endosome puis au lysosome via un récepteur mannose-6-phosphate (M6PR) (6). Le lysosome ne possède pas de récepteur mannose-6-phosphate, mais il est capable de fusionner avec les endosomes tardifs et les autophagosomes (5, 7). Il reçoit ainsi son contenu via des mécanismes d'endocytose ou via des voies de biosynthèse après fusion avec l'endosome tardif. La dégradation du matériel endosomal et autophagosomal se déroule au moment de la fusion avec le lysosome aboutissant à la formation des endolysosomes et des autolysosomes. La digestion des substrats endocytés et autophagiques débute dans les endolysosomes et les autolysosomes puis se termine dans le lysosome sous l'action des hydrolases acides. La dissociation entre l'hydrolase et le récepteur mannose-6-phosphate se fait dans

le lysosome en milieu acide. Le récepteur mannose-6-phosphate est ensuite recyclé vers l'appareil de Golgi pour recruter d'autres protéines lysosomales et les adresser au lysosome ou vers la membrane plasmique (**Figure 3**). D'autres mécanismes d'adressage au lysosome indépendants du mannose-6-phosphate existent. Ainsi, les protéines activatrices, cofacteurs des enzymes lysosomales, sont adressées au lysosome via un récepteur appelé sortiline (8). D'autre part, dans la mucopolidose de type II, les hydrolases solubles lysosomales ne possèdent pas de mannose-6-phosphate à cause du déficit en GlcNAc-phosphotransferase. Elles ne peuvent donc pas utiliser la voie de transport utilisant le mannose-6-phosphate (9). Cependant, dans les cellules non mésenchymateuses de ces patients, la majorité des hydrolases lysosomales est quand même adressée au lysosome suggérant l'utilisation de voies autres que celles du mannose-6-phosphate. La phosphatase acide lysosomale et la beta glucocérébrosidase ne possèdent pas non plus de résidus mannose-6-phosphate. La phosphatase acide lysosomale atteint l'endosome par endocytose. Elle contient un motif tyrosine dans son domaine cytosolique qui lui permet d'être endocytée. La bétagluco-cérébrosidase va se fixer au récepteur LIMP-2 qui va l'adresser au lysosome. La dissociation entre LIMP 2 et la glucocérébrosidase est possible grâce à l'acidité intra lysosomale (10).

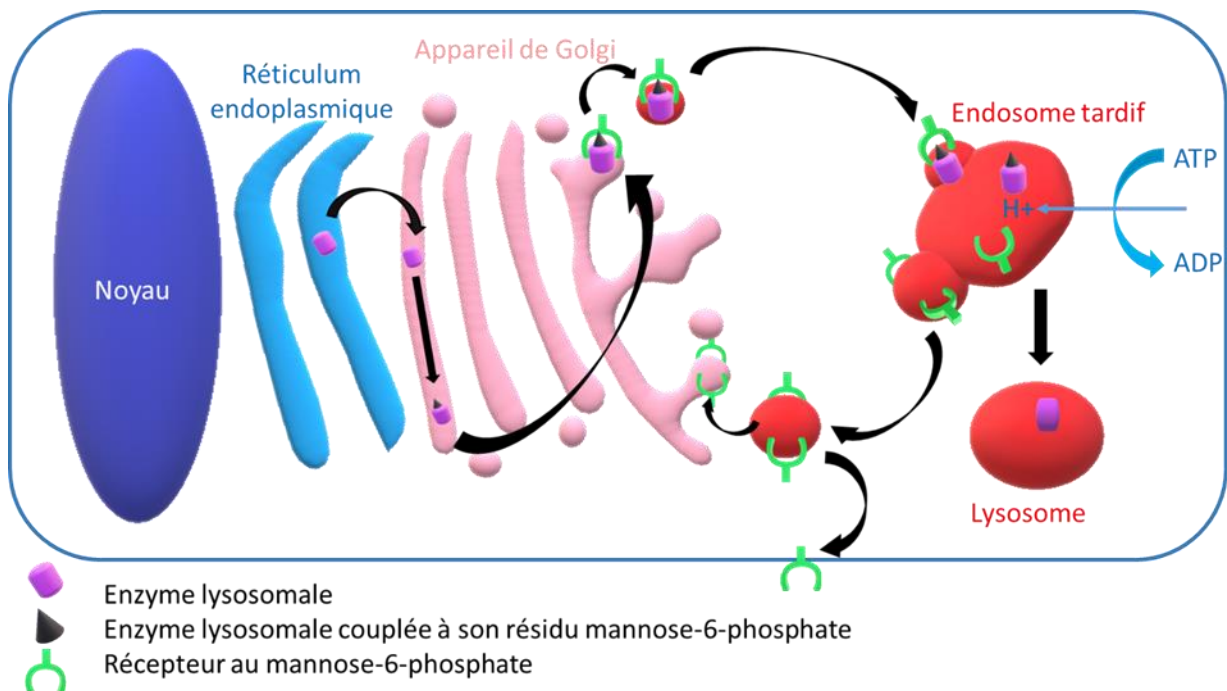


FIGURE 3 : TRAFIC DE L'ENZYME AU SEIN DU SYSTEME ENDOSOME LYSOSOSME. LES ENZYMES LYSOSOMALES SONT SYNTHETISEES DANS LE RETICULUM ENDOPLASMIQUE PUIS SUBISSENT DES MATURATIONS DANS L'APPAREIL DE GOLGI QUI PERMETTENT NOTAMMENT LA FORMATION D'UN RESIDU MANNOSE-6-PHOSPHATE. ELLES SONT PAR LA SUITE ADRESSEES VERS L'ENDOSOME APRES LIAISON A UN RECEPTEUR MANNOSE-6-PHOSPHATE. APRES ACIDIFICATION, LES HYDROLASES DEVIENNENT ACTIVES ET SE DETACHENT DE LEUR RECEPTEUR QUI VA ALORS ETRE RECYCLE POUR RETOURNER VERS L'APPAREIL DE GOLGI. CES RECEPTEURS PEUVENT AUSSI SE RETROUVER AU NIVEAU DE LA MEMBRANE CYTOPLASMIQUE.

Les récepteurs mannose-6-phosphate (M6PR)

Structure du récepteur

Les récepteurs mannose-6-phosphate (M6PR) sont présents dans presque toutes les cellules de mammifères. Deux types de ces récepteurs sont décrits : les récepteurs cation-indépendant CI-M6PR (300 kD) et les récepteurs cation-dépendant CD-M6PR (46 kD). Ces deux récepteurs sont glycosylés et fonctionnent sous forme d'homodimères reliés de manière non covalente (11). Le domaine luminal du récepteur CI-M6PR se compose de 15 unités répétées avec deux sites de liaison au mannose-6-phosphate alors que le récepteur CD-M6PR se compose de 159 acides aminés avec un seul site de liaison au mannose-6-phosphate. Le site de fixation reconnaît spécifiquement les motifs mannose-6-phosphate présents à la surface des précurseurs des hydrolases lysosomales nouvellement synthétisées.

Trafic des M6PR : transport et recyclage

Les récepteurs mannose-6-phosphate ont pour cible une grande majorité des hydrolases lysosomales nouvellement synthétisées à partir de l'appareil de Golgi. Ils sont localisés à l'état basal dans le réseau transgolgien (TGN), la membrane plasmique et l'endosome dans des zones riches en clathrine. Ces zones forment des vésicules de transport par un phénomène de bourgeonnement et ces vésicules se dirigent alors vers l'endosome tardif (**Figure 3**). Les protéines adaptatrices GGA (*Golgi localized gamma ear containing Arf-binding protein*) (12) et AP-1 (*Adaptor protein 1*) favorisent la concentration des M6PR dans les zones riches en clathrine. En cas de localisation du M6PR au niveau de la surface de la cellule, l'ensemble enzyme et récepteur M6PR est internalisé par l'intermédiaire de la protéine adaptatrice AP-2 (*Adaptator protein 2*) et rejoint l'endosome précoce (13).

L'acidité de l'endosome tardif favorise la dissociation entre le récepteur M6PR et son enzyme. Les M6PR sont ensuite réadressés au réseau transgolgien ou à la membrane cytoplasmique pour être recyclés et réutilisés par un système de rétro transport (6).

Rôles du récepteur M6PR

Le rôle majeur des récepteurs M6PR est l'adressage des enzymes lysosomales à l'endosome comme nous l'avons vu précédemment. Un rôle spécifique de chaque M6PR n'a pas été établi, mais certaines hydrolases lysosomales ont une affinité préférentielle pour le CD-M6PR ou pour le CI-M6PR. Le CD-M6PR a une liaison plus forte avec son ligand en présence de cations divalents. Les CI-M6PR permettent l'endocytose des hydrolases lysosomales ce qui n'est pas le cas des CD-M6PR en conditions physiologiques. Le CI-M6PR a également d'autres fonctions qui lui sont propres : régulation des taux circulants d'*Insulin-like growth factor II* notamment au cours du développement

embryonnaire (14); activation du *Transforming Growth factor β* (*TGF β*); activation de l'apoptose médiée par l'acide rétinoïque (15) ; activation de l'apoptose par internalisation du granzyme B (16). Le CI-M6PR est un potentiel agent suppresseur de tumeur du fait de son action sur ces différents facteurs de croissance.

1.1.2.2.2 La membrane lysosomale

Les hydrolases lysosomales sont isolées du reste du cytoplasme grâce à une membrane imperméable. Cette membrane permet de maintenir les hydrolases dans un environnement acide où leur efficacité est maximale et évite leurs éventuels effets indésirables en dehors du lysosome. La membrane est également imperméable aux composés monomériques dégradés qui nécessitent des protéines spécifiques pour leur transport. Les protéines membranaires lysosomales ont des rôles variés assurant le bon fonctionnement du lysosome : import et export de composants, acidification du contenu du lysosome, fusion avec d'autres vacuoles du cytoplasme notamment les endosomes, communication avec les autres organelles,... (3). Le déficit d'une seule de ces protéines peut induire une maladie sévère.

Par des techniques de marquage radioactif des substrats, une vingtaine de protéines de transport a été identifiée assurant le transport des acides aminés, des petites protéines, des sucres, des ions, des nucléotides, et des vitamines. Plus de la moitié des protéines membranaires est représentée par des protéines fortement glycosylées : LAMP1 et LAMP2 (*Lysosome associated membrane protein 1 et 2*) (5); LIMP1/CD63 et LIMP2 (*Lysosome-integral membrane protein 1 et 2*).

LAMP1 et LAMP2 forment une couche de type glycocalyx à l'intérieur de la membrane. Le domaine intra lysosomal de LAMP1 et LAMP2 représente la majeure partie de ces protéines. Elles ont un seul domaine transmembranaire et un domaine C-terminal cytosolique. Leur poids moléculaire se situe entre 40 et 45 kDa avant glycosylation et 120 kDa après glycosylation. La glycosylation leur permet de jouer leur rôle protecteur. Une souris KO pour LAMP1 a un phénotype quasi normal alors qu'une souris KO pour LAMP2 a un phénotype sévère avec une mortalité précoce (17). Chez les souris KO LAMP2, l'analyse tissulaire retrouve de nombreuses vacuoles autophagiques et une hypertrophie cardiaque. En pathologie humaine, la maladie de Danon est liée à des variants délétères du gène *LAMP2*. Cette maladie se caractérise par un retard global des acquisitions, un retard mental, une myopathie et une atteinte cardiaque. Les analyses tissulaires mettent en évidence de nombreuses vacuoles autophagiques musculaires et cardiaques (18, 19).

LIMP2 a un rôle de protection du fait de sa glycosylation, mais constitue également un récepteur de tri et de transport pour la β glucocérébrosidase. Sa surexpression peut induire une augmentation du nombre d'endosomes et de lysosomes. Elle peut perturber le trafic dans le système

endosome/lysosome. Elle contient deux domaines transmembranaires avec des extrémités N et C terminales situées dans le cytosol. En pathologie humaine, une altération du gène *SCARB2* codant pour LIMP2 est à l'origine du syndrome AMRF (*Action myoclonus renal failure*) associant une épilepsie myoclonique progressive et une atteinte rénale (protéinurie avec risque d'évolution vers un syndrome néphrotique) (20, 21).

D'autres protéines membranaires moins abondantes ont pu être identifiées après identification du gène impliqué dans la maladie (**Tableau 1**).

Protéine membranaire	Gène	OMIM	Pathologie associée
LAMP2	LAMP2	300257	Maladie de Danon
LIMP 2	SCARB2	602257	Syndrome AMRF (<i>Action myoclonus renal failure</i>)
Transporteur acide sialique/glucuronique	SLC17A5	269920	Maladie de Salla
Transporteur de la cystine	CTNS	219800	Cystinose
Héparane sulfate α glucosaminide N-acétyltransférase	HGSNAT	252930	Mucopolysaccharidose III C
Transporteur des cobalamines	LMBRD1	277380	Acidémie méthylmalonique avec homocystinurie type cblF
CLN 3	CLN3	204200	Céroïde lipofuscinose type 3
CLN7	CLN7	610951	Céroïde lipofuscinose type 7
Sortiline (récepteur de la prosaposine et de la sphingomyélinase acide)	SORT1	602458	Maladie d'Alzheimer, cancer

TABLEAU 1 : EXEMPLES DE PROTEINES DE LA MEMBRANE LYSOSOMALE ET LEUR IMPLICATION EN PHYSIOPATHOLOGIE.

Les protéines transmembranaires lysosomales possèdent des motifs tyrosine et dileucine. Elles peuvent être ciblées par les protéines adaptatrices telles que les protéines hétérotétramériques AP1 à AP4 (*Adaptor-protein complex*) ou les protéines de la famille GGA (*Golgi localizing gamma adaptating*). Ces deux familles de protéines adaptatrices permettent la ségrégation des protéines lysosomales et l'encapsidation dans des vésicules de clathrine (8).

1.2 L'endosome

1.2.1 Description de l'endosome

Les endosomes sont des organelles de plus grande taille que le lysosome. Deux types d'endosomes distincts sont décrits : l'endosome précoce et l'endosome tardif. L'endosome précoce a un pH modérément acide qui favorise la dissociation entre les ligands et leurs récepteurs. Il reçoit les composés extracellulaires endocytés au niveau de la membrane cytoplasmique. Après endocytose, il réalise un tri des différents composés en distinguant ceux à dégrader et ceux à recycler.

Les éléments à dégrader restent dans l'endosome précoce et sont séquestrés dans des vésicules intraluminales (IVL) de 50 à 80 nm de diamètre qui se forment par bourgeonnement interne. Les

composés à recycler sont dirigés vers l'endosome de recyclage via un système de tubules qui émergent de l'endosome. Ils sont adressés à la membrane ou à l'appareil de Golgi.

L'endosome précoce mature et se transforme progressivement en endosome tardif après une série de phénomènes de fusion et de scission des vésicules aboutissant à un tri protéique. Le nombre de vésicules intraluminales augmente. En parallèle, le pH diminue en lien avec l'activité d'une pompe à proton ATP-dépendante appelée V-ATPase. La dégradation des composés débute dans l'endosome tardif et se termine dans le lysosome.

1.2.2 Maturation de l'endosome

Il est parfois difficile de distinguer endosome précoce et endosome tardif, car les deux entités ne représentent que deux éléments d'un continuum. Néanmoins, certains marqueurs aident à les distinguer. Les endosomes précoces contiennent de la transferrine qui est recyclée et réadressée à la membrane plasmique. La transferrine est donc retrouvée en grande quantité dans l'endosome précoce alors que son taux est faible dans l'endosome tardif. Le nombre de vacuoles intraluminales est corrélé à la concentration de transferrine internalisée. L'endosome précoce contient généralement une à huit vésicules intraluminales et l'endosome tardif en contient au moins neuf. Les marqueurs protéiques EEA1 (*Early Endosomal Antigen 1*) et Rab5 sont spécifiques des endosomes précoces. Rab 11 apparaît progressivement au cours de la maturation de l'endosome précoce. Rab 5 disparaît pour être échangée contre Rab 9 et EEA1 disparaît.

L'endosome tardif évolue progressivement en lysosome par des phénomènes de maturation et de fusion avec des lysosomes préexistants. Les récepteurs au mannose-6-phosphate sont retrouvés dans l'endosome, mais pas dans le lysosome (3, 22).

II. L'autophagie

2.1 Définition

Le terme autophagie vient du grec qui signifie « se manger soi-même ». L'autophagie est un processus biologique des cellules eucaryotes décrit pour la première fois en 1966 (23). Elle aboutit à la dégradation lysosomale d'organelles (mitochondries altérées,...), de complexes protéiques de grande taille (ribosomes,...) et de certains agents infectieux qui compromettent le fonctionnement de la cellule et/ou qui sont à risque d'induire une dégénérescence cellulaire ou une tumeur. L'autophagie permet également le recyclage de certains composants cellulaires. Ce processus est indispensable au maintien de la vie cellulaire. On distingue 3 types d'autophagie : la microautophagie, l'autophagie médiée par les chaperons et la macroautophagie (couramment appelée autophagie). Ces trois voies font intervenir le lysosome pour dégrader le substrat cytoplasmique (**Figure 4**).

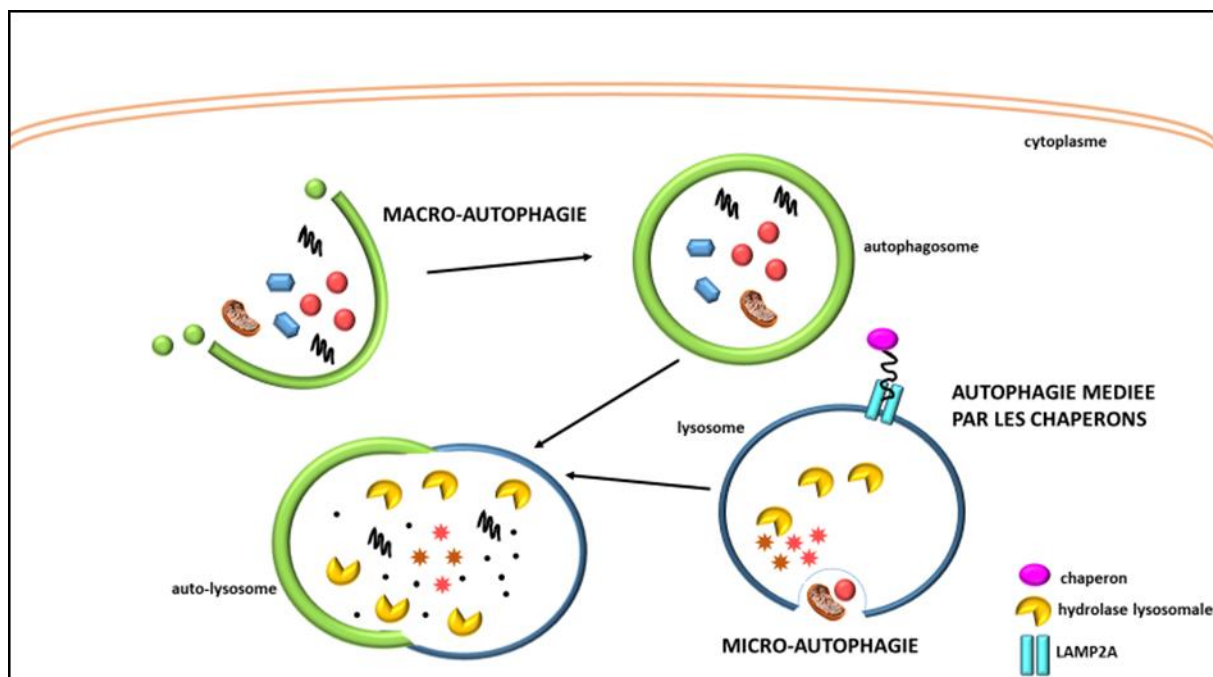


FIGURE 4 : LES DIFFERENTES VOIES AUTOPHAGIQUES : LA MACROAUTOPHAGIE, LA MICROAUTOPHAGIE, L'AUTOPHAGIE MEDIÉE PAR LES CHAPERONS. LES DIFFERENTES FORMES D'AUTOPHAGIE SONT DÉTAILLÉES DANS LES CHAPITRES SUIVANTS.

2.2 Les différentes voies autophagiques

2.2.1 Autophagie médiée par les chaperons (CMA)

Il s'agit d'une forme sélective d'autophagie qui cible les protéines cytosoliques contenant un motif peptidique particulier appelé KFERQ (24). Ce motif est présent chez environ 30 % des protéines

solubles. Ce type d'autophagie est médié par une protéine de membrane associée au lysosome (LAMP2)(25) et par des chaperons (26). Ces molécules chaperons sont des protéines de choc thermique Hsc70 (*Heat shock cognate 70*) encore appelées HSPA8 (*Heat Shock Protein A8*) (27). Elles reconnaissent et se fixent aux protéines solubles destinées à l'autophagie. Lorsque la protéine Hsc70 liée à son substrat cible la protéine membranaire LAMP2, celle-ci permet l'entrée du substrat dans le lysosome. La fixation sur LAMP2 permet en effet à la protéine d'être déroulée et internalisée dans le lysosome afin d'être dégradée par les enzymes lysosomales. Des mutations dans LAMP2A perturbent l'autophagie médiée par les chaperons.

L'autophagie médiée par les chaperons est régulée essentiellement par le niveau protéique en LAMP2A au niveau de la membrane lysosomale (28). Le statut nutritionnel influence également l'activation ou non de cette forme d'autophagie. En cas de carence en nutriments, la macroautophagie est activée en première ligne puis la CMA prend le relais pour fournir à la cellule les protéines indispensables à la survie cellulaire (29). Enfin, le stress oxydatif augmente la CMA par l'augmentation de la transcription de LAMP2 (30). Le vieillissement quant à lui diminue l'activité et la stabilité de LAMP2 à la membrane lysosomale et diminue la CMA (31).

L'autophagie médiée par les chaperons joue un rôle dans la réponse immunitaire en augmentant la présentation des épitopes par les molécules CMH (complexe majeur d'histocompatibilité) par surexpression de LAMP2 (32). Des dysfonctionnements de l'autophagie médiée par les chaperons sont impliqués dans plusieurs pathologies telles que les néphropathies (33, 34), la maladie de Parkinson (35), mais également certaines maladies lysosomales. La maladie de Danon est liée à un déficit en LAMP2 (36). L'autophagie médiée par les chaperons est également altérée dans la mucopolysaccharidose de type 4 où des mutations dans le récepteur mucopolysaccharidose-1 induisent une diminution de LAMP2 et une diminution des substrats adressés au lysosome (37).

2.2.2 Microautophagie

La microautophagie a été décrite pour la première fois en 1966 par De Duve et Wattiaux (23). La microautophagie ne nécessite pas la synthèse de vacuoles autophagiques (38). Le matériel à dégrader est directement adressé au lysosome par des phénomènes de pinocytose. Les constituants cytoplasmiques sont directement endocytés par les invaginations lysosomales. Ce type d'autophagie peut être sélectif ne touchant qu'un seul type de composants à dégrader ou non sélectif. Ainsi, chez la levure, la microautophagie peut dégrader spécifiquement le peroxysome (micropexophagie)(39), le noyau (micronucléophagie)(40) ou la mitochondrie (micromitophagie)(41).

Les mécanismes de régulation de la microautophagie sont similaires à ceux impliqués dans la formation des vésicules intra lumenales retrouvées dans les endosomes tardifs ou les corps

vésiculaires de la macroautophagie (38), mais des gènes spécifiques impliqués dans la régulation de la microautophagie sont également décrits.

Le processus de microautophagie est impliqué dans la survie cellulaire en maintenant l'intégrité des organites et l'homéostasie membranaire. Il a également un rôle dans le développement embryonnaire précoce (42). Néanmoins, la microautophagie chez l'homme reste mal connue. Les mécanismes liant microautophagie et maladies lysosomales sont encore peu connus, mais des liens physiopathologiques existent. Pour exemple, la microautophagie est altérée dans des myoblastes de patients atteints de la maladie de Pompe (43). Dans la maladie de Niemann-Pick de type C, la microautophagie est favorisée par l'augmentation des stérols dans la membrane vacuolaire (44).

2.2.3 La macroautophagie

2.2.3.1 Déroulement de la macroautophagie à l'échelle cellulaire

La macroautophagie est la forme la plus fréquente et la plus connue des types d'autophagie. Elle est impliquée dans la dégradation des protéines à demi-vie longue et des organites altérés. Plusieurs protéines nommées ATG (*Autophagy related genes*) conservées au cours de l'évolution sont nécessaires à l'exécution de l'autophagie (45). Ce processus comporte plusieurs étapes de biogenèse et fusion de plusieurs vésicules (**Figure 5**)(46).

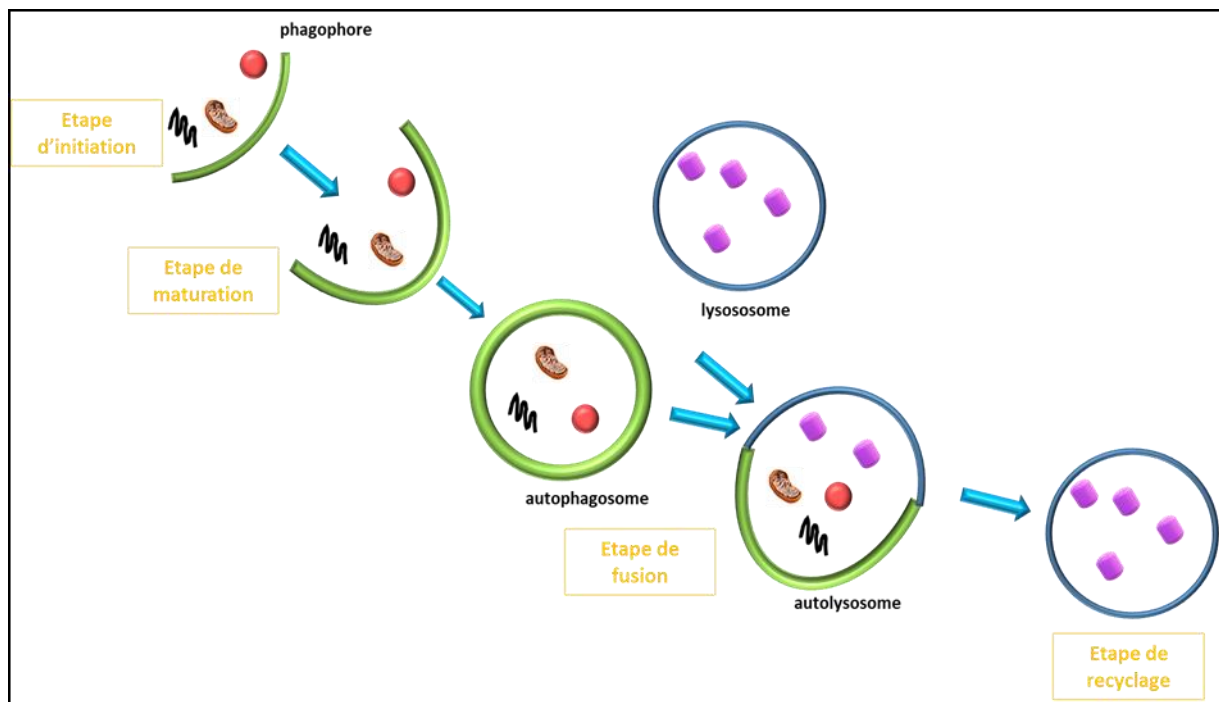


FIGURE 5 : LES ETAPES DU PROCESSUS DE MACROAUTOPHAGIE. LA MACROAUTOPHAGIE DEBUTE UNE PHASE D'INITIATION AVEC LA FORMATION D'UNE MEMBRANE D'ISOLATION QUI ABOUTIT AU PHAGOPHORE. LE PHAGOPHORE SE DEVELOPPE POUR FORMER UNE DOUBLE MEMBRANE QUI S'ALLONGE ET MATURE POUR FORMER L'AUTOPHAGOSOME. L'AUTOPHAGOSOME FUSIONNE AVEC LE LYSOSOME POUR CREER L'AUTOLYSOSOME. LE CONTENU DE L'AUTOLYSOSOME VA ETRE DIGERE PAR LES HYDROLASES LYSOSOMALES. (46)

Etapes d'initiation et d'élongation

La naissance de la première vésicule autophagique, l'autophagosome, débute avec la formation d'une membrane d'isolation appelée également phagophore qui vient entourer les éléments à dégrader. La membrane du phagophore proviendrait de la membrane plasmique, du réticulum endoplasmique, de la mitochondrie ou du compartiment intermédiaire entre le réticulum endoplasmique et l'appareil de Golgi sans que le mécanisme précis de son origine ne soit connu.

Etape de maturation

La membrane d'isolation se développe et forme un système de double membrane prenant alors le nom d'autophagosome. Les éléments à dégrader sont séquestrés dans ce système de double membrane.

Etape de fusion

L'autophagosome fusionne avec l'endosome tardif ou le lysosome pour former un amphisome ou un autolysosome respectivement. Le transport de l'autophagosome vers le lysosome dépend de la dynéine (47). Le lysosome déverse son contenu et notamment les hydrolases acides dans l'autophagosome.

Etape de digestion/recyclage

Le contenu de l'autolysosome va alors être dégradé via les hydrolases lysosomales et/ou recyclé. C'est l'étape de digestion. La macroautophagie permet la dégradation des lipides, des hydrates de carbone, des protéines, des ARN, des mitochondries et des fragments de réticulum endoplasmique. Le contenu du lysosome est ensuite relargué dans le cytoplasme par un système de perméases. Les acides aminés ainsi obtenus sont réutilisés et recyclés.

Reformation du lysosome après le processus d'autophagie

Les endolysosomes et les autolysosomes prolongent les structures tubulaires où les lysosomes et les protéines membranaires lysosomales sont concentrés. A l'extrémité de ces structures tubulaires, des vésicules mûrissent et s'acidifient pour la reformation du lysosome. Ce processus termine chaque cycle d'autophagie permettant la régénération de nouveaux lysosomes denses et matures capables de fusionner avec les endosomes et les autophagosomes (48). Ce processus est parfois altéré dans certaines maladies comme dans les pathologies de stockage de l'acide sialique où les hydrolases sont efficaces, mais où le lysosome est immature avec une densité anormale. Dans la maladie de Niemann-Pick de type C liée à un déficit en NPC2, la biogenèse du lysosome est également déficiente (49). Il peut y avoir un blocage de la maturation au stade autolysosome ou l'accumulation

d'autophagosomes due à un déficit en lysosomes denses et matures. Le processus de biogenèse du lysosome est régulé par mTOR (48).

2.2.3.2 Déroulement de la macroautophagie à l'échelle moléculaire

2.2.3.2.1 Initiation de l'autophagie

Via le complexe ULK

Le complexe ULK est essentiel pour initier l'autophagie. Il est composé par les protéines ULK1, ULK2, ATG13, FIP200 et ATG101 (50). Les protéines ULK1 et ULK2 sont des protéines sérine thréonine kinases. Le complexe ULK est régulé par le complexe mTORC1 (51) et par l'AMPK et va être activé en condition de déplétion nutritionnelle (52). ULK1 et ULK2 vont ainsi être activées en cas de carence de nutriments. Elles peuvent alors s'autophosphoryler et favoriser la phosphorylation de ATG13 et FIP200.

L'activation du complexe ULK peut également recruter un second complexe, le complexe PI3K (Phosphatidylinositol 3-Kinase), au niveau du site d'initiation de l'autophagie [60]. ULK1 induit l'activation de l'autophagie par la phosphorylation de Beclin-1 (53).

Via le complexe PI3K

Un second complexe favorise l'initiation de l'autophagie : le complexe PI3K (PI 3 kinase). Ce complexe comporte Vps34 (*Vacuolar protein sorting 34*), p150 et Beclin1. Ce complexe joue un rôle important dans le trafic vésiculaire et son activité enzymatique permet l'initiation du processus autophagique. Beclin1 a un rôle de liaison entre les processus d'autophagie et d'apoptose, en se liant à Bcl2, protéine anti apoptotique. Beclin1 peut se lier à différentes protéines ce qui modifie la fonction du complexe PI3K. La liaison de Beclin1 à ATG14, UVRAG (*UV irradiation resistance associated gene*) ou Ambra 1 a un effet pro-autophagique (47, 54, 55). La liaison de Beclin1 à Bcl2 (*B-cell lymphoma 2*)(56), Bcl-XL (*Bcl2-associated X protein*) (57) ou Rubicon (58) a un effet anti autophagique (**Figure 8**).

ATG14 joue un rôle important dans l'initiation de l'autophagie, car il favorise le recrutement du complexe au niveau du phagophore (59).

Interactions entre les deux complexes initiateurs de l'autophagie

Le complexe PI3K est retenu au cytosquelette via Ambra1. Par la phosphorylation d'Ambra 1, ULK1 peut libérer PI3K en cas d'activation de l'autophagie. Le complexe PI3K peut alors rejoindre le site d'initiation de l'autophagie où va débiter l'étape de nucléation (47).

2.2.3.2.2 Etape de maturation, formation de l'autophagosome

Cette étape se déroule via le recrutement de protéines ATG sur la membrane vésiculaire (60). Deux systèmes de conjugaison de type ubiquitine interviennent dans la formation de la vésicule.

Le début du processus (formation du phagophore) est sous le contrôle d'un système "ubiquitine-like". La protéine ATG12 est conjuguée de façon irréversible à ATG5 sous l'effet de l'enzyme d'activation E1-like (ATG7) et de l'enzyme de conjugaison E2-like (ATG10). Ainsi, ATG12 est transférée de la protéine ATG7 par l'intermédiaire d'ATG10 pour former une liaison covalente avec ATG5. Le complexe ATG12-ATG5 se lie avec ATG16L1 (61) (**Figure 6**).

Le deuxième système de conjugaison implique la famille de protéines ubiquitine-like ATG8 comprenant la protéine LC3 associée aux microtubules, GABA RAP la protéine associée au récepteur GABA et l'ATPase associée à l'appareil de Golgi (62, 63). La protéine LC3 va être modifiée à son extrémité C terminale par ATG4B et devient LC3-I. Elle va ensuite être conjuguée avec de la phosphatidyléthanolamine par ATG3 et ATG7 pour devenir LC3-II. (**Figure 6**)

Les complexes LC3-II et ATG12-ATG5-ATG16 initient l'autophagie en stimulant la formation du phagophore. Après la formation de l'autophagosome, le complexe ATG12-ATG5-ATG16 se détache de la membrane autophagique alors que LC3-II reste attaché à la membrane interne et peut parfois être dégradé par les hydrolases lysosomales. La protéine LC3-II est impliquée lors des premières étapes de l'autophagie puis dégradée au moment de la phase de digestion.

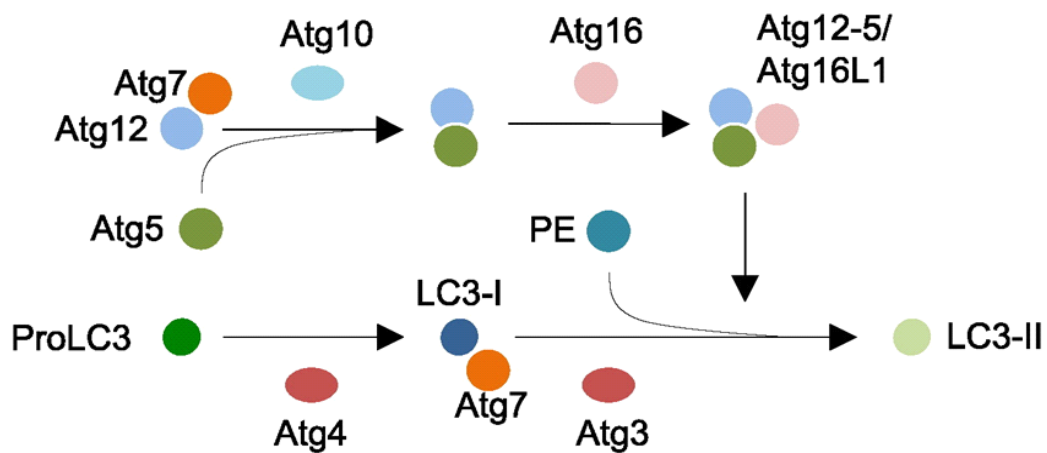


FIGURE 6 : ETAPE DE MATURATION DE LA MACROAUTOPHAGIE. FORMATION DES DEUX SYSTEMES DE CONJUGAISON UBIQUITINE LIKE VISANT LA FORMATION DE L'AUTOPHAGOSOME (64). L'ETAPE D'ELONGATION DE LA MEMBRANE D'ISOLATION FAIT INTERVENIR DEUX SYSTEMES DE TYPE UBIQUITINE. TOUT D'ABORD LE SYSTEME ATG12 QUI IMPLIQUE LES PROTEINES ATG10, ATG7, ATG5 ET ATG16L AFIN DE FORMER LE COMPLEXE ATG12/ATG5/ATG16L. DEUXIEMEMENT, LE SYSTEME LC3-PE QUI SE FORME GRACE A ATG3.

PE : PHOSPHATIDYLETHANOLAMINE

Etape de fusion

La fusion entre l'endosome tardif et l'autophagosome pour former l'amphisome est sous le contrôle de plusieurs molécules régulatrices : Beclin1, Rab7, UVRAG, Vps34, Vps15 et les protéines SNARE. Les protéines SNARE (*Soluble NSF attachment protein receptor*) régulent les processus de fixation et de fusion utilisés au cours de l'autophagie.

La fusion de l'autophagosome avec le lysosome est sous la dépendance de Rab7 (65) et de LAMP2 (17). Le transport des produits dégradés vers le cytosol se fait via un système de perméases dont ATG22.

2.2.3.3 Régulation du processus d'autophagie

L'autophagie est soumise à de nombreux niveaux de régulation. Des dizaines de gènes *ATG* codant pour des protéines ATG intervenant dans le processus l'autophagie ont été identifiés (66). Les protéines ATG sont des protéines kinases indispensables au bon fonctionnement cellulaire au cours du développement et de la vie cellulaire. Pour exemple la différenciation des globules rouges et des neurones dépendent de l'autophagie. L'invalidation de plusieurs gènes *ATG* a pour conséquence une mort *in utero* ou peu après la naissance illustrant le rôle primordial de l'autophagie au cours du développement et de la différenciation cellulaire (67).

Deux processus majeurs de régulation impliquent la protéine mTOR et le complexe Beclin1/PI3K.

2.2.3.3.1 Régulation par la voie mTOR et AMPK

Un des principaux régulateurs de l'autophagie est la protéine mTOR (*mammalian or mechanistic target of rapamycin*) qui est une protéine sérine/thréonine kinase. Elle phosphoryle ou déphosphoryle pour inhiber ou activer l'autophagie. La phosphorylation favorise également la reformation du lysosome en fin de processus d'autophagie (48, 68).

mTOR est impliqué dans deux complexes I et II. mTORC1 est impliqué dans la régulation de l'autophagie. mTORC1 est composé de la protéine kinase mTOR, de la protéine mLST8, de la protéine PRAS40 (*proline rich Akt substrate*), des inhibiteurs de kinase DEPTOR (*DEP domain containing mTOR interacting protein*) et des protéines régulatrices associées à mTOR (*raptor*). mTORC1 intègre un grand nombre de signaux intra et extra cellulaires visant à contrôler la synthèse protéique et à favoriser la croissance cellulaire. Ce complexe peut être recruté via ULK1, via TSC (complexe de la sclérose tubéreuse de Bourneville) ou via une kinase dépendante de l'AMP (AMPK) qui est activée par des taux diminués d'ATP dans la cellule. Ce processus est régulé par les complexes ULK1 et TSC qui favorisent la biogenèse de l'autophagosome.

mTORC1 s'adapte au statut nutritionnel de la cellule. L'activité du complexe mTORC1 est régulée en fonction du taux des nutriments, de l'énergie, de l'oxygène et des facteurs de croissance (69, 70)(Figure 7).

S'il n'y a pas de carence nutritionnelle ou sous l'action des facteurs de croissance, le complexe mTORC1 est activé. Il va réguler négativement le complexe protéique formé par ULK1-ATG13-ATG101-FIP200 en phosphorylant ULK1 et ATG13 ce qui inhibe l'autophagie (8).

En cas de carence nutritionnelle, mTORC1 est inhibé, l'autophagie est activée et l'anabolisme est inhibé. mTOR se dissocie d'ULK1 qui rejoint le site d'initiation de l'autophagosome ce qui permet le recrutement des protéines ATG (69). Ainsi, l'autophagie permet à la cellule de faire face à la situation de carence en favorisant le recyclage des composants intra cellulaires ce qui favorise la survie cellulaire. Dans cette situation, ULK1 et ATG13 sont déphosphorylés. Secondairement, mTORC1 phosphoryle et inactive DAP1 (*Death Associated Protein 1*) qui freine l'autophagie.

Pour fonctionner et être activé, mTORC1 doit reconnaître une séquence d'acides aminés qui permet la translocation du complexe sur la surface cytoplasmique du lysosome ce qui permet un contact de mTORC1 avec une GTPase activatrice appelée Rheb. Rheb active mTOR et par conséquent inhibe l'autophagie. L'AMPK active le complexe ULK, mais peut également inhiber mTORC1 lorsque le niveau d'ATP est bas (71, 72). mTORC1 est une cible pharmacologique qui pourrait être potentiellement utilisée à visée thérapeutique. La rapamycine est un inhibiteur bien connu du complexe mTORC1.

En physiopathologie, l'activité de mTOR est diminuée dans la céréoïde lipofuscinose juvénile (modèle murin) (9), dans la mucopolysaccharidose de type I (6), dans la maladie de Fabry, dans les cellules endothéliales humaines Niemann-Pick C1 ou C2 (10).

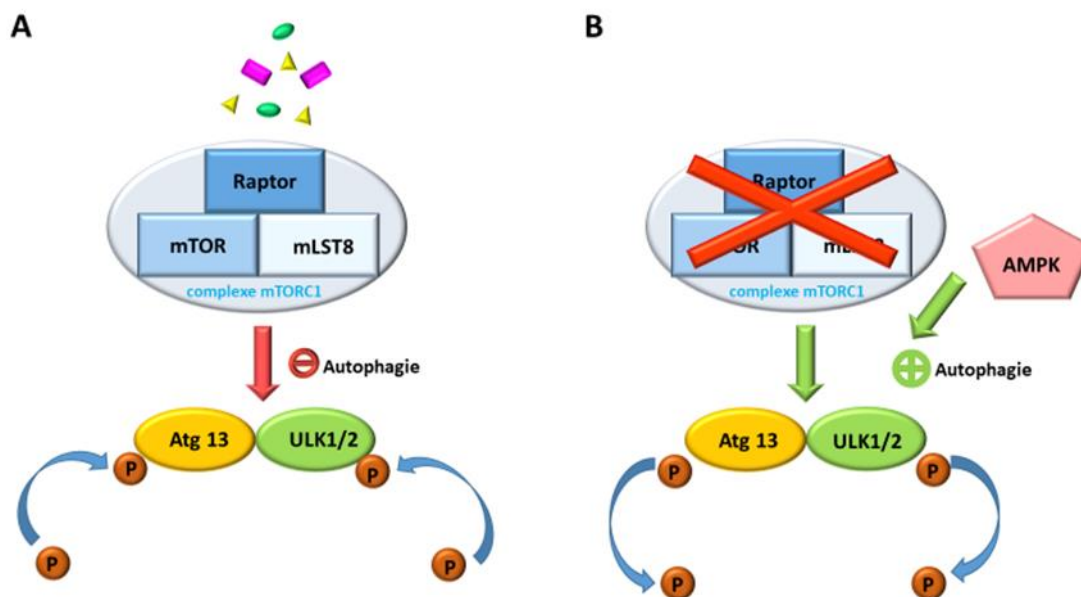


FIGURE 7 : REGULATION DE LA MACRO AUTOPHAGIE PAR LE COMPLEXE mTORC1.

A : EN CONDITION NUTRITIONNELLE RICHE : LE COMPLEXE ULK EST INHIBE PAR LE COMPLEXE mTORC1 VIA UN PROCESSUS DE PHOSPHORYLATION. L'AUTOPHAGIE EST INHIBE. B : EN CONDITION DE PRIVATION DE NUTRIMENTS. mTORC1 EST INHIBE PERMETTANT LA PHOSPHORYLATION DU COMPLEXE ULK. L'AUTOPHAGIE EST ACTIVEE. L'AMPK PEUT EGALEMENT ACTIVER L'AUTOPHAGIE. mTORC1 : MAMMALIAN TARGET OF RAPAMYCIN COMPLEX 1, AMPK : 5'ADENOSINE MONOPHOSPHATE-ACTIVATED PROTEIN KINASE.

2.2.3.3.2 Régulation via le complexe TSC

Un autre mécanisme est médié par le complexe TSC (complexe de la sclérose tubéreuse de Bourneville) formé de trois sous-unités : TSC1, TBC1D7 et TSC2. Ce complexe est régulé par des facteurs de croissance par l'intermédiaire d'Akt et PI3K. La sous-unité TSC2 agit comme une protéine activatrice GTPase pour Rheb induisant une réaction de déphosphorylation ce qui inhibe sa capacité à activer mTORC1. La carence énergétique induit le recrutement du complexe TSC.

2.2.3.3.3 Régulation indépendante de mTOR

La principale voie régulatrice de l'autophagie indépendante de mTORC1 est médiée par l'inositol intra cellulaire, le calcium, l'AMP cyclique, et la voie de signalisation JNK1/Beclin1/PI2KC3. La PI3 kinase est un important régulateur de la biogenèse de l'autophagosome (**Figure 8**) et de plusieurs cascades de signalisation indépendantes de mTOR : MAPK-ERK1-2, Stat2, CXCR4/GPCR, Akt/Foxo3.

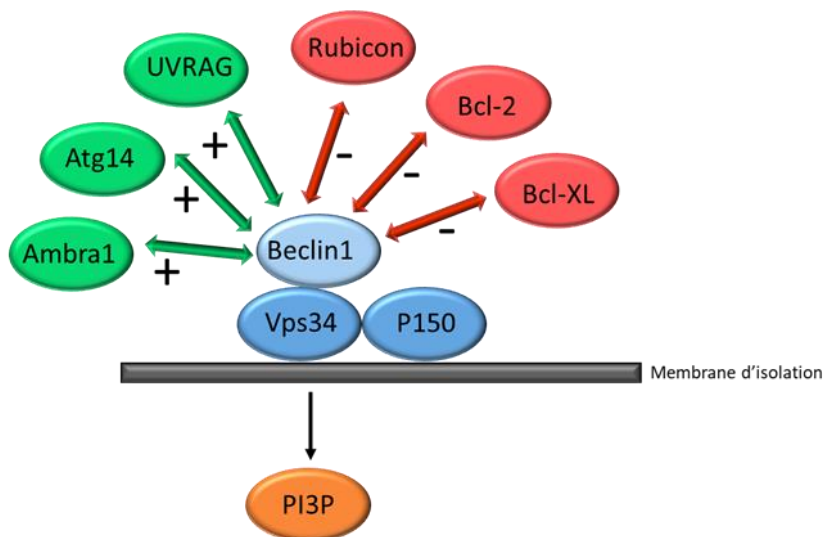


FIGURE 8 : REGULATION DE L'AUTOPHAGIE PAR LE COMPLEXE PI3K. LE COMPLEXE PI3K EST CONSTITUE DE BECLIN1, Vps34 ET P150. EN FONCTION DE LA NATURE DE LA MOLECULE SE LIANT A BECLIN 1, IL Y A ACTIVATION OU INHIBITION DE L'AUTOPHAGIE. LA LIAISON DE BECLIN1 A BCL2, BCL-XL OU RUBICON INHIBE LE COMPLEXE PI3K ET L'INITIATION DE L'AUTOPHAGIE. LA LIAISON A UVRAG, AMBRA 1 OU ATG14 ACTIVE LE COMPLEXE EPI3K ET L'ACTIVATION DE L'AUTOPHAGIE. D'APRES (73)

2.2.3.4 Autophagie et grandes fonctions biologiques

L'autophagie est impliquée dans de nombreux processus physiologiques et/ou physiopathologiques.

2.2.3.4.1 Adaptation de la cellule à son environnement

L'autophagie permet le maintien de l'homéostasie cellulaire en favorisant les adaptations de la cellule à son environnement. Elle agit comme un mécanisme de survie dans les situations de stress en maintenant l'intégrité cellulaire, en régénérant les précurseurs métaboliques et en favorisant l'élimination des débris subcellulaires (74, 75). En cas de carence nutritionnelle notamment en acides aminés, l'autophagie est induite pour que la cellule puisse bénéficier du recyclage des monomères de base indispensables à sa survie. (73)

2.2.3.4.2 Autophagie et développement

L'autophagie est décrite dès les premières étapes du développement embryonnaire juste après la fécondation. Elle est indispensable au bon développement de l'embryon. (76) L'absence de certaines protéines impliquées dans l'autophagie peut être létale ou provoquer des anomalies développementales. Les souris KO Beclin1 (Beclin1^{-/-}) meurent lors des premiers stades de l'embryogenèse. Les souris Beclin1^{+/-} développent précocement des tumeurs (77). L'inactivation du facteur Ambra 1 induit des anomalies du tube neural et une mort embryonnaire précoce. (78) L'autophagie est également indispensable lors des premières heures de vie pour faire face à la carence d'apports de nutriments après la section du cordon ombilical. Les souris KO pour ATG5 et ATG7 meurent d'ailleurs rapidement après la naissance ne pouvant pas faire face à la carence nutritionnelle. (79)

2.2.3.4.3 *Autophagie et vieillissement cellulaire*

Le vieillissement de la cellule est caractérisé par l'accumulation de composants cellulaires défectueux non fonctionnels et d'agrégats ubiquitaires. Il induit également des anomalies de l'ADN. Cela entraîne un stress oxydatif et un dysfonctionnement cellulaire. L'autophagie peut corriger ces dysfonctionnements en dégradant les composants vieillissants et les agrégats ubiquitaires. Elle agit comme un processus de contrôle qualité qui augmente au fur et à mesure du vieillissement cellulaire. Le blocage du processus autophagique diminue la durée de vie de la cellule. (80) Dans les cellules sénescentes, l'autophagie diminue via l'augmentation de l'expression de mTOR. La cellule devient alors moins fonctionnelle. Chez la Drosophile, la suractivation de l'autophagie dans le cerveau allonge la durée de vie en prévenant l'accumulation de protéines oxydées ou ubiquitinylées. (81)

2.2.3.4.4 *L'autophagie en prévention de la mort cellulaire*

Le processus d'autophagie interagit avec les autres mécanismes de mort cellulaire tels que l'apoptose ou la nécrose. Il existe de nombreux liens entre les processus d'autophagie et d'apoptose (82).

Le processus d'autophagie favorise la survie cellulaire mais dans certains cas il peut aboutir à une mort dite autophagique. En effet, certaines cellules présentent des signes de mort cellulaire et contiennent de nombreuses vacuoles autophagiques. Ce processus pourrait s'initier en cas d'inefficacité des mécanismes de survie de la cellule. D'autre part, quand l'apoptose est inhibée l'autophagie pourrait prendre le relais. En cas d'inhibition de la caspase 8, la cellule peut quand même évoluer vers un processus de mort cellulaire sauf en cas d'inhibition de Beclin1 et d'ATG7.(83, 84) Dans un modèle murin *knock-out* pour les gènes Bax et Bak (protéines de la famille Bcl2), l'apoptose est inhibée, mais les cellules peuvent évoluer vers une mort cellulaire non apoptotique associée à la présence d'autolysosomes. Cette mort cellulaire est inhibée en cas d'inhibition de Beclin1 et ATG5.(85) En présence d'activateurs de l'apoptose, la protéine ATG5 peut être clivée par les calpaïnes. Le fragment issu du clivage peut se lier à Bcl-xL dans la mitochondrie ce qui déclenche le relargage du cytochrome C et l'évolution de la cellule vers l'apoptose. Dans les cellules ayant une déplétion en sphingosine 1 phosphate, l'autophagie devrait être activée. Un traitement par Doxorubicine dans ces cellules induit l'apoptose via le clivage d'ATG5 par des calpaïnes (86).

Cependant, le rôle de l'autophagie dans la mort cellulaire n'est pas encore complètement connu. La mort cellulaire pourrait ne pas être liée directement à l'autophagie, mais pourrait s'accompagner d'autophagie (87).

2.2.3.4.5 Autophagie, immunité et infection

Dans l'organisme, deux types d'immunité coexistent : l'immunité innée et l'immunité acquise. L'immunité innée est une réponse immédiate dirigée contre un agent pathogène infectieux. Le système immunitaire détecte le non-soi et active l'inflammation via entre autres le recrutement de cellules immunitaires comme les cellules NLK (*Natural Killer*) et les phagocytes. Les agents infectieux sont phagocytés et ces cellules vont exposer l'antigène à leur surface via les complexes CMH I et CMH II. La protéine LC3 est retrouvée au niveau de la membrane du phagosome au moment des premières étapes de la phagocytose. Ce mécanisme permet l'activation d'une réponse immunitaire acquise spécifique qui garde en mémoire le contact avec l'antigène et qui fait appel aux lymphocytes T et B. Ce système met également en place des mécanismes humoraux comme la sécrétion du complément. La fixation de ligands sur les récepteurs « toll like » active l'autophagie. L'autophagie joue donc un rôle important dans la réponse immune adaptative avec la présentation de l'antigène et le développement des lymphocytes (88).

En cas d'infection, l'autophagie agit en parallèle de la réaction immunitaire via des cellules non immunitaires pour détruire les bactéries, les virus ou les parasites intracellulaires par le processus de xénophagie (89, 90). L'autophagosome contenant l'agent pathogène va fusionner avec le lysosome. La xénophagie nécessite la présence de Rab 7 pour permettre la fusion des vacuoles autophagiques. L'autophagie peut être activée suite à la reconnaissance d'un agent infectieux qui rentre dans la cellule.

Enfin, l'autophagie permet une tolérance immunitaire. L'expression de l'autophagie dans des cellules épithéliales thymiques délivre des protéines endogènes au CMH de classe 2 et contribue à la sélection des LT4.

2.2.3.4.6 Autophagie et maladies neurodégénératives

Les pathologies neurodégénératives sont des maladies caractérisées par la perte progressive des grandes fonctions neurologiques. Elles sont caractérisées par une dysfonction mitochondriale et par l'accumulation d'agrégats protéiques.(91) L'autophagie est dérégulée dans ces maladies. (92) Dans le tissu cérébral de patients atteints par la maladie d'Alzheimer, une accumulation d'autophagosomes a été décrite.(93) L'autophagie pourrait constituer un processus adaptatif de la cellule au processus neurodégénératif. Chez la souris, les anomalies génétiques induisant un déficit de protéines impliquées dans l'autophagie aboutit à une neurodégénérescence liée à l'accumulation d'agrégats protéiques.(94) La stimulation de l'autophagie en inhibant mTOR chez ces animaux peut atténuer les symptômes neurodégénératifs (95, 96).

La maladie de Huntington est associée à une accumulation d'une protéine mutante riche en glutamine : la huntingtine. Cette protéine forme des agrégats intra cytoplasmiques et intra nucléaires dans les neurones des patients. Ces agrégats peuvent être détruits par l'autophagie.

2.2.3.4.7 *Autophagie et obésité*

L'autophagie joue un rôle dans le métabolisme lipidique. L'accumulation de graisses dans l'organisme s'associe à une hypertrophie du tissu adipeux et à une inflammation chronique qui augmente le risque de diabète et de maladie cardio-vasculaire.

Chez la souris obèse ou insulino résistante, l'autophagie est diminuée comme le montre la diminution de l'expression de la protéine ATG7.(25) Une sensibilité plus importante du réticulum endoplasmique au stress est observée ainsi qu'une résistance à l'insuline. Ces effets néfastes sont diminués en activant l'autophagie. Cette activation permet une protection par rapport à l'intolérance glucidique en cas de régime riche en graisses (24). La modulation de l'autophagie dans le cerveau permet la modulation de l'appétit et de la prise alimentaire (97). L'autophagie augmente en cas d'exercice physique dans les cellules cardiaques, pancréatiques, squelettiques et dans le tissu adipeux.

La délétion des gènes codant pour les protéines de l'autophagie induit du stockage de lipides dans le foie et augmente le risque de diabète de type II. Les antidiabétiques oraux du type Metformine sont des activateurs de l'AMPK qui induisent l'autophagie.

2.2.3.4.8 *Autophagie et maladies héréditaires du métabolisme*

Une dérégulation de l'autophagie a été décrite dans plusieurs maladies lysosomales de surcharge **(Tableau 2)**.

Pathologie	Enzyme	Gène	Fonction	Conséquences	Altération de l'autophagie
Maladie de Fabry	Alpha galactosidase A	<i>GLA</i>	Glycoprotéine homodimérique qui hydrolyse les glycolipides/glycoprotéines	Accumulation de Gb3	Inhibition de l'autophagie, retard de maturation des autophagosomes (98)
Maladie de Gaucher	Glucocérébrosidase	<i>GBA</i>	Protéine lysosomale membranaire	Accumulation de glucocérébrosides	Blocage du flux autophagique (99, 100)
Maladie de Pompe (Glycogénose)	Alpha glucosidase	<i>GAA</i>	Clivage du glycogène	Accumulation de glycogène	Accumulation des autophagosomes (101)
Gangliosidose GM1	Beta galactosidase	<i>GLB1</i>	Clivage du beta galactose	Accumulation de gangliosides GM1	Accumulation des autophagosomes, activation de l'autophagie (102)
Mucopolysaccharidoses	Plusieurs enzymes	Plusieurs gènes	Métabolisme des GAG	Accumulation de glycosaminoglycanes	Défaut de fusion autophagosome-lysosome, accumulation des autophagosomes (103, 104)
Mucopolipidose de type IV	Mucopolipine 1	<i>MCOLN1</i>	Régulation de l'exocytose lysosomale	Déficit du canal récepteur ionique	Anomalie de l'autophagie, défaut de recyclage des mitochondries (105)
Céroïde lipofuscinose 3	Battenine	<i>CLN3</i>	Maintien d'un pH normal, fonction cathepsine D	Accumulation de céroïde lipofuscine	Défaut de maturation des vacuoles autophagiques, mitophagie anormale (106, 107)
Niemann Pick type A/B	Sphingomyélinase acide	<i>SMPD1</i>	Conversion de la sphingomyéline en céramide	Accumulation de sphingomyéline	Inefficacité de la clairance autolysosomale (108)
Niemann Pick type C	NPC1 et 2	<i>NPC1/NPC2</i>	Transport du cholestérol de l'endosome tardif au lysosome	Accumulation de cholestérol	Anomalie de formation de l'amphisome (109)
Maladie de Danon	Déficit protéine LAMP2	<i>Xq24</i>	Protéine membrane lysosomale, rôle dans la fusion avec les vacuoles autophagiques		Défaut de maturation des vacuoles autophagiques (110)

TABLEAU 2 DEREGLATION DE L'AUTOPHAGIE ET MALADIES LYSOSOMALES

III. Les maladies lysosomales de surcharge

3.1 Définition

Environ 70 maladies lysosomales de surcharge sont décrites (111, 112). Prises individuellement, ces maladies sont rares, mais elles sont probablement sous diagnostiquées. Leur prévalence combinée est de 1/5000 naissances (113). Elles ont en commun certains symptômes dits de « surcharge » liés à un défaut de dégradation de macromolécules ou de composés monomériques dans les organelles du système endosome/lysosome. Plusieurs classifications ont été utilisées pour distinguer ces différentes maladies. Elles peuvent être groupées en fonction de la nature du matériel qui

s'accumule : mucopolysaccharidoses, lipidoses, glycoprotéinoses, anomalies de stockage du glycogène, céroïdes lipofuscinoses (112) (**Tableau 3**).

Les mécanismes moléculaires pouvant induire une maladie lysosomale de surcharge sont nombreux. Toutes ces pathologies ne sont pas liées au déficit quantitatif ou qualitatif d'une hydrolase acide du lysosome.(3) Des anomalies de la biogenèse du lysosome, des anomalies des protéines de la membrane ou des anomalies du trafic vésiculaire ont également été décrites (114). Certaines maladies sont liées à un défaut d'une enzyme associée à la membrane comme dans la mucopolysaccharidose de type IIIC (115), à des altérations de la translocation ou de l'homéostasie ionique comme dans la mucolipidose de type IV (116).

Néanmoins le mécanisme physiopathologique des maladies lysosomales n'est pas complètement connu. Le déficit enzymatique et l'accumulation de cellules de surcharge ne permettent pas d'expliquer toute la symptomatologie.

Nom de la pathologie	OMIM	Enzyme/Protéine déficiente	Atteinte neurologique
Mucopolysaccharidoses			
MPS I (Hurler, Scheie, Hurler Scheie)	607014, 607015, 607016	α -L-Iduronidase	Oui
MPS II (Hunter)	309900	Iduronate-2-sulfatase	Oui
MPS III A (Sanfilippo A)	252900	Héparane-N-sulfatase	Oui
MPS III B (Sanfilippo B)	252920	α -N-Acétylglucosaminidase	Oui
MPS III C (Sanfilippo C)	252930	AcétylCoA α -glucosaminide N acétyltransférase	Oui
MPS III D (Sanfilippo D)	252940	N-Acétylglucosamine-6-sulfatase	Oui
MPS IV A (Morquio A)	253000	N-Acétylgalactosamine-6-sulfatase	Oui
MPS IV B (Morquio B)	253010	β -Galactosidase	Non
MPS VI (Maroteaux Lamy)	253200	N-Acétylgalactosamine-4-sulfatase	Oui
MPS VII (Sly)	253220	β -Glucuronidase	Oui
MPS IX	601492	Hyaluronidase	Oui
Sphingolipidoses			
Gangliosidose GM1 : forme infantile, juvénile ou adulte	230500	β -Galactosidase	Oui
Gangliosidose GM2 type I (Tay Sachs) : forme infantile, juvénile ou adulte	272750	β -Hexosaminidase A	Oui
Gangliosidose GM2 type II (Sandhoff) : forme infantile, juvénile ou adulte	268800	β -Hexosaminidase A et B	Oui
Gangliosidose GM2 type AB	272800	GM2 activateur	Oui
Gangliosidose GM3	604402	GM3 synthase	Oui
Leucodystrophie métachromatique	250100	Arylsulfatase A	Oui
Maladie de Krabbe	245200	Galactocérébrosidase / prosaposine	Oui
Maladie de Niemann Pick A et B	257200 607625	Sphingomyélinase acide	Oui
Maladie de Gaucher type I	230800	β -glucocérébrosidase	Non
Maladie de Gaucher type II et III	231000, 230900	Activateur saposine C	Oui
Maladie de Fabry	301500	α -Galactosidase A	Oui
Maladie de Farber	228000	Céramidase	Oui
Spingolipidose activator deficiency	176801	Sphingolipid activator	Oui

Nom de la pathologie	OMIM	Enzyme/Protéine déficiente	Atteinte neurologique
Glycogénoses			
Maladie de Pompe/Glycogénose type II	232300	α-glucosidase	Oui
Glycoprotéinoses			
Fucosidose	230000	α-L-Fucosidase	Oui
Aspartylglucosaminurie	208400	Aspartylglucosamidase	Oui
α-Mannosidose type I/II	248500	α-D-Mannosidase	Oui
β-Mannosidose	248510	β-D-Mannosidase	Oui
Sialidose	256550	α-D-Neuraminidase	Oui
Maladie de Schindler type I/II	104170	α-N-Acétalgalactosaminidase	Oui
Galactosialidose	256540	Neuraminidase et β galactosidase	Oui
Déficit en protéines membranaires/transporteurs			
Cystinose	219800	Cystinosine, déficit en transporteur de la cystine	Non
Maladie de Danon	269920/604369	Lysosome-associated membrane protein LAMP2	Oui
Maladie de Salla	604369	Sialine	Oui
Mucopolidose type IV	252650	Mucopolipine 1	Oui
Action myoclonus–renal failure syndrome	254900	Lysosome membrane protein 2	Non
Maladie Niemann Pick type C1 et C2	257220, 607625	NPC1, NPC2	Oui
Lipidoses			
Maladie de Wolman, maladie des esters du cholestérol	278000	Lipase acide lysosomale	Non
Déficit de modifications post traductionnelles			
Déficit multiple en sulfatases	272200	Sulphatase-modifying factor 2	Oui
Mucopolidose type II/type III	252500, 252600	N-acetylglucosaminyl-1 phosphotransférase	Oui
Céroïdes Lipofuscinoses			
Céroïde lipofuscinoses neuronales 1	256730	Palmitoyl protéine thioesterase	Oui
Céroïde lipofuscinoses neuronales 2,	204500	Tripeptidyl peptidase I	Oui
Céroïde lipofuscinoses neuronales 3, juvénile	294200	CLN3P	Oui
Céroïde lipofuscinoses neuronales 4	204300	CSP α	Oui
Céroïde lipofuscinoses neuronales 5	256731	CLN5P	Oui
Céroïde lipofuscinoses neuronales 6	601780	CLN6P	Oui
Céroïde lipofuscinoses neuronales 7	610951	CLN7	Oui
Céroïde lipofuscinoses neuronales 8	600143	CLN8P	Oui
Céroïde lipofuscinoses neuronales 9	609055	Protéine régulatrice de dihydrocéramide synthétase	Oui
Céroïde lipofuscinoses neuronales 10	610127	Cathépsine D	Oui
Céroïde lipofuscinoses neuronales 11	614706	Progranulin	Oui
Céroïde lipofuscinoses neuronales 12	606693	ATPase Cation Transporting 13A2	Oui
Céroïde lipofuscinoses neuronales 13	615362	Cathépsine F	Oui
Céroïde lipofuscinoses neuronales 14	611726	Potassium channel	Oui
Autres			
Hermansky–Pudlak disease type 1	203300	biogenesis of lysosomal organelles complex 3 subunit 1	Non
Hermansky–Pudlak disease type 2	608233	adaptor related protein complex 3 subunit beta 1	Non
Hermansky–Pudlak disease type 3	614072	biogenesis of lysosomal organelles complex 2 subunit 1	Non
Hermansky–Pudlak disease type 4	614073	biogenesis of lysosomal organelles complex 3 subunit 2	Non
Hermansky–Pudlak disease type 5	614074	biogenesis of lysosomal organelles complex 2 subunit 2	Non
Hermansky–Pudlak disease type 6	614075	biogenesis of Lysosomal organelles complex 2 Subunit 3	Non
Hermansky–Pudlak disease type 7	614076	dystrobrevin-binding protein 1	Non
Hermansky–Pudlak disease type 8	614077	biogenesis of lysosomal organelles complex 1 subunit 3	Non
Hermansky–Pudlak disease type 9	614171	biogenesis of lysosomal organelles complex 1 subunit 6	Non
Griscelli syndrome 1	214450	myosin VA	Oui
Griscelli syndrome 2	607624	RAB27A	Oui
Chédiak–Higashi disease	214500	lysosomal trafficking regulator	Oui

TABLEAU 3 : CLASSIFICATION DES MALADIES LYSOSOMALES SELON LA NATURE DES MOLECULES DE SURCHARGE

3.2 Présentation clinique générale

L'âge de début et le type de symptômes dépendent de la fonction de la protéine en cause, de la nature biochimique du matériel accumulé, et du type cellulaire où prédomine la surcharge (117). Certaines maladies lysosomales peuvent être diagnostiquées dès la période anténatale (hydramnios, anomalies osseuses, anasarque,...), mais la plupart sont asymptomatiques à la naissance et les signes cliniques apparaissent progressivement. Certaines pathologies peuvent parfois apparaître tardivement à l'âge adulte avec un phénotype atténué. La forme classique débute généralement dans l'enfance et s'aggrave progressivement avec des symptômes dégénératifs en l'absence de traitement (118).

Les symptômes les plus fréquemment retrouvés quelle que soit la pathologie sont les symptômes neurologiques : troubles du comportement, convulsions, perte des acquisitions psychomotrices après une phase d'acquisition normale, troubles neurosensoriels, troubles neuropsychiatriques, difficultés d'apprentissage scolaires.... Une étude menée au Royaume-Uni a révélé que les maladies lysosomales de surcharge représentaient la grande majorité des diagnostics dans une cohorte pédiatrique de maladies neurodégénératives (39). Les processus neurodégénératifs peuvent toucher plusieurs régions cérébrales : thalamus, cortex, hippocampe et cervelet.

Plusieurs hypothèses physiopathologiques ont été émises pour expliquer l'atteinte neurologique : le matériel de stockage exerce un effet délétère sur les sous-types neuronaux et certaines macromolécules sont synthétisées dans les neurones. L'activation de l'immunité innée est décrite dans le cerveau. Cette inflammation induit des lésions du système nerveux central. Une astrogliose est également retrouvée induisant une toxicité neuronale (119).

3.3 Thérapeutiques disponibles

Il n'existe pas de traitement curatif dans les maladies lysosomales. Les traitements disponibles ont pour but de ralentir ou d'arrêter la progression de la maladie. De nombreux traitements se sont développés sur les dernières années ayant pour cible les différentes organelles impliquées dans la physiopathologie de ces maladies.

Plusieurs stratégies thérapeutiques sont possibles (120, 121) (**Figure 9**) :

- atténuer le déficit enzymatique par l'enzymothérapie substitutive, l'allotransplantation de cellules souches hématopoïétiques, la thérapie pharmacologique de chaperon, la thérapie génique et la thérapie cellulaire.
- diminuer le substrat accumulé en favorisant son élimination et/ou en réduisant sa production.

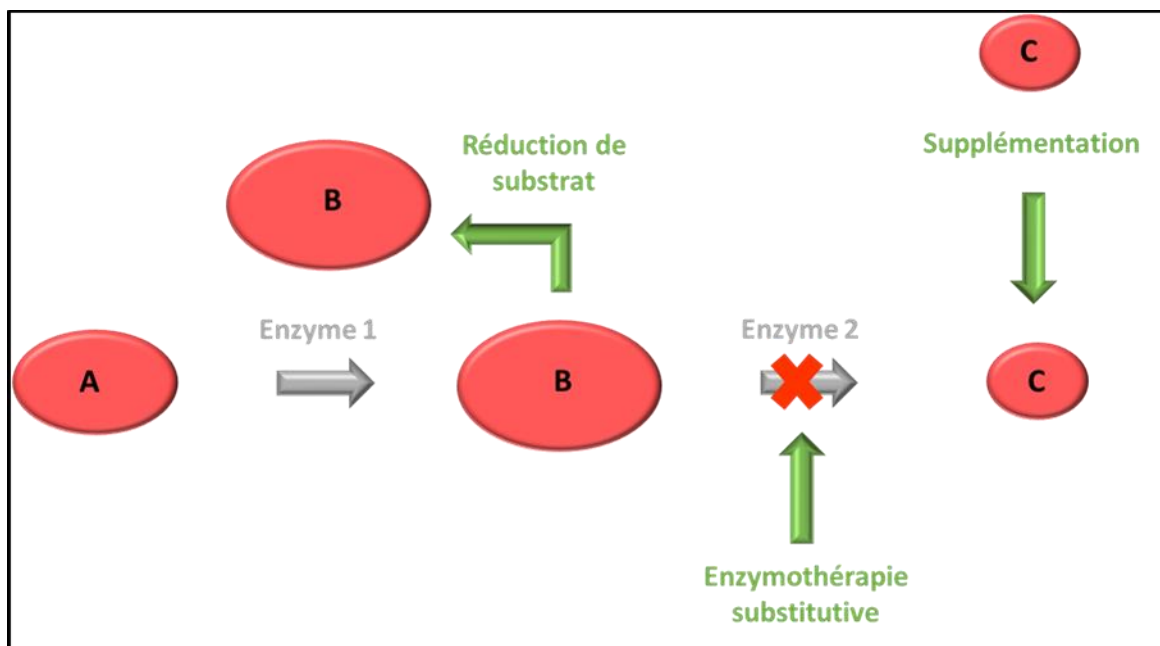


FIGURE 9 : PRINCIPES GÉNÉRAUX DE TRAITEMENT DANS LES MALADIES HÉRÉDITAIRES DU MÉTABOLISME. PAR DÉFICIT DE L'ENZYME 2, LE SUBSTRAT B S'ACCUMULE ET LE PRODUIT C DIMINUE. IL EST THÉORIQUÉMENT POSSIBLE DE SUPPLÉMENTER EN ENZYME 2 OU EN PRODUIT MANQUANT C. IL EST ÉGALEMENT POSSIBLE DE FAVORISER L'ÉLIMINATION DU SUBSTRAT B.

Les différents traitements disponibles sont détaillés ci-dessous. En l'absence de traitement curatif disponible, une part importante de la prise en charge thérapeutique est le traitement symptomatique des différentes complications liées à la maladie.

3.3.1 Enzymothérapie

L'enzymothérapie substitutive consiste à remplacer l'enzyme déficiente par une enzyme exogène le plus souvent par voie intraveineuse à intervalle régulier (hebdomadaire ou toutes les deux semaines). Ce traitement repose sur la possibilité des enzymes lysosomales d'être endocytées par reconnaissance de leurs résidus M6P grâce aux MPR situés au niveau de la membrane plasmique. L'enzyme de synthèse est adressée au lysosome où elle exerce son effet catabolique (122, 123).

La première enzymothérapie substitutive disponible a été développée pour la maladie de Gaucher de type 1. La glucocérébrosidase a été purifiée à partir de placenta humain. Elle est apportée au patient par des perfusions intra veineuse toutes les 2 semaines. Cette molécule a permis de modifier le pronostic des patients (124).

Après le développement de cette première enzymothérapie substitutive, d'autres traitements ont été développés pour la maladie de Fabry, pour la maladie de Pompe et pour plusieurs mucopolysaccharidoses. Elles apportent des bénéfices certains aux patients, mais ne permettent pas

d'améliorer tous les symptômes, car les molécules diffusent mal dans certains organes notamment le système nerveux et l'os. Le traitement ne passe pas la barrière hémato encéphalique (122).

Les principaux effets secondaires de ces traitements sont liés à l'immunogénicité avec le développement d'anticorps dirigés contre l'enzyme de synthèse. D'autre part, ces traitements sont coûteux et nécessitent des perfusions par voie intraveineuses à vie.

D'autres évolutions pour améliorer l'efficacité de l'enzymothérapie substitutive ont vu le jour récemment. En effet, des modifications chimiques ou des administrations intrathécale ou intracérébrale ont été proposées pour augmenter l'accès à des organes cibles (système nerveux central, os..) (111).

3.3.2 Transplantation de cellules souches hématopoïétiques

La transplantation de cellules souches hématopoïétiques (moelle osseuse ou sang de cordon) a été développée pour certaines maladies lysosomales. Les cellules injectées permettent la production de l'enzyme déficiente dans la moelle, mais également dans le système nerveux via les cellules microgliales (125). Elle est indiquée dans la mucopolysaccharidose de type I chez les patients diagnostiqués précocement avant l'âge de 2,5 ans et ayant un quotient de développement >70. Elle peut également être bénéfique dans certains sphingolipidoses telles que la maladie de Krabbe en cas de traitement précoce. En dehors de ces indications, les bénéfices sont très inconstants et la greffe n'est pas recommandée. Les indications peuvent être discutées au cas par cas (126).

3.3.3 Réduction de substrat

La thérapie par réduction de substrat a pour objectif de diminuer la production du substrat et son stockage. Ce traitement peut reposer sur l'inhibition des enzymes situées en amont de l'accumulation. La première molécule ayant été commercialisée est le Miglustat qui est un inhibiteur de la glucosylcéramide synthase (voie de synthèse des gangliosides). Il réduit partiellement la biosynthèse des glycosphingolipides. Le Miglustat est actuellement utilisé dans les maladies de Gaucher (127) et de Niemann-Pick C (128) afin de réduire le substrat disponible pour l'enzyme. Le Miglustat peut traverser la barrière hémato encéphalique. Cette molécule pourrait également avoir des bénéfices dans d'autres gangliosidoses, mais sans indication reconnue à ce jour devant l'absence de preuves scientifiques. Le traitement inhibe les disaccharides ce qui peut induire lors des premières semaines de traitements des effets secondaires digestifs (129).

3.3.4 Les approches variant-dépendantes (chaperons pharmacologiques)

La structure tridimensionnelle des protéines est essentielle au bon fonctionnement de celles-ci. Le repliement des protéines s'effectue au niveau du réticulum endoplasmique lors de leur synthèse. Certaines mutations peuvent conduire à la synthèse de molécules mal repliées. En cas de mauvais

repliement, les protéines sont dégradées par le protéasome grâce à un complexe appelé ERAD (*Endoplasmic Reticulum Associated protein Degradation*) (130).

Des molécules endogènes, dites chaperons ont la propriété d'aider au repliement correct des protéines. C'est le cas des protéines de choc thermique ou encore de la calnexine. Ces molécules synthétiques, dites chaperons pharmacologiques, ont été conçues pour agir comme des inhibiteurs de l'enzyme cible. Elles se lient avec une très forte affinité à la protéine mal repliée permettant ainsi d'aboutir à une conformation tridimensionnelle compatible avec son adressage au lysosome. L'excès de substrat au sein du lysosome et le pH acide de l'endosome tardif contribuent à libérer la molécule chaperonne du site actif. L'enzyme adressée est alors fonctionnelle pour dégrader le substrat accumulé. La protéine échappe, ainsi, à la dégradation par le système ERAD. Des études ont démontré qu'il n'est pas nécessaire de restaurer un taux physiologique intracellulaire d'enzyme active. Une activité enzymatique résiduelle de 10 % de l'activité physiologique peut s'avérer suffisante pour limiter la surcharge.

L'intérêt de ces chaperons pharmacologiques réside dans leur mode d'administration per os évitant, ainsi, la contrainte des perfusions de l'enzymothérapie substitutive. De plus, ces molécules présentent l'avantage d'avoir une meilleure biodistribution vers les tissus difficilement accessibles à l'enzymothérapie. Elles peuvent traverser la barrière hématoencéphalique. Ainsi, des essais menés sur un modèle murin de gangliosidose à GM1 ont démontré que le N-octyl-4-épi- β -valienamine est distribué de manière efficace dans le cerveau. Il permet d'augmenter l'activité de l'enzyme déficiente, de diminuer la surcharge et d'éviter l'aggravation des atteintes neurologiques (131).

Ce type de molécules n'est efficace que si les variants pathogènes induisent la production d'une protéine présentant un défaut de repliement. Seuls les patients présentant des variants dits sensibles pourront être concernés par ces traitements (par exemple, environ 10 à 15% des patients atteints de la maladie de Pompe (132)). Dans la maladie de Fabry, le Migalastat (1-désoxygalactonojirimycine DGJ) a obtenu une autorisation de mise sur le marché en 2017. Environ 30 % des patients présentent des variants sensibles au Migalastat (133, 134). Différents essais cliniques sont actuellement en cours dans les maladies de Pompe et de Gaucher. L'association de ces molécules à l'enzymothérapie a été proposée. Différentes études ont démontré une stabilisation par ces molécules de la β -Glucocérébrosidase, enzyme déficiente dans la maladie de Gaucher. Des essais, effectués sur des fibroblastes de patients atteints de la maladie de Pompe, ont démontré que la co-incubation de l' α -Glucosidase recombinante avec le N-butyl-désoxygalactonojirimycine augmente la stabilité de l'enzyme ainsi que son adressage vers le lysosome in vitro et son activité in vivo et in vitro. Le même type de résultat a été observé sur des fibroblastes de patients Fabry avec le DGJ (134).

3.3.5 Thérapie génique

La thérapie génique consiste à apporter, par l'intermédiaire d'un vecteur viral, un gène déficient au niveau d'un organisme. Ceci permet une production de la protéine codée par le transgène après son internalisation dans les cellules. Dans le cadre des maladies lysosomales, la thérapie génique va permettre la production de l'enzyme déficiente. Elle permet également d'obtenir une concentration en enzyme plus régulière par rapport à l'enzymothérapie. L'enzyme va pouvoir également être captée par les cellules qui n'ont pas intégré le transgène par l'intermédiaire du M6PR. La thérapie génique représente un intérêt pour les maladies lysosomales caractérisées par une symptomatologie cérébrale. En effet, l'introduction du transgène dans le SNC permet une production locale d'enzyme offrant une alternative à l'enzymothérapie dont la distribution dans le SNC est médiocre.

Différentes techniques d'injection sont possibles : soit une injection intracérébrale du vecteur soit une injection intrathécale. De nombreux essais ont été menés sur des modèles animaux et ont permis de prouver l'intérêt potentiel de la thérapie génique dans les maladies lysosomales (135). Des essais de phase I ou II sont planifiés ou en cours de réalisation dans la maladie de Hunter, la maladie de Hurler ou encore dans la maladie de Sanfilippo. Des résultats obtenus chez l'homme ont été publiés en 2008 dans la céréoïde lipofuscinose neuronale infantile tardive par injection intracérébrale du vecteur. Des effets indésirables ont été observés, mais non imputables explicitement au vecteur. Dix-huit mois après l'injection, les résultats semblent indiquer un ralentissement de l'atteinte neurologique par rapport à des sujets témoins (136).

3.4 Les mucopolysaccharidoses

Les mucopolysaccharidoses (MPS) sont des maladies lysosomales de surcharge caractérisées par une atteinte multiviscérale due à l'accumulation de glycosaminoglycanes (GAG) dans les lysosomes.

Les enzymes intervenant dans la dégradation des GAG sont déficitaires dans ces pathologies. Selon l'enzyme déficiente, sept types de mucopolysaccharidoses ont été décrites (**Tableau 4**).

Pathologie/OMIM	Enzyme déficiente	Substrat accumulé	Gène
MPSI / 607014	Alpha-L-iduronidase	DS, HS	<i>IDUA</i>
MPSII / 309900	Iduronate-2-sulfatase	DS, HS	<i>IDS</i>
MPSIIIA / 252900	Héparane-N-sulfatase	HS	<i>SGSH</i>
MPSIIIB / 252920	N-acétyl- α -glucosaminidase	HS	<i>NAGLU</i>
MPSIIIC / 252930	Acétyl-CoA : α -glucosaminidaseNacetyltransférase	HS	<i>HGSNAT</i>
MPSIIID / 252940	N-acétylglucosamine 6-sulfatase	HS	<i>GNS</i>
MPSIVA / 253000	N-acétylgalactosamine-6 sulfatase	KS, CS	<i>GALNS</i>
MPSIVB / 253010	β -galactosidase	KS	<i>GLB1</i>
MPSVI / 253200	Arylsulfatase B (N-acétyl galactosamine 4-sulfatase)	DS,CS	<i>ARSB</i>
MPSVI / 253220	β -glucuronidase	DS, CS, HS	<i>GUSB</i>
MPS IX / 601492	Hyaluronidase	AH	<i>HYAL1</i>

TABLEAU 4 : ENZYME DEFICIENTE, SUBSTRAT ACCUMULE ET GENE EN CAUSE DANS LES MPS. CS : CHONDROÏTINE SULFATE, DS : DERMATANE SULFATE, HS : HEPARANE SULFATE, KS : KERATANE SULFATE, AH : ACIDE HYALURONIQUE

Les mucopolysaccharidoses sont des maladies rares dont l'incidence varie entre 1/25000 et 1/400000 selon le type de maladie. La transmission est le plus souvent autosomique récessive sauf pour la MPS type II qui est liée à l'X. (137) Les symptômes sont progressifs, de sévérité variable secondaires à l'atteinte du système nerveux, du système musculo-squelettique, du foie, de la rate, des yeux et du cœur à des degrés variables selon le type de mucopolysaccharidose (137).

Il n'existe actuellement aucun traitement curatif des mucopolysaccharidoses. Leur traitement est essentiellement symptomatique et leur prise en charge est multidisciplinaire : traitement antalgique, prise en charge des troubles neuropsychiatriques, chirurgie orthopédique pour corriger les déformations articulaires, décompression médullaire, remplacement valvulaire cardiaque, traitement chirurgical du canal carpien, traitement chirurgical des hernies inguinales ou ombilicales, prise en charge rééducative globale (kinésithérapie, psychomotricité,...) (120).

Cependant, la recherche clinique est très active dans le domaine de la thérapeutique des maladies lysosomales et plusieurs traitements spécifiques sont disponibles pour atténuer les symptômes et leur progression.

Plusieurs traitements spécifiques sont utilisés en pratique clinique :

- l'enzymothérapie substitutive pour les MPS I, MPS II, MPS IVA, MPS VI et MPSVII
- la transplantation de cellules souches hématopoïétiques pour les MPS I

De nombreux traitements sont en cours d'étude dans les modèles animaux, mais également chez l'humain.

3.4.1 La mucopolysaccharidose de type I (maladie de Hurler, maladie de Scheie)

3.4.1.1 Présentation clinique

La mucopolysaccharidose de type I (MPSI), décrite pour la première fois en 1919, est une maladie lysosomale autosomique récessive causée par le déficit d'une enzyme : l'alpha-L-iduronidase (138). L'incidence estimée de cette pathologie est de 1/100 000. Ce déficit enzymatique induit une atteinte multisystémique due à l'accumulation de l'héparane sulfate et du dermatane sulfate.

La maladie est caractérisée par l'apparition progressive d'une organomégalie (hépatomégalie et splénomégalie), de hernies inguinales et ombilicales récidivantes, d'infections ORL (otite moyenne aiguë, sinusite,...) à répétition, d'une surdité de transmission (liée aux otites séro-muqueuses), d'une hypertrophie des amygdales et des adénoïdes, d'opacités cornéennes avec une diminution de l'acuité visuelle, d'une dysmorphie caractéristique avec des traits épaissis, un hirsutisme, et une macroglossie, de complications orthopédiques avec un enraidissement articulaire par des dysostoses multiples, d'un risque de compression médullaire, d'une cardiopathie, de valvulopathies, d'une insuffisance respiratoire, d'un retard moteur et cognitif.

Le phénotype de la MPSI est de sévérité variable selon l'activité enzymatique résiduelle. Dans la forme sévère, dite maladie de Hurler, les patients ne présentent généralement aucun signe clinique à la naissance. Au cours des premiers mois de la vie, le phénotype apparaît progressivement. Le diagnostic est généralement posé entre 4 et 18 mois devant l'atteinte multiviscérale progressive. Les symptômes sont sévères et les patients atteints de la maladie de Hurler décèdent généralement pendant la deuxième ou la troisième décennie en l'absence de traitement adapté (139). Il existe également des formes atténuées de MPSI sans atteinte neurocognitive : maladie de Scheie. Les signes apparaissent après 5 ans. Le diagnostic est généralement posé plus tardivement dans la 2ème décennie. Des formes intermédiaires dites Hurler-Scheie ont été décrites. Il apparaît actuellement que ces différents tableaux cliniques constituent un continuum phénotypique de la forme Hurler à la forme Scheie (139, 140).

3.4.1.2 Diagnostic biologique

Le diagnostic est évoqué devant la symptomatologie décrite ci-dessus. Il est confirmé classiquement par la mise en évidence d'une concentration de glycosaminoglycanes augmentée dans les urines avec essentiellement du dermatane sulfate et de l'héparane sulfate (141) et d'une activité enzymatique de l'alpha-L-iduronidase diminuée dans les leucocytes. L'étude est complétée par l'étude moléculaire du gène *IDUA* qui permet de réaliser une enquête familiale et de guider, le cas échéant, le conseil génétique. La MPSI se transmet sur un mode autosomique récessif. Le gène *IDUA* est localisé sur le chromosome 4p16.3. La maladie affecte les deux sexes.

3.4.1.3 *Prise en charge thérapeutique*

3.4.1.3.1 *Transplantation de cellules souches hématopoïétiques*

La transplantation de cellules souches hématopoïétiques (TCSH) peut être proposée avant l'âge de 2,5 ans chez les jeunes patients avec un niveau cognitif préservé (120, 142). Elle prévient le déclin cognitif, l'apparition des symptômes physiques en lien avec l'accumulation des GAG et le risque de décès. La persistance d'un déficit cognitif et de complications physiques résiduelles malgré la transplantation de cellules souches hématopoïétiques montre le besoin d'un traitement aussi précoce que possible avant l'apparition de complications avancées. Il existe, en effet, un temps de latence avant que la greffe ne soit efficace (125, 143, 144). La transplantation de cellules souches hématopoïétiques est un traitement efficace pour la MPSI, mais ce traitement est associé à taux élevé de mortalité et de morbidité. Une étude observationnelle prospective des patients transplantés pour MPSI est toujours en cours. Le but de cette étude est d'évaluer les avantages de la transplantation sur la survie et le besoin de soutien ventilatoire avant et après la greffe de cellules hématopoïétiques. La proportion de la prise de greffe et l'incidence des complications comme la réaction du greffon par rapport à l'hôte seront évaluées (145).

3.4.1.3.2 *Enzymothérapie substitutive*

L'atteinte multisystémique dans les MPSI est traitée par une enzymothérapie substitutive par laronidase (enzyme humaine recombinante alpha-L-iduronidase). (146, 147). Ce traitement est administré par perfusion intra-veineuse hebdomadaire. Il est efficace sur l'atteinte systémique avec une diminution des symptômes ORL et respiratoires mais l'enzyme ne passant pas la barrière hémato-encéphalique, ce traitement n'est pas efficace sur l'atteinte neurologique (148). D'autre part, de fortes doses de traitement sont difficiles à administrer du fait de réactions immunes contre l'enzyme recombinante. L'enzymothérapie substitutive peut être combinée à la transplantation de cellules souches hématopoïétiques.

3.4.2 Mucopolysaccharidose de type II ou maladie de Hunter

3.4.2.1 *Présentation clinique*

La MPSII, également appelée maladie de Hunter, est une maladie transmise par l'X due à la déficience d'une enzyme, l'iduronate-2-sulfatase, induisant l'accumulation d'héparane sulfate et de dermatane sulfate dans les lysosomes.

L'incidence de la maladie de Hunter dans le monde est de 0,30-0,71/100 000 naissances vivantes (149).

La présentation clinique est similaire à celle de la MPSI même si les atteintes orthopédiques et viscérales sont généralement moins marquées. Il s'agit d'une maladie multisystémique progressive

associée à un risque de décès précoce (150). Le diagnostic est habituellement posé entre 18 mois et 4 ans. Il existe cependant deux sous-types de cette maladie selon la sévérité des symptômes : la forme sévère et la forme atténuée ou légère. Dans la forme sévère, l'atteinte neurocognitive est marquée avec une dégradation entre 6 et 10 ans. (150)

3.4.2.2 Diagnostic biologique

Le diagnostic est évoqué devant la présentation clinique. Il est classiquement confirmé par le dosage de l'héparane sulfate et du dermatane sulfate dans les urines, et par le dosage de l'activité enzymatique de l'iduronate-2-sulfatase dans les leucocytes. Enfin, la recherche de variant le gène *IDS* confirme le diagnostic.

3.4.2.3 Prise en charge thérapeutique

L'atteinte multisystémique est traitée par l'administration d'une perfusion intraveineuse hebdomadaire d'Idursulfase (151, 152). Une revue Cochrane évaluant le traitement par l'Idursulfase a inclus 96 patients atteints du syndrome de Hunter âgés de 4,9 à 30,9 ans. Les auteurs ont comparé un traitement avec une perfusion hebdomadaire d'idursulfase à 0,5 mg/kg avec une perfusion hebdomadaire de placebo. Cette revue a montré une amélioration du périmètre de marche et une diminution des GAG urinaires. Cependant, aucun effet bénéfique significatif sur la qualité de vie ni sur la mortalité n'a été observé. De plus, le traitement était sans effet sur les signes neurologiques, car l'enzyme ne traverse pas la barrière hémato-encéphalique (153). Le traitement par l'idursulfase en perfusion intraveineuse hebdomadaire est actuellement recommandé à la dose de 0,5 mg/kg chez les patients atteints de MPSII. (142, 154)

La transplantation hématopoïétique est un traitement efficace des MPSI, mais les résultats pour les MPSII sont variables.(155) La morbidité est élevée avec un léger effet sur les troubles neurocognitifs. Ce traitement n'est pas recommandé dans cette maladie. (120, 142, 156)

3.4.3 Mucopolysaccharidose de type III ou maladie de Sanfilippo

3.4.3.1 Présentation clinique

Cette maladie a été décrite pour la première fois dans les années 1960. Elle est liée à la carence de l'une des enzymes de la dégradation de l'héparane sulfate avec 4 sous-types de MPSIII selon le type d'enzyme déficiente (157). **(Tableau 3)**

Les caractéristiques cliniques sont essentiellement neurologiques avec des troubles psychomoteurs ou comportementaux au premier plan. L'évolution est généralement rapidement progressive. Elle débute par des pleurs inexplicables chez les nourrissons associés à une stagnation des acquisitions psychomotrices après une première phase de développement normal. L'enfant présente alors un retard du langage, des troubles comportementaux et des troubles du sommeil (158). L'atteinte

viscérale est moins importante que dans les MPSI et II et survient plus tardivement. Néanmoins, les patients présentent des signes de surcharge : macrocéphalie, faciès grossier, hépatomégalie, splénomégalie, infections oto-rhino-laryngologiques fréquentes, dysostose légère. Le pronostic est sombre avec la mort précoce dans la deuxième décennie dans les formes sévères (159, 160).

3.4.3.2 Diagnostic biologique

La maladie de Sanfilippo est caractérisée par une augmentation de l'excrétion urinaire d'héparane sulfate (HS). Le sous type de maladie est précisé grâce au dosage enzymatique dans les leucocytes ou les fibroblastes (159). Pour les types IIIA et IIID, la mesure de l'activité d'une autre sulfatase doit être réalisée pour exclure un déficit multiple en sulfatases. Les MPSIII sont transmises sur un mode autosomique récessif avec 4 gènes distincts identifiés (157). Le **Tableau 3** reporte le nom de l'enzyme déficiente, du substrat accumulé et du gène en cause pour chaque type de MPSIII.

3.4.3.3 Prise en charge thérapeutique.

Aucun traitement efficace n'est disponible actuellement pour les MPSIII. Le traitement est uniquement symptomatique avec un pronostic sombre (161). Des essais cliniques basés sur la thérapie génique sont en cours.

3.4.4 Mucopolysaccharidose de type IV ou maladie de Morquio

3.4.4.1 Présentation clinique

Les premiers symptômes de la maladie de Morquio ont été décrits en 1929. (162) Cette maladie inclut deux sous types : type A (Morquio A) et type B (Morquio B). La maladie de Morquio A est due à un déficit en N-acétylgalactosamine-6-sulfatase (gène *GALNS*). La N-acétylgalactosamine-6-sulfatase permet la dégradation du kératane sulfate et du chondroïtine-6-sulfate. (163) Le kératane sulfate s'accumule alors dans les yeux et dans le cartilage. La maladie de Morquio B est due à un déficit en bêta-D-galactosidase qui hydrolyse des résidus de galactose des dissaccharides composant le kératane sulfate. (164)

Comme pour les autres MPS, la sévérité des symptômes est variable et plusieurs formes cliniques ont été décrites allant du phénotype sévère au phénotype atténué.

La prévalence de la maladie de Morquio A est de 1/250 000 (165). C'est une maladie multisystémique progressive de sévérité variable. Certaines formes sont sévères et il existe parfois des signes anténataux (anasarque). Dans les formes intermédiaires, les signes cliniques sont progressifs et apparaissent entre 1 et 5 ans. Les déformations osseuses sont généralement les premiers symptômes : pectus carinatum, genu valgum, cyphose, scoliose, platyspondylie, déformation des os longs. Les patients présentent un retard statural avec une petite taille finale et une dysplasie spondylo-épiphysaire. L'une des complications graves est la compression de la moelle épinière par hypoplasie

de l'odontoïde, hyperlaxité des ligaments avec instabilité des deux premières vertèbres cervicales, et dépôt de GAG dans l'espace extradural antérieur. Il existe une hyperlaxité des poignets contrastant avec une limitation des articulations proximales. Des symptômes extra osseux apparaissent plus tard : surdité, valvulopathie, opacification de la cornée, infections oto-rhino-laryngologiques, syndrome pulmonaire restrictif. Le statut cognitif est normal.(163, 165)

La prévalence de la maladie de type B Morquio est plus rare. Les anomalies squelettiques prédominent avec la déformation spinale et le retard de croissance. Il n'y a pas de déficience neurodégénérative. (166)

Les symptômes osseux sont progressifs et irréversibles, sauf si le traitement est commencé tôt avant les premiers signes.

3.4.4.2 Diagnostic biologique

La sécrétion urinaire de kératane sulfate est inconstante, mais peut orienter le diagnostic. Le diagnostic est confirmé par le dosage de l'activité de la N-acétylgalactosamine-6-sulfatase pour la maladie de Morquio A et par le dosage de l'activité de la bêta-D-galactosidase dans les leucocytes ou les fibroblastes activés pour la maladie de Morquio B. La transmission de la maladie est autosomique récessive. Une étude génétique complète le diagnostic avec la recherche de variants dans les gènes *GALNS* (Morquio A) et *GLB1* (Morquio B).

3.4.4.3 Prise en charge thérapeutique

L'élosulfase alpha est une thérapie de remplacement enzymatique pour le traitement de la maladie de Morquio A depuis 2014. Ce traitement est administré chaque semaine à la dose de 2 mg/kg (167, 168). Les études cliniques de phase III randomisées contrôlées chez les patients de plus de 5 ans atteints ont montré une amélioration du test de marche de 6 minutes, de la fonction respiratoire, et de la croissance staturale avec le traitement par elosulfase alpha versus placebo. (168) Une réduction significative de l'excrétion urinaire du kératane sulfate a également été observée. (169) Les mêmes résultats pour la croissance et la réduction de l'excrétion urinaire kératane sulfate ont été observés chez les patients de moins de 5 ans. (170) Cependant, l'enzymothérapie substitutive dans la maladie de Morquio A est coûteuse et a peu d'effets sur les anomalies squelettiques qui sont les principaux symptômes de cette maladie.

3.4.5 Mucopolysaccharidose de type VI ou Maladie de Maroteaux-Lamy

3.4.5.1 Présentation clinique

La mucopolysaccharidose de type VI ou maladie de Maroteaux Lamy est une pathologie autosomique récessive due à un déficit en arylsulfatase B. Une altération dans le gène *ARSB* induit une accumulation de dermatane de sulfate pathogène dans les cellules (171). C'est une maladie rare dont

la prévalence est difficile à estimer, car elle varie considérablement selon les études épidémiologiques (172-174).

Les caractéristiques cliniques sont variables. La forme sévère se caractérise par des symptômes multi systémiques progressifs avec une dysplasie squelettique prédominante : restriction de la croissance avec petite taille finale, dysostose multiple, scoliose, déformation thoracique et risque de compression de la moelle épinière. La maladie provoque une limitation progressive du périmètre de marche. L'atteinte cognitive est généralement absente. D'autres symptômes sont décrits : une insuffisance respiratoire progressive, une maladie valvulaire cardiaque, une hépatomégalie, une splénomégalie, des sinusites et des otites répétées, une surdité mixte, des apnées du sommeil, une opacification de la cornée, des hernies ombilicales et inguinales, un syndrome du canal carpien. Les patients ont une dysmorphie progressive avec des traits du visage grossiers, un hirsutisme et une macroglossie. Dans les formes sévères de MPS VI, les premiers symptômes apparaissent avant l'âge de 2 ou 3 ans et le risque de décès précoce est élevé. D'autres formes atténuées sont décrites avec des symptômes moins sévères et une durée de vie moins limitée (174).

3.4.5.2 Diagnostic biologique

Le diagnostic est évoqué sur la symptomatologie clinique décrite ci-dessus. Le dosage des GAG met en évidence une accumulation de dermatane sulfate dans les urines. Le diagnostic est confirmé par une activité enzymatique de la β -glucuronidase effondrée dans les leucocytes ou les fibroblastes en culture. La mesure d'une autre sulfatase doit être réalisée pour exclure un déficit multiple en sulfatases. Le bilan est complété par la recherche d'un variant pathogène dans le gène *ARSB*. (174)

3.4.5.3 Prise en charge thérapeutique

Des recommandations pour la prise en charge des patients atteints de maladie de Maroteaux-Lamy ont été publiées en 2007 (175). L'arylsulfatase B recombinante humaine est disponible pour le traitement enzymatique substitutif (Galsulfase, Naglazyme®). La galsulfase ne traverse pas la barrière hématoencéphalique et sa diffusion est faible en cas de mauvaise vascularisation. Ce traitement est peu efficace dans la cornée, le cartilage et le système nerveux central. Des études cliniques de phase 3 ont été réalisées chez l'homme avec Naglazyme® montrant une diminution de l'excrétion urinaire des GAG et une amélioration du périmètre de marche. Le traitement est administré par perfusion hebdomadaire de 1 mg/kg. Des études ont également montré une amélioration de la fonction respiratoire (capacité vitale forcée, volume expiratoire forcé) après plusieurs semaines de traitement par la galsulfase. (176) Enfin, le traitement par galsulfase est efficace sur la croissance staturale et sur le retard pubertaire.(177)

La greffe de moelle osseuse ou de cellules hématopoïétiques a été réalisée avec des effets décevants chez des patients MPS VI. La transplantation rétablit l'activité enzymatique et réduit l'excrétion urinaire des GAG, améliore la fonction cardiovasculaire, la mobilité articulaire, l'acuité visuelle et réduit l'hépatomégalie et la splénomégalie. (178, 179). Elle n'est pas efficace sur les atteintes musculosquelettiques et sur les opacités cornéennes. Ce traitement invasif est associé à des risques. Il n'est pas recommandé en pratique clinique.

3.4.6 Mucopolysaccharidose de type VII ou maladie de Sly

3.4.6.1 Présentation clinique

Cette maladie est liée à un déficit de l'activité de l'enzyme β -glucuronidase entraînant une accumulation de chondroïtine, dermatane et héparane sulfate dans les tissus. Sa transmission est autosomique récessive avec des variations pathogènes retrouvées dans le gène *GUSB*. La MPSVII est rare avec une incidence globale estimée à 1/250 000 naissances.

Plusieurs formes sont décrites. Les plus sévères peuvent être diagnostiquées en anténatal avec un tableau d'anasarque fœto-placentaire. La MPS VII peut se révéler plus tardivement avec des patients sans atteinte cognitive. Les atteintes sont multisystémiques comme dans les autres mucopolysaccharidoses avec un retard mental, une atteinte squelettique et une viscéromégalie.

3.4.6.2 Diagnostic biologique

Le diagnostic repose sur le dosage des GAG urinaires qui met en évidence une augmentation de la chondroïtine, du dermatane et du kératane sulfate, et sur le dosage de l'activité de la β -glucuronidase qui est effondrée. La confirmation se fait par l'analyse moléculaire du gène *GUSB*.

3.4.6.3 Prise en charge thérapeutique

Une enzymothérapie substitutive par Recombinant Human Beta-glucuronidase (rhGUS) Ultragenix est disponible.

3.5 La maladie de Fabry

3.5.1 Epidémiologie et génétique

La maladie de Fabry (OMIM 301 500) est une pathologie lysosomale décrite pour la première fois en 1898 par Johannes Fabry (180) et William Anderson (181). L'incidence de la maladie est de 1/40 000 naissances vivantes chez le garçon (182). Une étude récente de dépistage néonatal retrouvait une incidence plus élevée à 1/3100 tous sexes confondus mais des biais d'interprétation sont possibles et ce chiffre pourrait être surestimé (183). Plus de 50 % des femmes seraient asymptomatiques ou sous diagnostiquées. La maladie de Fabry n'a pas de spécificité ethnique.

La transmission est liée à l’X avec des variants pathogènes retrouvés dans le gène *GLA* codant pour l’ α -galactosidase A. Ce gène est situé en Xq22.1. Sa longueur est de 12 kb et il contient 7 exons. Il code pour une protéine contenant 429 acides aminés qui aboutit à une glycoprotéine mature contenant 370 acides aminés (184) (**Figure 10**). Plus de 900 variants ont été décrits dans la *base de données Human Gene Mutation Database* (<http://www.hgmd.cf.ac.uk/ac/gene.php?gene=GLA>) avec différents types de variants : faux ou non-sens, petites et grandes délétions, anomalies de l’épissage, petites et grandes insertions (185). Les patients présentant une symptomatologie sévère avec une atteinte multiviscérale ont une mutation induisant une absence totale de la protéine. Certains variants sont associés à une persistance d’une activité résiduelle de l’enzyme et sont associés à un phénotype atténué. (184)

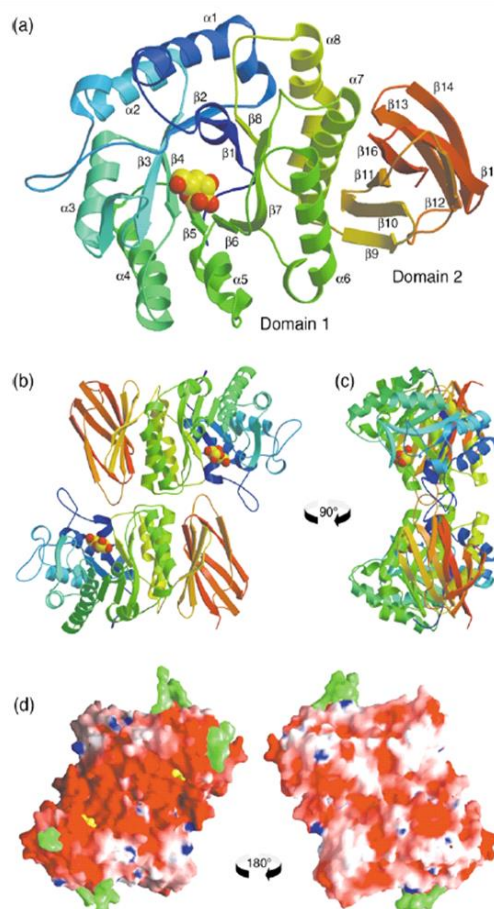


FIGURE 10 : STRUCTURE TRIDIMENSIONNELLE DE L’ A-GALACTOSIDASE A. EN (A) EST REPRESENTEE LA STRUCTURE DU MONOMERE COMPOSE DE DEUX DOMAINES. LE SITE ACTIF EST LOCALISE AU NIVEAU DU DOMAINE 1. EN (B) ET (C) EST REPRESENTE L’HOMODIMERE SOUS DEUX ANGLES DIFFERENTS. D’APRES GARMAN ET GARBOCZI 2004

3.5.2 Mécanismes physiopathologiques

3.5.2.1 Accumulation du globotriaosylcéramide (Gb3)

Les symptômes de la maladie de Fabry seraient liés entre autres à l'accumulation de glycolipides dans les cellules de l'organisme, principalement le globotriaosylcéramide (Gb3). L'alpha galactosidase A catalyse le clivage hydrolytique du galactose terminal du Gb3. Elle est présente dans toutes les cellules de l'organisme sauf les globules rouges. Un déficit en alpha galactosidase A peut donc induire une atteinte multiviscérale. (184)

L'origine du Gb3 lysosomal peut être endogène, mais également exogène dans la circulation sanguine via les LDL (186). Le Gb3 circulant va être capté par les cellules possédant des récepteurs ayant une affinité pour les lipoprotéines. Il est généralement détecté à un stade tardif dans la maladie quand les capacités de captation du Gb3 sont dépassées.

L'accumulation de Gb3 induit une toxicité chronique. Les vaisseaux d'un patient atteint de maladie de Fabry ont une réactivité anormale liée à une dysfonction endothéliale. Des techniques de tomographie par émetteur de positon et par IRM ont montré une hyperperfusion cérébrale. La vasoréactivité cérébrale est allongée (187, 188).

Les effets pathogènes dans la maladie sont liés entre autres à l'accumulation du Gb3, mais la physiopathologie semble plus complexe qu'une simple surcharge et plusieurs grandes fonctions biologiques sont modifiées dans la maladie de Fabry. Ainsi, chez des patients ayant une forme classique de maladie de Fabry, le Gb3 représente seulement 3 % de la masse du cœur. D'autre part, le taux de Gb3 circulant n'est pas corrélé à la sévérité de la maladie (189).

La forme déacylée du Gb3, le lysoGb3 (globotriaosylsphingosine) est également augmentée chez certains patients Fabry. Le lysoGb3 est un inhibiteur de l' α -galactosidase A et de l' α -galactosidase B. In vitro, il augmente la prolifération des cellules musculaires lisses. Le lysoGb3 pourrait donc être à l'origine d'une augmentation de l'épaisseur des parois vasculaires(190). Cependant, aucune corrélation n'est retrouvée entre le taux de lysoGb3, l'âge du patient, la sévérité de la cardiomyopathie hypertrophique chez les patients hémizygotés. Chez les femmes hétérozygotes, en revanche, il existe une corrélation entre le taux de lysoGb3 et l'hypertrophie myocardique (189). Lorsqu'il existe une activité enzymatique résiduelle et une atteinte prédominante au niveau cardiaque, le taux sanguin de lysoGb3 est normal. Les souris KO (activité résiduelle nulle) ont une faible expression phénotypique de la maladie, mais un taux de lysoGb3 élevé.

Le rôle exact du lysoGb3 dans la maladie de Fabry reste donc à définir.

3.5.2.2 Retentissement de la maladie à l'échelle tissulaire

Lors de l'analyse anatomopathologique, certaines lésions sont retrouvées spécifiquement dans la maladie de Fabry alors que d'autres lésions sont secondaires et non spécifiques de la pathologie. Dans tous les types cellulaires, mais plus particulièrement dans les cellules endothéliales, les cellules musculaires lisses, les cellules de l'endocarde et du myocarde, les cellules rénales et les cellules du système nerveux, des inclusions lysosomales ou des dépôts lipido-protéiques sont visualisés en microscopie (191) (192).

La maladie de Fabry est associée à une vasculopathie avec des atteintes artériosclérotiques des vaisseaux qui apparaissent épaissis.(192) La vasculopathie perturbe la fluidité de l'écoulement sanguin et induit un risque augmenté d'accident vasculaire cérébral. Les vaisseaux sont également dilatés avec une augmentation du diamètre de leur lumière surtout dans les artères cérébrales de la fosse postérieure. Les angiokératomes quant à eux sont liés à une dilatation des capillaires du derme. (193)

La cardiomyopathie hypertrophique se constitue progressivement parallèlement à l'accumulation du Gb3. Les valves s'épaississent également avec un risque de fibrose à plus long terme (192). Dans le rein, le mésangium s'élargit progressivement. Une fibrose du glomérule rénal et des tubules est décrite à un stade plus avancé (194). Dans le cerveau, des lésions secondaires aux accidents ischémiques sont observées (195).

3.5.2.3 Autophagie

L'autophagie est altérée dans la maladie de Fabry induisant, entre autres, un blocage de fusion entre autophagosome et lysosome avec une accumulation de vacuoles autophagiques dans les tissus (98, 196). Cette altération serait en lien avec la diminution de l'activité du complexe mTOR secondaire à la diminution de l'activité de l' α -galactosidase A (197). L'altération de ce flux autophagique semble corrélée à la sévérité des symptômes. Ainsi, chez une patiente ayant une symptomatologie marquée, il a été mis en évidence une altération du flux autophagique intra cellulaire alors que le flux autophagique était conservé chez ses deux sœurs ayant une symptomatologie moins marquée (198).

3.5.2.4 Stress oxydatif

Chez le patient Fabry, il existe un déséquilibre entre les mécanismes de défense contre les radicaux libres avec une altération du métabolisme du glutathion et une augmentation de la production des radicaux libres et du NO et ce de manière corrélée à l'accumulation du Gb3 (199). Dans une étude clinique en double aveugle, des taux élevés de 3-nitrotyrosine ont été retrouvés chez les patients Fabry versus patients témoins dans le plasma et dans les vaisseaux du derme et du parenchyme cérébral. L'augmentation de la nitrotyrosine a également été confirmée dans les cardiomyocytes

(200, 201). Les lésions retrouvées dans la maladie de Fabry pourraient donc être dues à la production de radicaux libres. Les premières études sur l'impact de l'enzymothérapie substitutive sur la production des radicaux libres chez les patients Fabry ne retrouvaient pas de bénéfice malgré la diminution de l'accumulation du Gb3. Une bithérapie associant des anti-oxydants à l'enzymothérapie pourrait être bénéfique (199).

3.5.2.5 Inflammation et impact sur la coagulation

L'inflammation est activée en continu dans la maladie de Fabry. Plusieurs études ont mis en évidence une augmentation des leucocytes et des marqueurs inflammatoires (202). Les principales cytokines proinflammatoires sont exprimées en excès chez les patients Fabry (TNF α , Il-1 β) (203). L'immunité innée est activée via des interactions entre le Gb3 et le lysoGb3 et les *Toll like receptors* (TLR) (204). Dans le rein, il a été démontré une augmentation du TGF β et d'autres molécules favorisant la fibrose (205). Le myocarde des patients Fabry est infiltré par de nombreux macrophages (206). Dans le tissu endothélial, une augmentation de l'expression des récepteurs de l'angiotensine 1 et 2, du NF κ B, et des molécules oxydatives associée à une diminution de la synthèse du NO a été décrite (207). Ce phénotype pro-inflammatoire a également été décrit au cours d'autres maladies lysosomales indépendamment du type de substrat accumulé (208). Cette activation chronique de l'inflammation peut induire des phénomènes d'apoptose et de production de radicaux libres compromettant le bon fonctionnement tissulaire.

Dans le sang des patients atteints de la maladie de Fabry, les molécules prothrombotiques, les inhibiteurs de l'activateur du plasminogène, les molécules d'adhésion vasculaires (dont le CD11b) et les sélectines sont augmentées. Le taux de thrombomoduline est diminué. Des anomalies de la fibrinolyse et de l'angiogenèse sont également observées (209). Il n'y a pas eu d'étude fonctionnelle spécifique du clou plaquettaire chez les patients Fabry mais une augmentation de l'agrégation plaquettaire et du risque thromboembolique est décrite dans cette pathologie. L'administration de vitamine E et d'un antiagrégant plaquettaire pourrait améliorer ces anomalies (210). L'insuffisance rénale étant habituellement associée à une augmentation de la réactivité plaquettaire, les patients Fabry ont probablement une dysfonction plaquettaire. Une augmentation significative des marqueurs d'activation plaquettaire (facteur plaquettaire 4 et β thromboglobuline) liée à l'atteinte rénale a d'ailleurs été retrouvée chez des patients Fabry d'une cohorte néerlandaise ce qui conforte cette hypothèse (211).

3.5.2.6 Modifications hormonales

Il a été également montré une augmentation de l'expression des récepteurs aux androgènes dans un modèle murin de maladie de Fabry. Le blocage de ces récepteurs suite à une castration ou via un antagoniste prévient les atteintes cardiaques et rénales chez la souris Fabry (212).

3.5.3 Présentation clinique

3.5.3.1 *Histoire naturelle de la maladie*

Les premiers symptômes apparaissent le plus souvent dans l'enfance, l'adolescence ou chez le jeune adulte. La maladie de Fabry est une maladie progressive avec une atteinte multiviscérale. Les premiers signes à apparaître sont généralement les acroparesthésies survenant périodiquement puis l'hypohydrose, les angiokératomes, les opacités cornéennes et lenticulaires. La maladie est parfois diagnostiquée tardivement à l'occasion d'une complication aiguë comme un accident vasculaire cérébral (AVC). L'espérance de vie des patients est diminuée avec une médiane de survie à 55 ans chez les hommes et à 70 ans chez les femmes. Le risque de décès précoce est lié à l'insuffisance rénale, aux complications cardiaques ou cérébrovasculaires (189).

3.5.3.2 *Symptômes neurologiques et cérébro-vasculaires*

Le déficit en alpha galactosidase A provoque des douleurs neuropathiques avec des acroparesthésies en lien avec une neuropathie des fibres nerveuses de petite taille. L'atteinte neuropathique des petites fibres augmente le seuil de perception de la douleur, du froid et du chaud. La neuropathie épargne les fibres de plus gros calibre. La vitesse de conduction nerveuse et la proprioception sont conservées. (213)

Ces symptômes débutent souvent dans l'enfance et sont présents chez une majorité des patients. Ils augmentent avec l'âge, sont très marqués vers l'âge de 40 ans puis diminuent progressivement (214) (189). Des facteurs déclenchant les douleurs sont fréquemment retrouvés comme le stress, l'exercice physique, les changements de température, l'humidité et la fatigue. Les douleurs sont majoritairement distales, débutant aux pieds puis s'étendant aux mains avec parfois une généralisation secondaire. Elles peuvent durer de quelques minutes à plusieurs jours (215) (216). Des douleurs continues sont également décrites. Des douleurs articulaires sont aussi décrites chez les patients atteints de la maladie de Fabry. Enfin, des douleurs abdominales sévères mimant une appendicite ou une colique néphrétique sont possibles (217).

Une neuropathie auditive avec baisse progressive de l'acuité auditive accompagne souvent les autres symptômes. Sa sévérité est corrélée à l'atteinte des autres nerfs périphériques et du cerveau signant l'avancée de la pathologie. D'autres symptômes ORL sont fréquemment associés tels que des vertiges ou des acouphènes pouvant débuter dès l'enfance (218).

Les patients atteints de la maladie de Fabry ont un risque augmenté d'accident vasculaire cérébral. Il peut survenir à tout âge y compris dans l'enfance. Le risque est également augmenté chez les femmes hétérozygotes (219). La survenue d'un AVC précoce chez un patient sans facteur de risque personnel ou familial doit faire rechercher une maladie de Fabry (220). Les accidents vasculaires

cérébraux sont principalement ischémiques et liés à l'atteinte vasculaire de la maladie. Des embolies d'origine cardiaque sont également décrits (221). L'IRM cérébrale peut révéler des lésions essentiellement de la substance blanche alors que le patient est asymptomatique (222).

3.5.3.3 Atteinte cardiaque

Plusieurs complications cardiaques surviennent chez les patients atteints de la maladie de Fabry. (223) Le système de conduction du cœur est infiltré par les cellules en surcharge ce qui peut causer des troubles de conduction (raccourcissement de l'intervalle PR), des troubles de repolarisation (décalage du segment ST, inversion de l'onde T) et des troubles du rythme (bradycardie, tachycardie ventriculaire, fibrillation atriale). Ces complications peuvent limiter les capacités à faire un effort physique, nécessitant parfois la pose d'un pace maker (224). L'insuffisance de la valve mitrale peut être retrouvée dès l'enfance ou l'adolescence (225, 226). La cardiomyopathie hypertrophique associe une hypertrophie du ventricule gauche associée à une hypertrophie du septum interventriculaire. Elle se constitue progressivement, plus précocement chez les hommes homozygotes que chez les femmes hétérozygotes. Elle concerne 50 % des hommes et un tiers des femmes et constitue l'une des principales causes de morbi-mortalité dans cette pathologie. Un suivi cardiaque régulier avec ECG et échographie cardiaque pour recherche de complications sont recommandés dans la prise en charge de la maladie de Fabry (227).

3.5.3.4 Symptômes dermatologiques

Des lésions vasculaires typiques appelées angiokératomes sont retrouvées chez les patients atteints de la maladie de Fabry (**Figure 11**). Ces lésions apparaissent précocement dans la maladie et sont facilement reconnaissables : papules kératosiques bleu marine ou rouge foncé dans les couches superficielles de la peau. Elles peuvent être planes ou en relief et ne s'effacent pas à la vitropression. Les angiokératomes sont bilatéraux et sont situés de façon préférentielle entre l'ombilic et les genoux dans la zone dite du caleçon et au niveau des muqueuses (bouche, conjonctives) (228). Ces lésions sont retrouvées chez 66 % des hommes et chez 36 % des femmes (229)(54). Les lésions cutanées augmentent avec la progression de la maladie et sont corrélées avec la sévérité de l'atteinte systémique (230).

L'hypohydrose ou l'anhydrose sont des signes précoces et quasi constants. Les patients décrivent des difficultés à s'adapter à la chaleur et à tolérer l'exercice physique. A contrario, des cas d'hyperhydrose sont également décrits et sont retrouvés chez 11,9 % des femmes et 6,4 % des hommes dans une étude Européenne de cohorte (231).



FIGURE 11 : ANGIOKERATOMES CHEZ UN PATIENT ATTEINT DE LA MALADIE DE FABRY. (189)

3.5.3.5 Symptômes ophtalmologiques

Des opacités cornéennes avec une cornée verticillée caractéristique sont retrouvées chez une majorité des hommes hémizygotés et chez de nombreuses femmes hétérozygotes (232, 233). Ainsi, dans la cohorte FOS, une cornée verticillée était décrite chez 77 % des femmes et chez 73 % des hommes (234). Les lésions débutent dans la couche sous épithéliale puis s'étendent du centre vers la périphérie de la cornée. Les opacités sont disposées en verticilles et sont de couleur crème (233) (Figure 12).

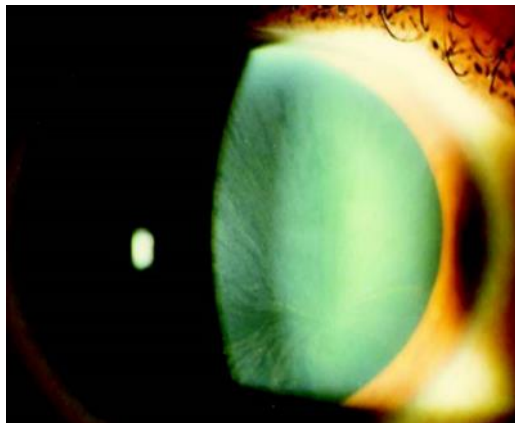


FIGURE 12 : CORNEE VERTICILLEE CHEZ UN PATIENT ATTEINT DE LA MALADIE DE FABRY (234)

L'examen de la conjonctive et le fond d'œil peuvent mettre en évidence des vaisseaux conjonctivaux et rétiens tortueux qui sont prédictifs de l'atteinte systémique de la maladie (233). Des opacités lenticulaires avec l'apparition progressive d'une cataracte sont décrites. Trente pour cent des patients hémizygotés ont une atteinte lenticulaire sous forme d'un dépôt avec une opacité lenticulaire caractéristique. La sévérité de l'atteinte ophtalmologique est corrélée avec la sévérité de la maladie de Fabry (234, 235).

3.5.3.6 Complications rénales

L'atteinte rénale est précoce dès la deuxième décennie sous forme d'une protéinurie décrite chez 43 % des hommes et 26 % des femmes. (236) L'insuffisance rénale s'installe progressivement du fait de la dysfonction du tissu rénal (accumulation des glycosphingolipides, bloc autophagique, stress oxydant...). L'insuffisance rénale survient généralement entre la 3^{ème} et la 5^{ème} décennie, mais des formes plus précoces sont possibles. Une tubulopathie peut également survenir avec une polyurie mimant un diabète insipide. Les patients ont parfois recours à la dialyse ou à la transplantation rénale dans les formes sévères (236, 237).

3.5.3.7 Particularités morphologiques

Les patients avec une forme classique de maladie de Fabry (activité nulle de l'alpha galactosidase A) ont théoriquement des particularités morphologiques prédominantes au niveau de la face : un front enfoncé, des sourcils fournis et broussailleux, des lobules proéminents des oreilles et des oreilles en rotation postérieure, un angle nasal marqué, un grand nez, des reliefs périorbitaires marqués, des lèvres charnues, un prognathisme et des traits grossiers (238).

3.5.3.8 Autres phénotypes et maladie de Fabry chez les femmes hétérozygotes

Lorsque l'activité résiduelle de l'alpha galactosidase A est entre 25 % et 30 % de l'activité normale, les patients ont un phénotype modéré avec des premiers signes de la maladie plus tardifs. L'atteinte multiviscérale est moins marquée rendant le diagnostic plus difficile. L'atteinte cardiaque est souvent au premier plan (239).

La transmission de la maladie est liée à l'X, mais la maladie s'exprime également chez les femmes hétérozygotes avec une pénétrance variable. La survenue de symptômes est liée à l'inactivation du chromosome X (240). Les femmes hétérozygotes ont une mosaïque avec des cellules normales et des cellules avec un déficit en alpha galactosidase A. Les symptômes sont généralement atténués, de survenue plus tardive, et l'atteinte multiviscérale est moins marquée. Néanmoins, certaines femmes hétérozygotes peuvent présenter des complications graves de la maladie avec des risques de défaillance cardiaque, cérébro-vasculaire et rénale à long terme (241).

3.5.4 Diagnostic de la maladie de Fabry

Le diagnostic est évoqué devant les signes cliniques. Tous les symptômes ne sont pas forcément présents et la maladie peut parfois se présenter initialement avec une seule atteinte d'organe ce qui peut rendre le diagnostic difficile. L'activité de l'enzyme alpha galactosidase A peut être dosée dans des leucocytes ou des fibroblastes obtenus à partir de la biopsie cutanée du patient. Une activité enzymatique inférieure à 25 % de la normale permet de poser le diagnostic. Quand l'activité enzymatique est entre 25 et 30 %, le diagnostic est suspecté. Chez les femmes hétérozygotes,

l'interprétation du dosage enzymatique est plus difficile en raison de la mosaïque liée à l'inactivation du chromosome X. Ainsi, une activité enzymatique normale ne permet pas d'éliminer le diagnostic de maladie de Fabry chez la femme. L'analyse sera systématiquement complétée par l'étude moléculaire du gène *GLA* pour confirmer le diagnostic.

La recherche de biomarqueurs dans la maladie de Fabry est importante. En effet, ces biomarqueurs peuvent être utilisés pour faciliter le diagnostic notamment dans les formes atténuées ou en cas de variants de signification inconnue ; pour préciser le pronostic de la maladie et pour suivre la réponse au traitement.

Le Gb3 ne constitue pas un bon biomarqueur dans les formes atténuées ou chez les femmes hétérozygotes, car sa sécrétion plasmatique ou urinaire est inconstante et ne corrèle pas avec l'efficacité thérapeutique. Son dérivé déacylé, le lysoGb3, est plus discriminant, mais il peut également être normal dans les formes atténuées de la maladie (242). L'identification de nouveaux biomarqueurs par l'utilisation d'études « omiques » pourrait permettre d'optimiser le diagnostic et la prise en charge des patients atteints de la maladie de Fabry.

Une fois le diagnostic posé un bilan complet de la maladie sera à réaliser pour évaluer la sévérité de la maladie (227, 243) :

Bilan neurologique : IRM cérébrale avec angio IRM, questionnaire d'évaluation de la douleur

Bilan ophtalmologique : examen spécialisé avec au minimum fond d'œil et examen à la lampe à fente

Bilan rénal : urée et créatinine sanguine, bilan hydro électrolytique complet, recherche d'une protéinurie et d'une hématurie, cystatine C (marqueur de la filtration glomérulaire), échographie rénale. Ponction biopsie rénale en fonction des résultats des premiers bilans

Bilan cardiaque : ECG, échographie cardiaque-doppler, IRM cardiaque, holter ECG, scintigraphie myocardique d'effort. Coronarographie si suspicion d'atteinte coronarienne

Bilan ORL : audiométrie, PEA

Bilan respiratoire : explorations fonctionnelles respiratoires et radiographie ou tomodensitométrie thoracique selon la symptomatologie

Exploration osseuse : ostéodensitométrie pour recherche d'une ostéoporose ou d'une ostéopénie

Bilan biologique complémentaire : bilan lipidique (cholestérol total, HDL-cholestérol, LDL-cholestérol, triglycérides), bilan phosphocalcique (phosphatases alcalines, parathormone, vitamine D, calcémie, phosphorémie), hémogramme, à la recherche d'une anémie carentielle (fer, folates, vitamine B12)

3.5.5 Prise en charge thérapeutique de la maladie de Fabry

Le traitement a pour objectif de ralentir et/ou de stopper la progression de la maladie de Fabry. La prise en charge est multidisciplinaire du fait de l'atteinte multisystémique (227, 243).

L'enzymothérapie substitutive est disponible depuis 2001 (244). Il existe deux formes d'enzymothérapie substitutive : l'agalsidase alpha (Replagal®, Takeda) et l'agalsidase beta (Fabrazyme®, Sanofi Genzyme). L'Agalsidase alpha est administré à la posologie de 0,2 mg/kg par voie intraveineuse hebdomadaire. L'Agalsidase Beta est perfusée par voie intra veineuse à la dose de 1 mg/kg toutes les 2 semaines. L'efficacité et la tolérance de ces traitements ont été prouvées par des études de cohorte (245).

L'enzymothérapie substitutive a été étudiée chez l'enfant dans des essais cliniques en ouvert. Le traitement était bien toléré et apportait des bénéfices avec une diminution des douleurs neuropathiques, une diminution des symptômes digestifs, une stabilisation des fonctions cardiaques et rénales associés à une diminution de l'excrétion du Gb3. En revanche le traitement n'avait pas de bénéfice sur l'incidence des accidents vasculaires cérébraux. Les patients développent des anticorps (IgG) dirigés contre l'enzymothérapie, mais ceux-ci n'affecteraient pas l'efficacité du traitement (246-248).

Des molécules chaperons seules ou en association à l'enzymothérapie substitutive existent. Le migalastat (Galafold®) est une molécule chaperon efficace dans la maladie de Fabry. Elle stabilise la fonction ventriculaire gauche, la fonction rénale et diminue les symptômes digestifs (249). Cette approche est disponible uniquement pour certains variants spécifiques (250).

D'autres traitements sont complémentaires de ces thérapeutiques spécifiques (227, 243) et seront adaptés en fonction des différentes complications retrouvées dans le bilan de la maladie.

Sur le plan cardio-vasculaire, un traitement antiagrégant plaquettaire peut être utilisé en prévention des accidents vasculaires cérébraux. Les facteurs de risque cardio-vasculaires doivent être strictement contrôlés : normalisation de la tension artérielle (TA), arrêt du tabagisme, normalisation d'une éventuelle hyperlipidémie ... Les troubles du rythme ou de la conduction peuvent nécessiter un traitement anti arythmique, un traitement chirurgical et/ou la mise en place d'un stimulateur cardiaque.

Sur le plan néphrologique, les inhibiteurs de l'enzyme de conversion et les inhibiteurs de l'angiotensine sont prescrits en cas de protéinurie à visée néphroprotectrice. Ils favorisent la préservation de la fonction rénale en association avec l'enzymothérapie substitutive (251). L'insuffisance rénale peut parfois nécessiter le recours à la dialyse ou à une greffe rénale (237).

Les douleurs neuropathiques peuvent être traitées par des antalgiques tels que la lamotrigine, la carbamazépine ou la gabapentine (252).

La recherche de nouveaux traitements pour la maladie de Fabry est très active comme dans les autres maladies lysosomales. Des thérapies par réduction de substrat ont montré des effets bénéfiques dans un modèle murin (252). Des études utilisant la réduction de substrat (inhibiteur de la glucosylcéramide synthase) ou des techniques de thérapie génique sont actuellement en cours chez l'homme.

IV. La médecine de Précision

4.1 Définition

La médecine de précision également appelée médecine personnalisée a pour objectif une prise en charge du patient adaptée à ses caractéristiques personnelles. Elle utilise les informations obtenues par le biais des différentes explorations cliniques, biologiques, radiologiques.... Ces données individuelles vont nourrir des bases populationnelles qui à leur tour peuvent être utilisées pour une prise de décision au niveau individuel. La mise en évidence de biomarqueurs spécifiques et discriminants permet d'aider au diagnostic et guide la prise en charge thérapeutique. Les sciences « omiques » qui se sont développées ces dernières années ont permis l'émergence de cette médecine personnalisée. Les sciences « omiques » permettent d'extraire des données à partir d'un système biologique de façon globale grâce à des approches ciblées ou non ciblées. Ces nouvelles techniques permettent des analyses à haut débit et en multiplex et différent, ainsi, des approches conventionnelles qui sont réductionnistes et séquentielles (253).

Les approches « omiques » aboutissent à un grand nombre de données qui doivent être converties sous une forme exploitable. Des techniques de bioinformatique et biomathématiques sont appliquées. Ces outils spécialisés permettent l'analyse comparative des données et permettent de transformer les informations brutes en données interprétables par le médecin qui prend en charge le patient.

4.2 Les supports de l'information biologique

Les approches « omiques » peuvent étudier les différents supports de l'information biologique : les gènes par la génomique, les acides ribonucléiques par la transcriptomique, les protéines par la protéomique et le métabolome par la métabolomique (**Figure 13**). Des méthodes de séquençage haut débit ont été développées (séquençage nouvelle génération ou NGS) et permettent l'analyse des acides nucléiques (ADN et ARN) en déterminant leur ordre d'enchaînement dans une séquence donnée. L'analyse des protéines et des métabolites se fait essentiellement par Résonance

Magnétique Nucléaire (RMN) ou par Spectrométrie de Masse (MS). Nous passerons en revue ci-dessous brièvement les principales sciences « omiques ».

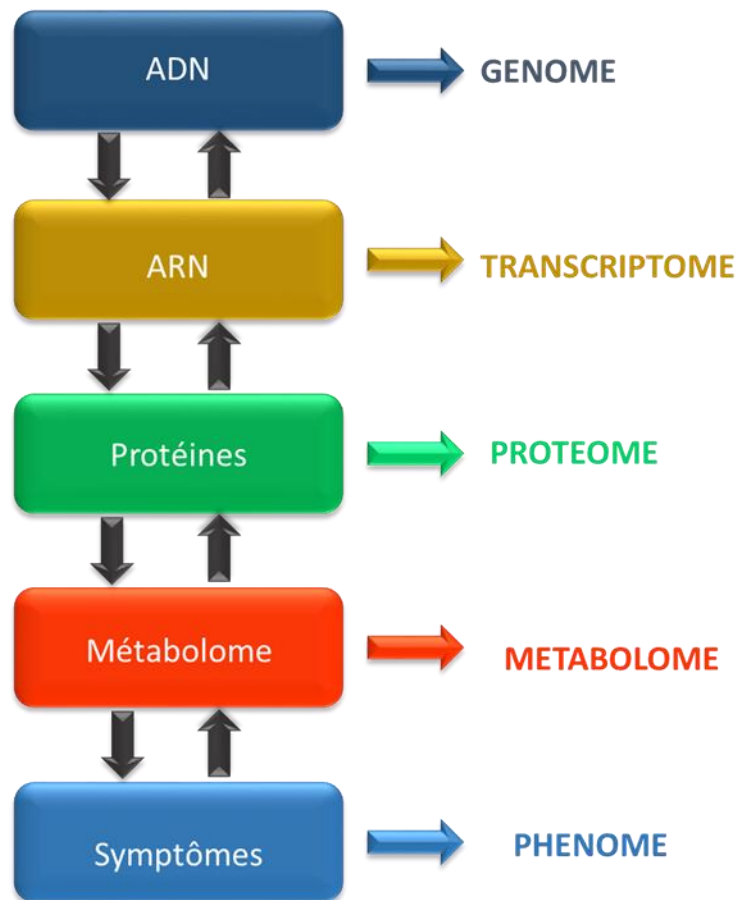


FIGURE 13 : LES DIFFERENTS SUPPORTS DE L'INFORMATION BIOLOGIQUE ET LES SCIENCES « OMIQUES » CORRESPONDANTES QUI PERMETTENT LEUR EXPLORATION

4.2.1 La génomique

Le génome est l'ensemble des acides désoxyribonucléiques (ADN) présents dans un organisme. L'acide désoxyribonucléique (ADN) est principalement retrouvé dans les noyaux des cellules (ADN cellulaire) mais également dans les mitochondries (ADN mitochondrial). La séquence de l'ADN varie peu dans le temps (254, 255). Le génome humain a été séquencé en entier en 2001 par la technique de Sanger. Par la suite, des techniques de séquençage haut débit ont été développées. Le séquençage haut débit en multiplex est actuellement utilisé en pratique clinique pour l'étude du génome. L'analyse ciblée de plusieurs gènes d'intérêt est possible par le séquençage d'un panel de gènes. L'étude de l'ensemble des séquences exoniques et des jonctions intron-exon est réalisée par un

séquençage d'exome (*Whole exome sequencing* ou WES) et l'analyse de l'ensemble du génome par le *Whole genome sequencing* (WGS).

Le génome peut subir des modifications épigénétiques de type méthylation qui seront transmises aux générations suivantes. Ces phénomènes épigénétiques sont secondaires à l'environnement dans lequel vit le patient. Celles-ci peuvent modifier le phénotype et le devenir des futures générations (256).

4.2.2 La transcriptomique

4.2.2.1 Définition

Le transcriptome est composé des acides ribonucléiques (ARN). L'ARN présente une variabilité avec une expression différente des gènes selon le type de cellule, selon l'environnement et selon les processus biologiques en cours. On distingue plusieurs types d'ARN (par exemple les ARN dédiés à la synthèse protéique ou les ARN de régulation (miARN)). La transcriptomique permet d'étudier la quantité d'expression des gènes.

La transcriptomique est l'étude de l'ensemble des ARN (ARNm, ARNr, ARNt, ARNnc) présents dans une cellule à un temps donné. La transcriptomique repose sur la quantification systématique des ARN transcrits et permet d'obtenir des informations sur les séquences exprimées. Elle peut permettre de comparer l'ARN entre deux populations (par exemple malades versus contrôles sains), d'évaluer la réponse d'une cellule à son environnement (par exemple réponse à une infection, à un traitement,...), d'évaluer l'évolution de la transcription au cours du développement (par exemple au cours de l'embryogénèse) ou de comparer les ARN dans des types cellulaires différents.

La transcriptomique fait appel à plusieurs techniques haut débit : microarrays (puces), *RNA sequencing*,...

Le *RNA sequencing* permet d'étudier l'ensemble des ARN présents dans une cellule ou un fluide biologique. Il séquence les ADNc et les compare à un génome ou à un transcriptome de référence. Cette technique est basée sur le fait que plus un gène est exprimé dans une condition, plus les transcrits correspondants sont synthétisés et plus les ADNc correspondants ont une chance d'être séquencés. Le *RNA sequencing* fournit des dizaines à des centaines de millions d'informations à partir d'une seule analyse. L'interprétation des données obtenues nécessite un travail laborieux d'analyse de bioinformatique et de biostatistiques.

4.2.2.2 Technique du RNA sequencing

La première étape consiste à choisir une plateforme de séquençage. La plupart des plateformes utilisent la technique de séquençage par synthèse par le biais d'une ADN polymérase ou d'une ligase.

Elles peuvent séquencer un seul type de molécule ou de multiples copies d'une molécule d'ADN ou d'ARN. Le choix de la plateforme de séquençage dépend des objectifs de l'étude.

Pour les plateformes de séquençage d'une seule molécule, le protocole de préparation des échantillons est simple et peut être facilement utilisé quand l'échantillon d'ARN est de petite taille. Ce type de plateforme permet également d'éviter l'étape d'amplification par PCR.

Les plateformes de séquençage de plusieurs molécules utilisent généralement des étapes d'amplification. Les informations obtenues sur les niveaux d'expression des ARN sont plus nombreuses mais ces techniques nécessitent plus de points de contrôle et l'analyse bio-informatique est plus laborieuse.

Les plateformes de séquençage d'une seule molécule sont associées à un taux d'erreur intrinsèque élevé d'environ 5 %, essentiellement des insertions et des délétions. Un taux élevé d'erreur rend difficile l'appariement des lectures du séquençage avec le génome de référence. Un taux minimal d'erreur de séquençage est important pour les séquençages de micro RNA qui sont des molécules de petite taille pour éviter la perte de lecture de données brutes. Les plateformes de type Illumina ou SOLiD permettent une plus grande profondeur de séquençage et de détecter plus facilement les transcrits peu exprimés.

L'assemblage des différentes séquences d'ARN permet la lecture de l'ensemble du transcriptome avec les séquences d'ARN messagers et les séquences d'ARN non codantes.

4.2.2.2.1 Préparation des librairies

La préparation des librairies de RNA sequencing nécessite la conversion de l'ARN cellulaire en molécules pouvant être séquencées. L'ARN ribosomal peut constituer jusqu'à 80 % de l'ARN de la cellule. Le séquençage de ces séquences d'ARN peut réduire la profondeur de lecture induisant une perte de lecture des ARN peu exprimés. L'ARN ribosomal peut être éliminé par dégradation enzymatique utilisant des nucléases spécifiques (257); ou par des méthodes de déplétion basées sur de l'hybridation (258). Cela permet de garantir une profondeur adéquate pour les transcrits les plus rares. La préparation de la librairie dépend de la taille souhaitée. L'ARN peut être fragmenté par hydrolyse chimique ou par digestion enzymatique jusqu'à la taille appropriée pour la plateforme de séquençage choisie. Si les ARN étudiés sont de petite taille notamment pour les miARN, les étapes de fragmentation ne sont pas nécessaires. Pour les ARN plus grands, une fragmentation est nécessaire pour obtenir des ARN entre 200 et 250 nt pour pouvoir être lus par les plateformes Illumina ou SOLiD.

Une fois la taille souhaitée obtenue, les ARN sont convertis en ARN complémentaires par une reverse transcriptase utilisant des amorces aléatoires. Les oligonucléotides adaptateurs sont ensuite

ligaturés à l'ADNc pour permettre l'amplification et le séquençage. Pour les plateformes SOLiD ou Illumina, des adaptateurs spécifiques liés aux extrémités 3' ou 5' avant l'étape de la transcriptase reverse et la PCR permet l'identification de la direction du brin d'ARN d'origine.

L'étape de préparation de l'échantillon est importante car elle peut être associée à l'introduction de biais notamment si la transcriptase reverse copie l'ADNc du premier brin mais également du deuxième brin; ou du fait de la ligature aléatoire (259, 260).

4.2.2.2 Immunoprécipitation et séquençage

Dans l'étape d'immunoprécipitation, un anticorps dirigé contre une protéine spécifique d'intérêt est utilisé pour récupérer les ARN liés à la protéine étudiée. La récupération de l'ARN lié est une étape importante. Si l'anticorps se dissocie de la protéine étudiée ou la protéine de l'ARN, aucun signal ne sera observé. A contrario, si l'ARN s'associe à la protéine après la lyse cellulaire, des artéfacts peuvent être détectés. In vivo, la technique de *crosslinking* permet de remédier à ces difficultés (261). Cette technique permet de détecter les interactions ARN protéine cinétiquement stables. Les complexes ARN protéines sont ensuite purifiés par électrophorèse pour réduire les ARN non spécifiques.

Bien que simple en théorie, le séquençage et l'immunoprécipitation de l'ARN est une technique complexe en pratique. L'étape de *crosslinking* n'est pas très performante et seule une petite quantité d'ARN est disponible pour la construction de la librairie. Des nucléosides photoactivables peuvent être utilisés pour augmenter l'efficacité de l'étape de nucléation mais ils peuvent rendre plus difficile l'appariement de séquences de lecture avec le génome de référence (262).

4.2.2.3 Analyse des données

Le séquençage de l'ARN aboutit à un nombre important de données. Ainsi, une plateforme HiSeq2000 produit jusqu'à 50 G de données dans un *run* de séquençage. Ces données doivent être analysées pour identifier les correspondances avec le transcriptome mais également être assemblées et quantifiées sous forme de transcrits.

L'analyse des données comprend deux étapes principales. La première étape consiste à supprimer les artéfacts et les erreurs de l'ensemble des données. Les erreurs de séquençage peuvent être supprimées ou corrigées à partir du score de qualité. Lorsque le génome étudié est volumineux, un filtrage des données est possible avant la phase d'alignement. La deuxième étape consiste à aligner les données sur le génome de référence. L'alignement des données et la technique d'analyse dépendent de la plateforme choisie et des objectifs du *RNA sequencing*.

Certaines plateformes utilisent leur propre pipeline d'analyse des données mais il est également possible d'utiliser des programmes analytiques indépendants.

Le séquençage d'ARN est une technique puissante mais une validation expérimentale des données est indispensable pour confirmer les données obtenues. Certaines données peuvent être vérifiées par des expérimentations sur un locus d'intérêt : utilisation d'une technique de *RNA sequencing* spécifique ciblée, d'une PCR quantitative avec des amorces stratégiquement choisies ou une amplification rapide des extrémités de l'ADNc (260).

4.2.3 La protéomique

4.2.3.1 Définitions

Le protéome est constitué de l'ensemble des protéines d'un système biologique à un moment t. A partir du même génome, le nombre de protéines différentes générées est important. Pendant et après le processus de traduction, les protéines vont subir des modifications dites co et post-traductionnelles qui seront variables en fonction du type cellulaire et en fonction du rôle attribué à la protéine. La protéomique a pour objectif de caractériser l'expression, la structure, la fonction et les interactions des protéines présentes dans un échantillon biologique (**Figure 14**) (263). Une meilleure connaissance des protéines aide à la compréhension des mécanismes moléculaires des processus biologiques. La protéomique permet également de rechercher des biomarqueurs spécifiques d'une pathologie. Le développement de la protéomique s'est accéléré grâce au développement de la spectrométrie de masse pour l'étude des macromolécules ce qui valut à John Fenn et Koichi Tanaka le prix Nobel de Chimie en 2002 (264-266). Avant l'arrivée de la spectrométrie de masse, l'analyse des protéines était laborieuse et le biologiste devait purifier une grande quantité de protéines dans un échantillon avant de débiter leur étude. Deux approches sont possibles dans ces analyses : l'approche non ciblée qui va étudier l'ensemble des protéines contenues dans un échantillon biologique de façon exhaustive. Cette première étude va permettre d'émettre des hypothèses pour élaborer une étude plus ciblée sur certaines protéines. La résonance magnétique nucléaire (RMN) et la spectrométrie de masse (MS) haute résolution sont utilisées pour cette approche non ciblée. L'analyse ciblée est possible par spectrométrie de masse et par des techniques d'immunodosage. La partie suivante détaille brièvement plusieurs techniques utilisées en protéomique. Un chapitre est consacré à la technique ciblée multiplex MSD® qui a été utilisée dans ce travail de thèse.

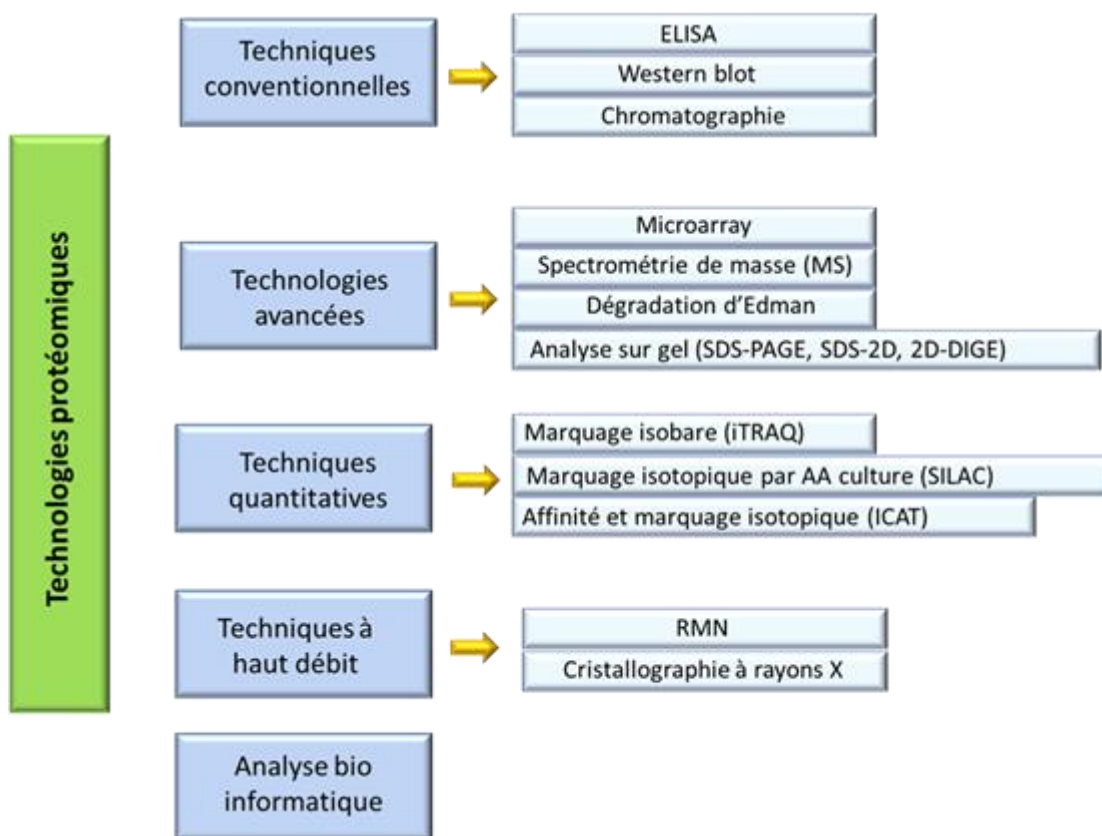


FIGURE 14 : LES DIFFERENTES TECHNIQUES UTILISABLES EN PROTEOMIQUE

RMN= RESONANCE MAGNETIQUE NUCLEAIRE, MS= SPECTROMETRIE DE MASSE, SILAC=STABLE ISOTOPE LABELING WITH AMINO ACIDS IN CELL CULTURE, ICAT= ISOTOPE-CODED AFFINITY TAG, iTRAQ= ISOBARIC TAG FOR RELATIVE AND QUANTITATIVE QUANTITATION

4.2.3.2 Techniques utilisées pour l'étude protéomique

4.2.3.2.1 Technique d'analyse haut débit : MS et RMN

Les techniques de spectrométrie de masse permettent de séparer et d'identifier les différentes protéines contenues dans un système biologique par la mesure du rapport m/z (masse/charge) des protéines avec une grande sensibilité (267). La spectrométrie de masse est largement utilisée pour les investigations cliniques et donc plus accessible que la RMN.

La résonance magnétique nucléaire est basée sur la détection des transitions de spin selon l'état énergétique des noyaux atomiques placés dans un champ magnétique intense. Cette approche est plus particulièrement utilisée à la recherche. Cette technique est moins sensible que la MS et lui est souvent complémentaire. Le **Tableau 5** (partie métabolomique) compare plus précisément ces deux techniques. La technique de MS est détaillée dans la partie métabolomique.

4.2.3.2.2 Techniques ELISA et Western Blot

Certaines techniques ciblées dites d'immunodosage font appel à des anticorps dirigés contre des protéines d'intérêt. Il s'agit par exemple des techniques de type ELISA et western blot. Ces techniques ciblées permettent de détecter la présence d'une protéine individuelle dans un échantillon. La technique ELISA est une technique sensible d'immunodosage utilisée en routine en pratique clinique. Elle se base sur la détection d'un antigène ou d'un anticorps présent sur une surface en utilisant une enzyme couplée à un anticorps. L'activité enzymatique est mesurée. Elle est proportionnelle à la quantité d'antigène ou d'anticorps présent dans l'échantillon (268). Le western blot est une technique permettant la détection de protéines présente en faible quantité dans un échantillon. Les protéines sont d'abord séparées par une technique d'électrophorèse puis transférées sur une membrane de nitrocellulose. La protéine d'intérêt est ensuite détectée par l'utilisation d'un anticorps dirigé contre la protéine d'intérêt (269).

4.2.3.2.3 Technique de dosage ciblée des protéines par technologie multiplex (MSD®)

La protéomique peut également faire appel à d'autres outils permettant l'analyse rapide d'un grand nombre de protéines sur un échantillon biologique. Nous détaillerons ici une technique de dosage ciblée des protéines via une technologie multiplex (MSD®). Cette technique permet la détection de molécules ciblées dans un échantillon biologique par électrochimiluminescence sur un système de plaques modélisées (**Figure 15**). Les échantillons biologiques sont déposés dans la plaque puis incubés avec des anticorps « tagués » dirigés contre les protéines d'intérêt pendant un temps défini. La stimulation électrique dans les différents puits de la plaque va induire un signal lumineux qui est détecté par des photodétecteurs ultra sensibles. La quantification de ce signal évalue la quantité de protéine spécifique dans l'échantillon biologique. L'expérimentation est suivie par un travail de bioinformatique et de biostatistiques car les données générées sont massives. Cette technique a été appliquée dans plusieurs domaines (270, 271). Un des avantages de cette technique ciblée est sa très bonne sensibilité associée à sa spécificité. De plus, cette technologie permet une analyse reproductible. Cette technique a été utilisée au cours de ce travail de thèse.

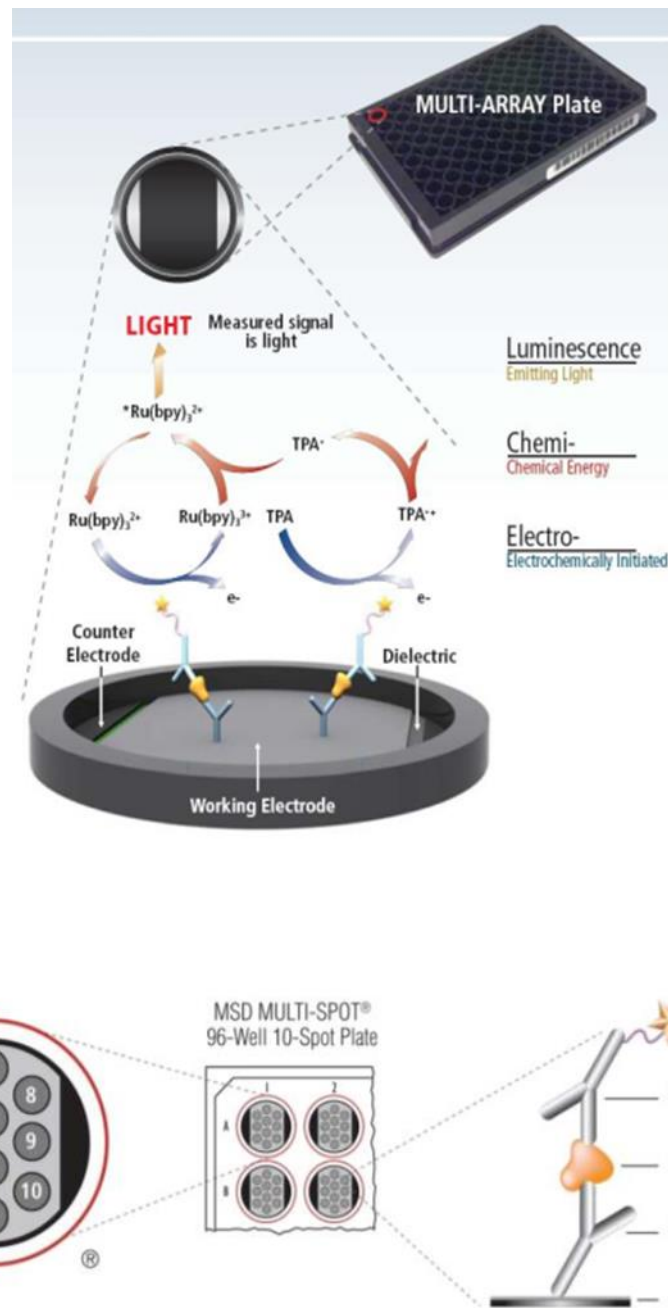


FIGURE 15 : ANALYSE MULTI ARRAY MSD®. CHAQUE Puits POSSEDE DES ELECTRODES ET DES ANTICORPS DE CAPTURE SUR LESQUELS VONT SE FIXER LES MOLECULES D'INTERET. DES ANTICORPS « TAGUES » SONT INCUBES AVEC L'ECHANTILLON BIOLOGIQUE PENDANT UN TEMPS DEFINI. LA STIMULATION ELECTRIQUE DANS LES DIFFERENTS Puits DE LA PLAQUE VA INDUIRE UN SIGNAL LUMINEUX QUI EST DETECTE PAR DES PHOTODETECTEURS ULTRA SENSIBLES.

4.2.4 La métabolomique

Le terme « métabolome » a été utilisé pour la première fois en 1998 par S.G Oliver (272). Il se définit comme l'ensemble des métabolites synthétisés par un organisme donné. En pratique, il désigne l'ensemble des métabolites présents dans un fluide biologique, un tissu ou une cellule à un moment t (273). Les métabolites sont de petites molécules organiques issues ou à l'origine d'une réaction enzymatique. Le métabolome se caractérise par une grande variabilité constituant l'expression la

plus distale du génome. Les molécules du métabolome ont subi un grand nombre de modifications (glycosylation, phosphorylation, acétylation, hydroxylation,...). L'expression des métabolites varie également en réponse à l'environnement qui l'entoure (273).

4.2.4.1 Définition

La métabolomique est la science « omique » qui permet d'analyser le métabolome en étudiant les caractéristiques biochimiques et moléculaires de ce dernier et ses modifications liées à des facteurs environnementaux ou génétiques (274). Cette science « omique » permet de caractériser les voies biochimiques impliquées dans un processus physiopathologique. Elle se développe de plus en plus avec des applications dans différents domaines de la biologie comme en témoigne le nombre de publications sur les dernières années (273, 275-277).

Deux approches sont utilisées en métabolomique : l'approche non ciblée et l'approche ciblée (273). L'approche non ciblée analyse toutes les molécules présentes dans un fluide ou un tissu biologique de la façon la plus exhaustive possible. Elle a pour objectif d'obtenir un maximum de données pour générer des hypothèses qui pourront servir pour une analyse ciblée ultérieure. La résonance magnétique nucléaire et la spectrométrie de masse haute résolution sont utilisées pour cette approche non ciblée. Ces techniques seront détaillées plus loin. L'analyse ciblée va analyser spécifiquement des métabolites choisis. Elle vient compléter l'analyse initiale non ciblée. Cela nécessite de connaître au préalable les molécules d'intérêt. En clinique, elle est utilisée couramment pour la quantification de molécules spécifiques comme c'est le cas par exemple pour le dépistage de la phénylcétonurie chez le nouveau-né avec le dosage spécifique du taux de phénylalanine sur un grand nombre d'échantillons de patients (278).

4.2.4.2 Technique de l'analyse métabolomique

4.2.4.2.1 Recueil et préparation des échantillons

L'échantillon doit être choisi en fonction des objectifs de l'analyse métabolomique afin de recueillir les informations les plus pertinentes en réponse à la question posée. L'étape de préparation des échantillons est primordiale car elle détermine la qualité de l'analyse métabolomique sur le plan qualitatif et quantitatif. La technique choisie doit être simple, rapide et reproductible. Elle doit également être la moins sélective possible vis-à-vis des métabolites afin de conserver un maximum de données biochimiques (279).

Après choix du tissu ou du fluide étudié, la première étape consiste à homogénéiser le prélèvement afin de limiter la variabilité liée à la nature de l'échantillon lui-même. Pour la préparation des échantillons issus de tissus solides, la cryopulvérisation ou un broyage avec des billes sont souvent utilisés (277, 279). La cryopulvérisation permet entre autres le maintien de la stabilité de

l'échantillon. L'homogénéisation peut être optimisée par l'utilisation de billes inertes qui vont augmenter la surface de contact entre le solvant et l'échantillon.

La seconde étape est l'extraction visant à sélectionner dans l'échantillon les métabolites analysables et à éliminer les molécules interférentes qui peuvent rendre l'analyse difficile (**Figure 16**). Pour les urines, l'analyse peut débuter après simple centrifugation et dilution sans cette étape d'extraction (280). Pour les autres tissus ou fluides, une étape d'extraction peut être nécessaire. En cas de faible concentration prévisible de certains métabolites, l'extraction peut également permettre de concentrer ces derniers. La technique utilisée pour réaliser l'extraction dépend de l'objectif du travail et de la nature de l'échantillon biologique (281). Si l'analyse de l'échantillon vise une approche globale, l'extraction vise à garder analysable un nombre maximal de métabolites. Si l'analyse est dite ciblée sur un type de molécule en particulier, l'extraction vise à éliminer les composants indésirables de la matrice (277, 279, 282).

Le choix du solvant d'extraction a pour but d'optimiser la performance de l'extraction sans altérer les métabolites à analyser. Pour les métabolites polaires, le solvant le plus souvent utilisé pour précipiter les protéines est le méthanol (283). Pour les métabolites lipophiles, un système biphasique est utilisé pour que les composés lipophiles et polaires soient répartis entre les phases organiques et aqueuses. D'autres techniques telles que l'extraction solide-liquide qui enrichit progressivement l'extrait avec les métabolites cibles sont également utilisables (284).

Le prélèvement doit se faire en limitant au maximum le stress ou le traumatisme afin d'éviter des perturbations du métabolome et de limiter la lipolyse in vivo. Les échantillons doivent être extraits le plus rapidement possible après prélèvement sur un organisme vivant afin de préserver la nature des métabolites à étudier. Si l'homogénéisation et l'extraction doivent se faire de façon différée après le prélèvement, l'échantillon doit être immédiatement congelé dans de l'azote liquide ou de la carboglace puis stocké à -80 °C dans un récipient bien fermé et opaque afin d'éviter la dégradation des métabolites. Les conditions de conservation doivent être contrôlées afin de garantir la qualité de l'analyse ultérieure (285).

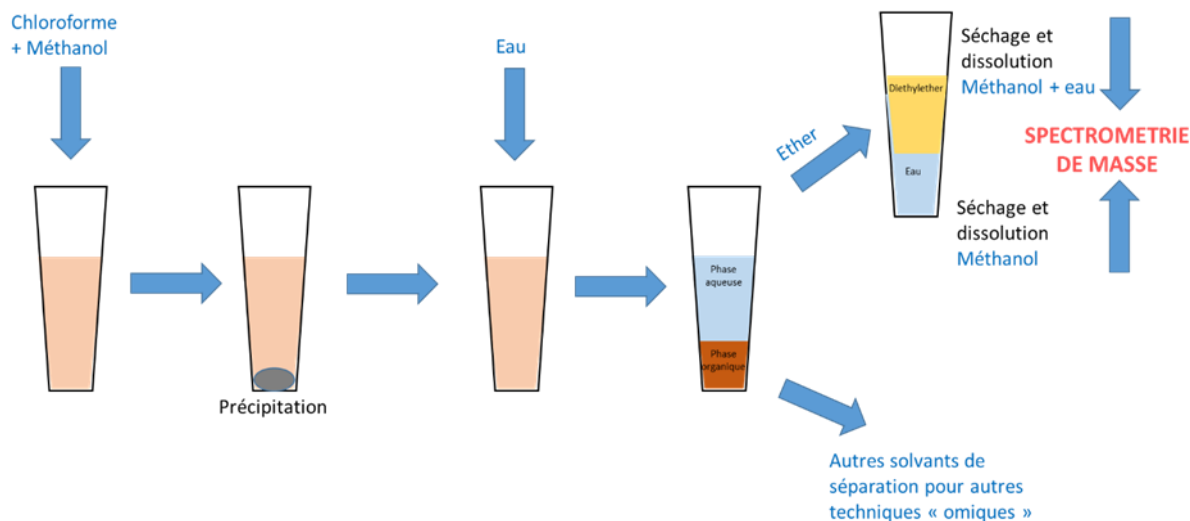


FIGURE 16 : PREPARATION DES ECHANTILLONS POUR LES ETUDES DE METABOLOMIQUE. ETAPE D'EXTRACTION DES METABOLITES. DU METHANOL EST UTILISE COMME SOLVANT D'EXTRACTION ET PERMET DE FAIRE PRECIPITER LES PROTEINES. L'AJOUT D'EAU DANS LE TUBE PERMET D'OBTENIR DEUX PHASES : UNE PHASE AQUEUSE ET UNE PHASE ORGANIQUE. ENFIN, L'AJOUT D'UN SOLVANT DE SEPARATION PERMET D'ISOLER LES MOLECULES A ETUDIER.

4.2.4.2.2 *Analyse métabolomique par résonance magnétique nucléaire ou par spectrométrie de masse*

La résonance Magnétique Nucléaire

La résonance magnétique nucléaire est utilisée pour une approche métabolomique globale. Elle est basée sur la détection des transitions de spin selon l'état énergétique des noyaux atomiques soumis à un champ magnétique intense. Elle permet une analyse rapide, reproductible, et de qualité tout en respectant la nature des molécules et en limitant le temps de préparation des échantillons. L'analyse est quantitative. L'analyse des spectres apporte également des informations précieuses sur la structure des molécules. Néanmoins, cette technique est coûteuse et son principal défaut réside dans son manque de sensibilité (286-288). Elle doit souvent être couplée avec la spectrométrie de masse. L'association des deux techniques permet une analyse plus complète du métabolome.

Analyse comparative de la spectrométrie de masse et de la résonance magnétique nucléaire

Les deux techniques sont complémentaires avec des caractéristiques qui leur sont propres. Elles peuvent être utilisées en parallèle pour optimiser l'analyse métabolique. Le Tableau 5 résume les principales caractéristiques de chacune des techniques (**Tableau 5**)

	RMN	MS
Sensibilité	Faible Peut être améliorée par la RMN haute résolution ou 2D	Bonne Peut-être altérée par la suppression de certains ions dans les échantillons complexes
Mesure	Analyse de la totalité de l'échantillon en une seule mesure	Nécessite plusieurs analyses chromatographiques pour les différentes classes de métabolites
Reproductibilité	Très bonne	Modérée Préparation plus complexe des échantillons
Préparation de l'échantillon	Préparation minimale de l'échantillon	Nécessité de plusieurs colonnes de chromatographies en phase liquide avec conditions d'ionisation optimales
Temps d'expérimentation	5 minutes pour la RMN 1D	Moins de 3 minutes pour la perfusion directe mais plus de 10 minutes pour une analyse simple par GC-MS ou LC-MS
Analyse tissulaire	Possible, en utilisant la HRMAS-RMN Analyse tissulaire directe	Indirecte Nécessite une extraction tissulaire Utilisation possible sur tissu par technique MALDI-MS
Analyse ciblée	Technique peu performante pour l'analyse ciblée	Meilleure technique pour l'analyse ciblée
Utilisable pour études <i>in vivo</i>	Oui	Non
Etude moléculaire dynamique	La RMN peut être utilisée pour l'étude moléculaire dynamique et l'étude de la diffusion moléculaire	Non

TABLEAU 5 : TABLEAU COMPARATIF DES DEUX TECHNIQUES D'ANALYSE METABOLOMIQUE : RMN ET MS (D'APRES (287)).

RMN 1D : RESONNANCE MAGNETIQUE NUCLEAIRE EN UNE DIMENSION, GC-MS : SPECTROMETRIE DE MASSE AVEC CHROMATOGRAPHIE EN PHASE GAZEUSE, LC-MS : SPECTROMETRIE DE MASSE AVEC CHROMATOGRAPHIE EN PHASE LIQUIDE, HRMAS : HIGH-RESOLUTION MAGIC-ANGLE SPINNING, MALDI-MS : MATRIX ASSISTED LASER DESORPTION IONISATION-MASS SPECTROMETRY.

La spectrométrie de masse

Le principe de la spectrométrie de masse repose sur l'ionisation des molécules et la mesure du rapport m/z (masse/charge) de ces molécules. Cette technique permet, ainsi, la détection d'un grand nombre de métabolites avec une sensibilité élevée de l'ordre du micromole. Le processus d'ionisation va permettre la séparation des molécules chargées qui seront détectées par la formation d'un courant électrique. Ce courant électrique est converti en spectre de masse (**Figure 17**).

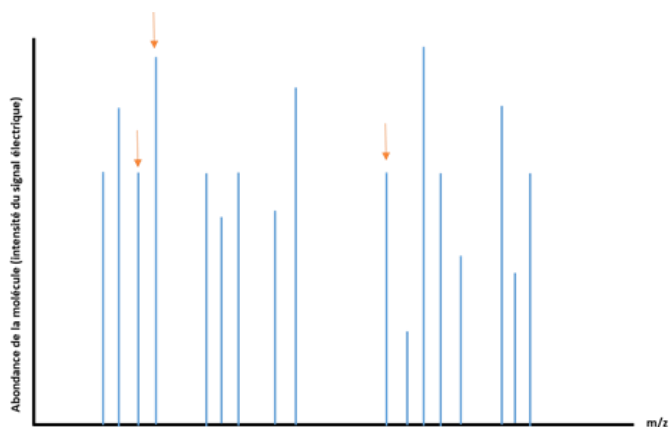


FIGURE 17 : REPRESENTATION SCHEMATIQUE D'UN SPECTRE DE MASSE. CHAQUE PIC CORRESPOND A UN ION DONNE. LE RAPPORT MASSE/ CHARGE CORRESPOND A LA SIGNATURE IONIQUE SPECIFIQUE D'UNE MOLECULE DONNEE. M/Z : RAPPORT MASSE SUR CHARGE.

Les différentes étapes d'une analyse en spectrométrie de masse

L'analyse par spectrométrie de masse comporte plusieurs étapes : l'injection, l'ionisation, la séparation/analyse et la détection. Ces différentes étapes sont détaillées plus loin.

Injection

L'échantillon peut être introduit dans la source d'ionisation directement ou indirectement après une étape de séparation préalable (par électrophorèse ou par chromatographie). Le prélèvement peut être introduit sous forme gazeuse, solide ou liquide. L'introduction directe permet une analyse rapide à haut débit sur un échantillon natif ou sur un extrait. L'introduction indirecte est conseillée en cas d'échantillon complexe. Elle permet de s'affranchir de l'effet matrice. Le choix de la technique de séparation varie selon la nature du métabolite étudié et l'objectif de l'étude (289). La spectrométrie de masse associée à la chromatographie gazeuse est plus adaptée aux molécules volatiles et semi-volatiles alors que la chromatographie liquide est adaptée aux solutions liquides. L'électrophorèse est utilisée pour les molécules chargées. L'association de la spectrométrie de masse à une technique de séparation permet la séparation d'espèces similaires qui ne pourraient pas être séparées par mesure de masse seule.

Ionisation

L'ionisation peut se faire par désorption ou par nébulisation, sous vide ou à pression atmosphérique. Cette étape est importante car elle détermine la nature des molécules qui seront analysées par spectrométrie de masse. La molécule va acquérir des propriétés physico-chimiques qui lui sont propres et qui permettent l'étape de séparation.

Différentes sources d'ionisation peuvent être utilisées (206) : *Matrix assisted laser desorption ionisation* (MALDI), *Electrospray ionisation* (ESI), *Atmospheric pressure chemical ionization* (APCI),

Atmospheric pressure photoionization (APPI). L'ionisation électronique (EI) est souvent utilisée en GC-MS et l'ionisation par electrospray est souvent utilisée en LC-MS.

Source MALDI

Elle permet d'ioniser les molécules via un laser ultraviolet (290). La matrice utilisée doit être sensible à l'ultraviolet, être une source de protons et se vaporiser sous le laser. La source MALDI est adaptée à l'ionisation des molécules fragiles telles que les polymères et les protéines avec modifications post traductionnelles et les acides nucléiques. L'analyse après ce type d'ionisation est généralement simple car les ions produits sont souvent monochargés. Les spectres de masse sont donc peu complexes à analyser.

Source ESI (electrospray ionisation)

Cette source d'ionisation a été développée dans les années 1980. L'ionisation par *electrospray* repose sur la formation de microgouttelettes contenant des molécules chargées à partir d'une solution solvants-molécules. Les molécules chargées peuvent être obtenues à partir de biomolécules complexes (265, 291). La solution contenant les solvants et les biomolécules est déposée dans un capillaire à haute température au sein de la chambre d'ionisation. Sous l'effet de la chaleur, des microgouttelettes se forment à l'extrémité du capillaire. La solution solvant s'évapore et les charges positives des molécules induisent une diminution de la taille des gouttelettes et une séparation des molécules (**Figure 18**). Les molécules ionisées produites par cette technique sont le plus souvent multichargées ce qui facilite l'analyse ultérieure.

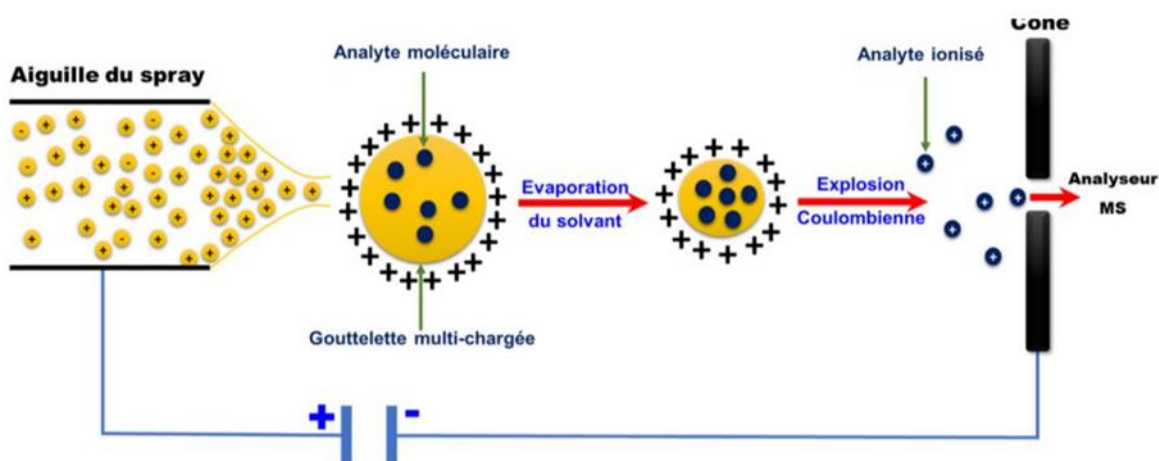


FIGURE 18 : PRINCIPE DE L'IONISATION PAR ELECTROSPRAY. LE SOLUTE A ETUDIER EST INTRODUIT DANS LA SOURCE PAR UN CAPILLAIRE. UNE TENSION ELEVEE Y EST EXERCEE. UN EXCES DE CHARGE SE FORME A SA SURFACE DU FAIT D'UN GRADIENT ELECTRIQUE ENTRE LE CAPILLAIRE ET LA CONTRE-ELECTRODE. SUR LA POINTE DU CONE, DES GOUTTELETTES CHARGES SE FORMENT. ELLES S'EVAPORENT AVANT L'ENTREE DANS LE SPECTROMETRE DE MASSE POUR PRODUIRE DES MOLECULES POUVANT ETRE SEPARÉES ET ANALYSEES SELON LEUR RAPPORT M/z. (292)

Cette technique d'ESI a permis le développement des sciences « omiques » basées sur la spectrométrie de masse telles que la protéomique ou la métabolomique (273). Son impact significatif dans le domaine de la biochimie a d'ailleurs permis à Fenn de recevoir le prix Nobel de Chimie en 2002.

Analyse

Après leur formation, les ions sont concentrés et accélérés par des lentilles électroniques. Ils gagnent en énergie cinétique. Un analyseur de masse va ensuite les séparer selon leur rapport masse/charge (m/z). L'analyseur est caractérisé par sa précision en masse et par sa résolution en masse. La précision mesurée en parties par million (ppm) correspond à la mesure de la différence entre les rapports m/z théoriques et les rapports réellement mesurés pour une molécule donnée. La résolution est la capacité à distinguer deux molécules ayant des rapports m/z proches. Il existe plusieurs d'analyseurs de masse : l'analyseur temporel ou l'analyseur spatial.

Analyseur temporel

Analyseur à temps de vol (time of flight, TOF)

Cet analyseur repose sur le concept de temps de vol de la molécule. La vitesse de déplacement d'une molécule chargée dans un champ électromagnétique uniforme dépend de son rapport m/z (moindre accélération pour les ions les plus lourds). Une molécule lourde va se déplacer plus lentement qu'une molécule légère. Les molécules sont détectées de façon distincte en fonction de leur temps de vol.

Les analyseurs temps de vol permettent une analyse très rapide avec une vitesse de séparation de l'ordre de la microseconde. Au cours des dernières années des spectromètres de masse haute résolution à temps de vol de plus en plus performants ont été développés. Les analyseurs TOF-MS actuels possèdent une haute résolution, des vitesses d'acquisition élevées et une bonne sensibilité permettant l'obtention de spectres de masse riches en analyse ciblée comme en analyse non ciblée (293). L'analyseur TOF est souvent couplé à une source MALDI.

Dans cette technique, les ions quittent la source d'ionisation par un cône d'échantillonnage et rentrent dans une zone sous vide. Ils passent par un cône d'extraction puis par un guide d'ions à onde progressive (*Travelling-Wave Ion Guide* ou TWIG) qui permet leur focalisation avant l'entrée dans le quadripôle. Le quadripôle va servir de guide pour les ions en mode MS et de filtre en mode MS/MS. Ensuite, les ions rentrent dans la Triwave® Celle-ci se compose d'une cellule trappe de type TWIG, d'une cellule de mobilité ionique et d'une cellule de transfert de type TWIG. Par la suite, les ions parviennent à l'analyseur à temps de vol avec une vitesse réduite sous l'effet d'un champ électrique perpendiculaire au faisceau d'ions. Les ions vont par la suite se diriger vers un réflectron qui les renvoie vers le détecteur selon une trajectoire qui dépend des caractéristiques de précision et de

résolution de l'analyseur en temps de vol. Certains appareils permettent d'atteindre une précision inférieure à 3 ppm et une résolution en masse de 30 000 en routine.

Analyseurs spatiaux

La séparation des molécules chargées peut également être réalisée par passage à travers un champ électrique multidimensionnel et dynamique (analyseur quadripolaire, piège à ions) ou selon le mouvement périodique au sein d'un champ dynamique électromagnétique (analyseur Orbitrap et à résonance ionique cyclotronique).

Séparation par passage dans un champ électrique dynamique

Cette technique de séparation peut faire appel à un analyseur quadripolaire ou à un piège à ions.

Analyse quadripolaire : L'analyseur de type quadripôle comporte quatre électrodes fonctionnant par paires. Une tension continue et une tension alternative sont appliquées sur les électrodes ce qui crée un champ électrique dynamique. Une gamme pour le rapport m/z est prédéfinie par la valeur des tensions. Les molécules dont le rapport m/z est compris dans la gamme auront une trajectoire stable et pourront être détectées. Les molécules dont le rapport ne rentre pas dans la gamme auront une trajectoire instable et vont rentrer en collision avec les électrodes.

Linear Ion Trap ou piège à ions linéaire : Le piège à ions linéaire est formé de trois quadripôles. Son fonctionnement est très proche de celui de l'analyseur quadripolaire. L'entrée et la sortie du piège à ions sont soumises à une tension électrique continue ce qui crée un champ électrique emprisonnant les molécules ionisées dans la partie centrale. Dans cette partie, une tension alternative à haute fréquence crée un champ électrique radial. Le rapport m/z de la molécule détermine son mouvement au sein de ce champ. Les ions sont alors envoyés au détecteur ou détruits par collision. Cet analyseur linéaire peut stocker un grand nombre de molécules ionisées, et permet une acquisition rapide en haute résolution.

Le piège à ions 3 D : L'analyseur 3 D possède 3 électrodes : une électrode chapeau d'entrée et une électrode chapeau de sortie ; et une électrode annulaire. Les molécules ionisées subissent un champ électrique 3 D. Selon leur rapport m/z elles sont stables et sont détectées; ou elles sont instables et entrent en collision avec les électrodes. L'analyseur 3D a pour avantage sa petite taille avec des performances proches de celles de l'analyseur linéaire.

Séparation par mouvement périodique au sein d'un champ électrique/magnétique

Piège électrostatique Orbitrap® : L'analyseur Orbitrap® est une variante du piège à ions. Il est constitué de deux électrodes : une électrode creuse et une électrode coaxiale en fuseau. Les molécules ionisées vont avoir un mouvement oscillant et orbital autour de l'électrode coaxiale sous

l'effet d'un champ électrostatique. La fréquence de l'oscillation et la trajectoire des molécules dépendent de leur rapport m/z . Le détecteur intégré à l'analyseur permet l'élaboration d'un spectre de masse. Ce type d'analyseur permet de mesurer une large gamme de masse avec des résultats précis en haute résolution (294, 295).

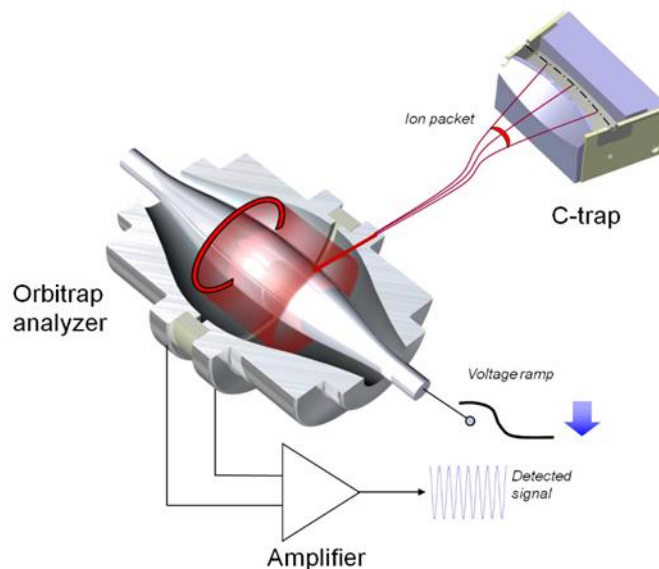


FIGURE 19 : PIÈGE À IONS INJECTEUR ET PIÈGE À IONS ÉLECTROSTATIQUE ORBITRAP®. LE TRAJET DES IONS EST SYMBOLISÉ PAR LE TRAIT ROUGE. ILLUSTRATION THERMO FISHER SCIENTIFIC.

Analyseur à résonance ionique cyclotronique : Cet analyseur se compose d'une cellule de résonance cyclotronique composée de deux plaques d'excitation, deux plaques de piégeage au sein d'un champ magnétostatique et deux plaques de détection. Les plaques d'excitation créent un champ électrique oscillant. Un champ magnétostatique orthogonal au champ électrique est appliqué. Les molécules ionisées sont piégées et vont avoir un mouvement circulaire uniforme cyclotronique. La fréquence cyclotronique dépend de $1/(m/z)$. Un signal électrique périodique est détecté et l'analyse du signal permet la lecture du spectre de masse. Ce type d'analyseur a les mêmes avantages que l'Orbitrap® (296, 297).

Détection

Après leur séparation selon leur rapport m/z , les molécules ionisées arrivent au détecteur qui va produire un signal électrique. Certains analyseurs contiennent leur propre module de détection (exemple de l'Orbitrap®) mais pour les autres, un détecteur doit être associé à l'analyseur. Le détecteur a pour principale fonction de compter les ions. Le détecteur le plus fréquemment utilisé est le multiplicateur d'électrons couplé à une dynode de conversion. L'énergie cinétique acquise par les ions est dirigée sur la dynode. Cela induit la production d'électrons qui vont être multipliés avec un gain de 10^6 - 10^7 . Pour les analyseurs utilisant la technologie de mouvement périodique dans un

champ électrique (Orbitrap®, résonance ionique cyclotronique), une transformée de Fourier permet la production d'un spectre de fréquence à partir du signal émis par les ions. Les analyseurs à temps de vol utilisent pour la détection un système de micro canaux qui multiplient les électrons.

La spectrométrie de masse en tandem

La combinaison de modules de séparation et d'analyse avec des capacités de fragmentation des molécules ionisées ont permis d'augmenter la performance des appareils de spectrométrie de masse avec des informations de plus en plus précises sur les molécules étudiées. La spectrométrie de masse en tandem (MS/MS) est réalisée avec des appareils contenant deux modules couplés de séparation/analyse (**Figure 20**). Ces appareils associent alors une grande sensibilité de détection et une précision de mesure pour l'identification et la quantification.

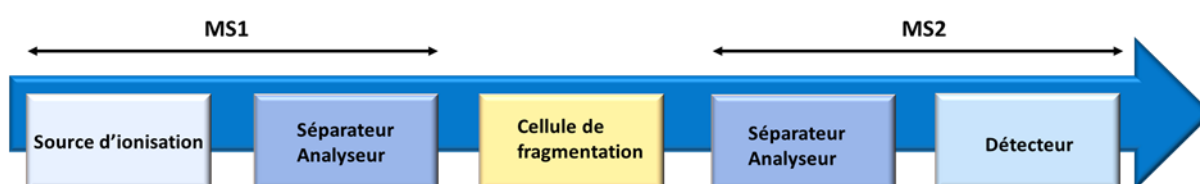


FIGURE 20 : REPRESENTATION SIMPLIFIEE D'UN SPECTROMETRE DE MASSE EN TANDEM.

4.2.5 Traitement des données

4.2.5.1 Préparation de l'analyse statistique

Le traitement des données constitue une partie majeure du travail des approches « omiques ». Les informations obtenues par l'analyse métabolomique sont multiples et des logiciels de traitement des données sont indispensables pour les étudier (298, 299). Plusieurs logiciels sont commercialisés ou disponibles en accès libres. Les données brutes doivent dans un premier temps être nettoyées. La suppression du bruit de fond se fait par des algorithmes de filtrage qui sont plutôt homogènes quel que soit le logiciel. En revanche, des différences sont observées en ce qui concerne les étapes d'extraction et d'alignement des signaux (300).

Certains logiciels permettent d'évaluer la qualité de l'extraction des données. L'homogénéité des échantillons est évaluée par le coefficient de variation des pics dans les répliqués techniques (plusieurs passages du même échantillon dans le spectromètre). Le pourcentage de données manquantes, la précision de la mesure de masse, les effets de dérive inter-séries, le rapport signal sur bruit sont habituellement utilisés pour étudier la reproductibilité analytique. La qualité des données obtenues est également évaluée par l'utilisation d'échantillons de contrôle qualité.

Les données brutes sont constituées de vecteurs de points acquis au cours du temps. Chaque point correspond à un rapport m/z , une intensité, une section efficace de collision et un temps de

rétenion. Plusieurs systèmes de filtrage sont disponibles pour la suppression du bruit de fond. Ceux-ci doivent veiller à préserver les informations pertinentes.

Les données brutes continues sont ensuite converties sous forme de pic. Chaque pic représente un ion. L'étape suivante consiste à calculer les centroïdes des pics dans la gamme m/z , puis à rechercher les temps de rétention des pics chromatographiques et/ou du spectre de mobilité ionique dans toute la gamme. Les données issues des différents échantillons sont comparées après avoir groupé les pics et aligné les temps de rétention/dérive. En effet, des fluctuations liées à la technique peuvent survenir au cours de la manipulation. Cette étape d'alignement est donc indispensable pour obtenir des données interprétables. Ainsi, le même ion est comparable entre les différents échantillons. Pour corriger le temps de rétention, une des possibilités est l'ajout d'une molécule repère de référence (standard interne) dans l'échantillon afin d'aligner les pics (277). La molécule de référence est soigneusement choisie afin d'être dans la gamme du temps de rétention. Néanmoins, un excès de molécules d'étalonnage peut également supprimer certaines données du spectre. Dans l'idéal, les méthodes d'alignement ne faisant pas appel à un composé de référence sont à privilégier. Certains logiciels utilisent les résultats de détection de pics pour faire correspondre les pics similaires sans utiliser de molécule de référence. La procédure est réalisée deux ou trois fois. Un résultat sera considéré comme reproductible si la dérive du temps de rétention est assez corrigée.

L'étape suivante consiste à identifier les métabolites à partir des données analytiques acquises. Cela permet la traduction de données en information biologique (301). L'utilisation conjointe de pièges ioniques permet l'obtention d'informations structurales complémentaires facilitant la reconnaissance des molécules (302). L'identification des métabolites fait appel à l'interrogation de bases de données numériques. De nombreuses bases ont ainsi été développées afin de faciliter leur identification. Ces bases de données stockent des informations sur les propriétés physico-chimiques et biologiques des composés. D'autre part, le site OMICtools répertorie les logiciels disponibles pour l'analyse des données métabolomiques. Les bases de données vont permettre de lister les molécules connues qui ont un m/z correspondant à celui de la molécule à annoter. Cependant, une identification avec certitude nécessite de caractériser cette molécule en réalisant par exemple une analyse de ce pic par MSMS et confirmer ainsi que le nom choisi dans la liste correspond bien à celui de la molécule à annoter.

A l'issue de cette étape certains métabolites vont être identifiés alors que d'autres resteront inconnus ne correspondant à aucun métabolite connu dans les bases de données numériques. Cela constitue une des limites de l'analyse non ciblée.

4.3 Analyse des données

4.3.1 Analyse univariée

Cette étape a pour objectif de réduire la dimension des données acquises par un système de filtrage des variables pour ne conserver que les données les plus pertinentes pour l'analyse statistique ultérieure. Les variables sont classées et sont filtrées en fonction de leur significativité statistique (303). Classiquement, un test de Student avec un facteur correctif des tests multiples et un seuil de 0,05 % est utilisé.

4.3.2 Analyse multivariée

Les analyses métabolomiques génèrent un grand nombre de données quantitatives et qualitatives. L'utilisation de logiciels de bioinformatique et biomathématiques est nécessaire pour répondre au caractère multidimensionnel des données. Les analyses statistiques multivariées permettent l'extraction et l'interprétation d'informations biologiques complexes (292).

4.3.2.1 Approche non supervisée : Analyse multivariée descriptive

4.3.2.1.1 Analyse en composante principale (ACP)

Cette méthode a pour but de comprimer et d'explorer les données de façon globale. Chaque échantillon correspond à un point. Des variables définissent ce point au sein d'un espace. Chaque variable permet de définir les coordonnées d'une observation. L'ACP a pour objectif de mettre en évidence les directions de plus grande dispersion des observations dans l'espace. Ces directions sont celles qui contiennent le plus de variance et donc le plus d'information. Aucune hypothèse n'est faite sur les liens possibles entre les échantillons et les variables. Cette analyse est donc non supervisée.

La décomposition de la matrice de l'analyse en composante principale aboutit à des matrices des contributions factorielles (loadings) et des coordonnées factorielles (scores) à partir des données originales (**Figure 21**). D'un point de vue géométrique, l'ACP est une méthode de rotation des données faisant en sorte que l'observateur soit le mieux placé pour évaluer les liens entre les individus. Les scores peuvent placer les échantillons sur des plans construits à partir des composantes principales. Cela permet de dépister la formation de groupes ou la présence d'observations aberrantes ; et de rechercher des répartitions structurées des objets.

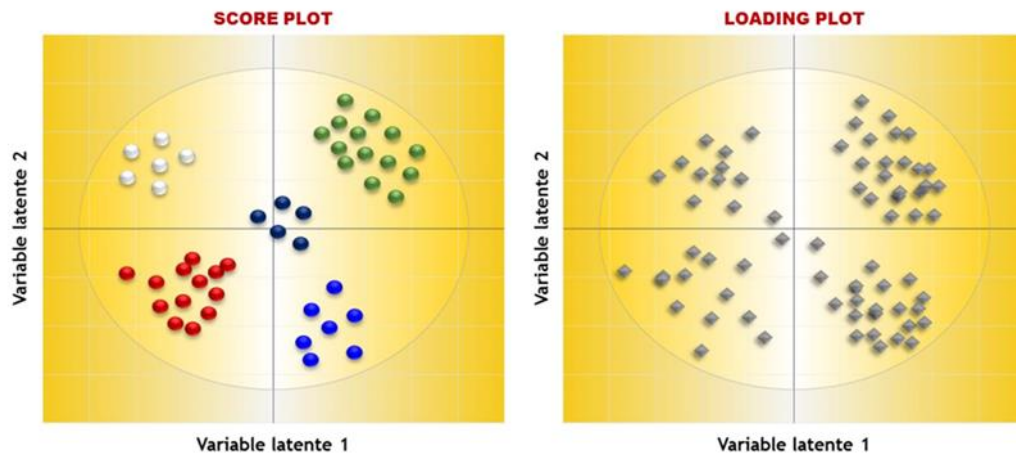


FIGURE 21 : EXEMPLE DE COORDONNEES FACTORIELLES (SCORE PLOT) ET CONTRIBUTIONS FACTORIELLES (LOADING PLOT) OBTENUES A PARTIR D'UNE MATRICE DE DONNEES. LES DIAGRAMMES DE DISPERSION DOIVENT ETRE INTERPRETES SIMULTANEMENT POUR INTERPRETER LES TENDANCES DE REGROUPEMENT DES VARIABLES OU LA PRESENCE D'OBSERVATIONS ABERRANTES.

4.3.2.2 Approche supervisée

4.3.2.2.1 *Partial Least Squares Regression Discriminant Analysis (PLSDA) and Orthogonal Partial Least Square Discriminant Analysis (OPLSDA)*

Le but de cette analyse est de construire des modèles prédictifs basés sur une régression multivariée. Les liens entre deux matrices peuvent être analysés en utilisant la Partial Least Square regression (PLS) à travers un modèle linéaire multivarié. Avec cette technique, des variables colinéaires incomplètes ou bruitées peuvent être analysées (304).

Les variables prédictives (X) sont généralement des spectres. Elles peuvent également être des mesures physico-chimiques. Les réponses (Y) sont de nature variée (concentration des substances, propriétés physico-chimiques, activités biologiques). En biologie, la réponse est souvent catégorielle définissant l'appartenance à un groupe (témoins sains/malades, traités/non traités). L'analyse PLSDA doit discriminer les variables les plus riches en information permettant de séparer les groupes observés. La figure 22 montre un score plot de PLSDA (**Figure 22**). L'OPLSDA (Orthogonal Partial Least Square Discriminant Analysis) est une variante de la PLS. La variation de X (expérimentale ou biologique) non corrélée avec Y est éliminée. Cela permet d'optimiser les performances prédictives des modèles statistiques générés (305).

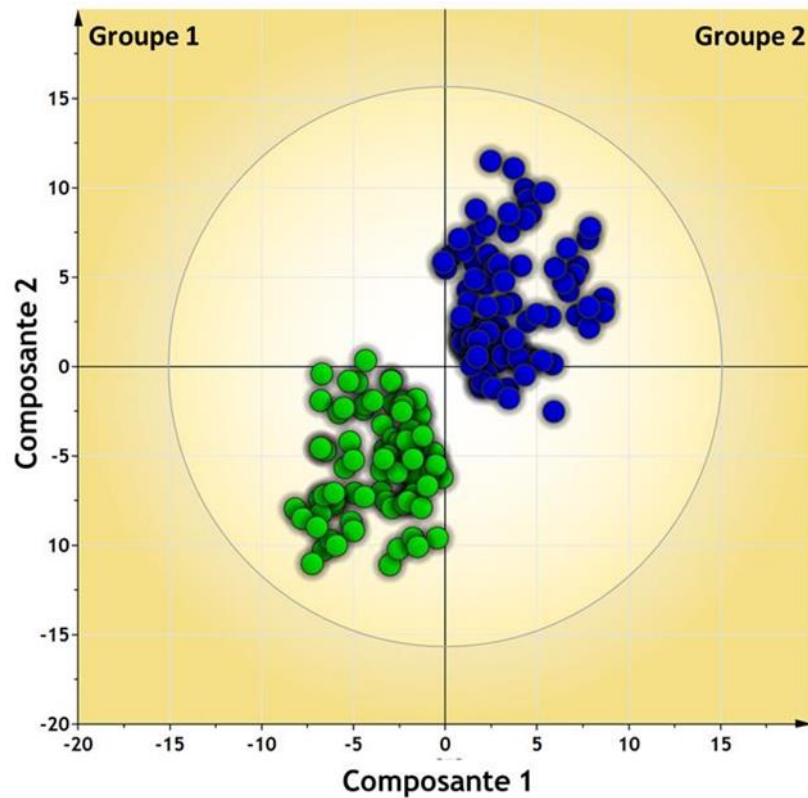


FIGURE 22 : REPRESENTATION PLSDA MONTRANT LA SEPARATION DE DEUX GROUPES (GROUPE 1 ET GROUPE 2)

4.3.3 Validation des modèles

La validation des modèles générés est primordiale pour l'analyse de données permettant sa fiabilité prédictive ultérieure.

4.3.3.1 Validation croisée

La première étape est la validation croisée permettant de définir la dimensionnalité optimale du modèle en définissant le nombre optimal de variables latentes. En augmentant le nombre de variables, les modèles sont de plus en plus précis. Si la *Partial Least Squares* est utilisée, le but est de se rapprocher au maximum des valeurs de la variable réponse et de limiter les écarts par rapport à la prédiction.

La variabilité intrinsèque de l'échantillon (bruit) non informative ne doit pas être incluse dans le modèle car elle diminue la performance des prédictions. A contrario, un modèle avec un nombre de variables latentes inférieur à ce qu'il faut pour inclure toute la variabilité pour limiter les erreurs de prédiction n'est pas idéal non plus. L'utilisation de la *Partial Least Squares* nécessite un équilibre judicieux entre un bon pouvoir de prédiction et l'extraction de la variabilité des matrices. Elle dépend fortement du nombre de variables latentes à inclure dans le modèle.

La validation croisée estime l'exactitude des prédictions pour un modèle avec un nombre de variables latentes croissant. La matrice est ainsi séparée en deux parties : une partie (échantillons d'étalonnage) sert à mettre au point des modèles d'étalonnage avec plusieurs variables latentes différentes et une autre partie (échantillons de validation) sert à évaluer les erreurs de prédiction.

Après cette validation croisée, la PRESS (*Predicted Residual Error Sum of Squares*) qui correspond à la somme des carrés des erreurs résiduelles de prédiction est calculée. L'objectif pour la *Partial Least Squares* est d'avoir la plus petite valeur de PRESS possible.

4.3.3.2 Test de permutation

Pour estimer la confiance sur un modèle de régression dans le cadre de la PLS, un test de permutation est applicable. Ce test va modifier au hasard le positionnement des valeurs de y et établir le modèle entre X et le vecteur Y sur plusieurs échantillons d'étalonnage. La PRESS est calculée pour tous les couples étalonnage/prédiction. Le PRESS du vrai modèle de validation croisée est alors comparé avec la distribution des erreurs des modèles perturbés.

4.3.3.3 CV ANOVA

Le coefficient de variation ANOVA est un test de significativité de l'hypothèse nulle de résidus égaux de deux modèles comparés. Il évalue la fiabilité de la PLS et de l'OPLS en recherchant une variabilité significative des résidus prédictifs.

4.3.3.4 Validation externe

Les méthodes de validation croisée interne permettent d'évaluer la qualité du modèle mais une validation externe permet de renforcer la fiabilité des données obtenues. Les échantillons utilisés pour la validation externe ne doivent pas avoir été utilisés pour la validation du modèle.

4.3.4 Evaluation des variables discriminantes

Après avoir validé le modèle par les techniques précédemment décrites, l'enjeu est de sélectionner les variables discriminantes pour répondre à l'objectif de l'étude.

Cette stratégie peut utiliser les tests statistiques les plus courants univariés mais peut également faire appel à des analyses multivariées. En métabolomique, l'objectif est de sélectionner des molécules pouvant être utilisées comme biomarqueurs et d'évaluer la puissance prédictive du modèle construit.

Le processus de filtrage permet de classer les sous-ensembles de variables avant de former les modèles prédictifs. La technique de modélisation est alors répétée en prenant les métabolites les plus prédictifs.

Certains logiciels permettent de sélectionner de façon intuitive les variables potentiellement discriminantes. Les variables choisies sont alors classées selon leur pouvoir discriminant. Le VIP (*Variable Importance*) est un coefficient qui représente le lien entre les variables X et Y. Plus le VIP est élevé, plus la significativité de la différence d'intensité de la molécule ionisée entre les échantillons étudiés est importante.

OBJECTIFS DE LA THESE

Les maladies lysosomales représentent une cinquantaine de maladies ayant en commun des symptômes dits de « surcharge » avec une atteinte multiviscérale liés à un défaut de dégradation de macromolécules ou de composés monomériques dans les organelles du système endosome/lysosome. Certaines de ces pathologies sont accessibles à des traitements spécifiques mais ces thérapeutiques sont uniquement symptomatiques et ne guérissent pas les patients. Même si le phénomène de surcharge peut expliquer entre autres la symptomatologie observée, la physiopathologie de ces maladies est complexe et non précisément connue. De nombreux processus biologiques tels que l'autophagie ou l'inflammation semblent impliqués.

La médecine de précision également appelée médecine personnalisée a pour objectif une prise en charge du patient adaptée à ses caractéristiques personnelles. Elle utilise les informations obtenues par des études biologiques individuelles mais aussi populationnelles. Les sciences « omiques » qui se sont développées sur les dernières années ont permis l'émergence de cette médecine personnalisée. Les sciences « omiques » permettent d'extraire des données biologiques à partir d'un système biologique de façon globale avec une approche ciblée ou non ciblée. Ces nouvelles techniques sont différentes des approches conventionnelles qui sont réductionnistes et séquentielles.

Un des objectifs de cette thèse était d'appliquer des techniques « omiques » à deux groupes de maladies lysosomales afin d'explorer leurs mécanismes physiopathologiques.

La première partie du travail a été appliquée aux mucopolysaccharidoses (MPS). L'étude s'est intéressée à la recherche de profils métaboliques spécifiques des mucopolysaccharidoses par des techniques de métabolomique non ciblée puis ciblée. La mise en évidence de nouveaux biomarqueurs spécifiques des MPS pourrait en effet améliorer le diagnostic, le suivi et la prise en charge thérapeutique de ces pathologies complexes.

La deuxième partie de ce travail a été réalisée sur la maladie de Fabry. Différentes techniques « omiques » ont été utilisées : transcriptomique, protéomique, métabolomique. Comme pour les mucopolysaccharidoses, l'objectif était de rechercher de nouveaux biomarqueurs spécifiques pour le diagnostic et le suivi de la maladie mais aussi de mieux comprendre les processus biologiques et les voies métaboliques impliqués dans la pathologie.

L'autre objectif de ce travail de thèse était indépendant des techniques « omiques » avec une approche plus mécanistique visant à étudier le processus d'autophagie dans la maladie de Fabry. Des premiers travaux menés au sein de l'équipe ont montré une altération du flux autophagique avec l'accumulation de vacuoles autophagiques dans des cellules de patients atteints de la maladie de Fabry. Nous avons donc étudié les effets de la modulation du processus autophagique sur le devenir

et l'activité de l'enzyme substitutive dans des cultures cellulaires. L'accumulation du Gb3 et l'adressage de l'enzyme au lysosome après inhibition du processus d'autophagie ont été explorés.

Pour la partie « omiques » mon travail a consisté à aider au recueil et à la préparation des échantillons d'urines pour la spectrométrie de masse haute résolution pour la partie MPS ; réaliser les cultures cellulaires, réaliser les études de protéomique sur les échantillons de cellules et de plasma et les interpréter pour la partie Fabry. Les études de métabolomique ont été réalisées en collaboration avec un autre laboratoire.

Pour la partie « mécanistique » dans la maladie de Fabry, j'ai réalisé les cultures cellulaires, les dosages de Gb3 et les études d'immunohistochimie pour l'étude de l'adressage de l'enzyme au lysosome.

TRAVAIL EXPERIMENTAL

I. Présentation de l'étude METALYS

1.1 Objectifs

Le projet METALYS a pour objectif de rechercher un profil "omique" spécifique chez des patients atteints de maladies lysosomales par comparaison à un groupe contrôle. Nous avons pour cela collecté des échantillons provenant de patients atteints de différentes maladies lysosomales et notamment de mucopolysaccharidoses et de maladie de Fabry.

1.2 Analyse métabolomique des mucopolysaccharidoses par spectrométrie de masse haute résolution

1.2.1 Collecte des échantillons

Des échantillons urinaires de patients atteints de mucopolysaccharidoses appartenant à des collections biologiques déclarées dans des centres experts (Necker-Enfants Malades, Trousseau, Lille, Amiens, Lyon et Rouen) et 66 échantillons urinaires contrôles ont été analysés (**Figure 23**).

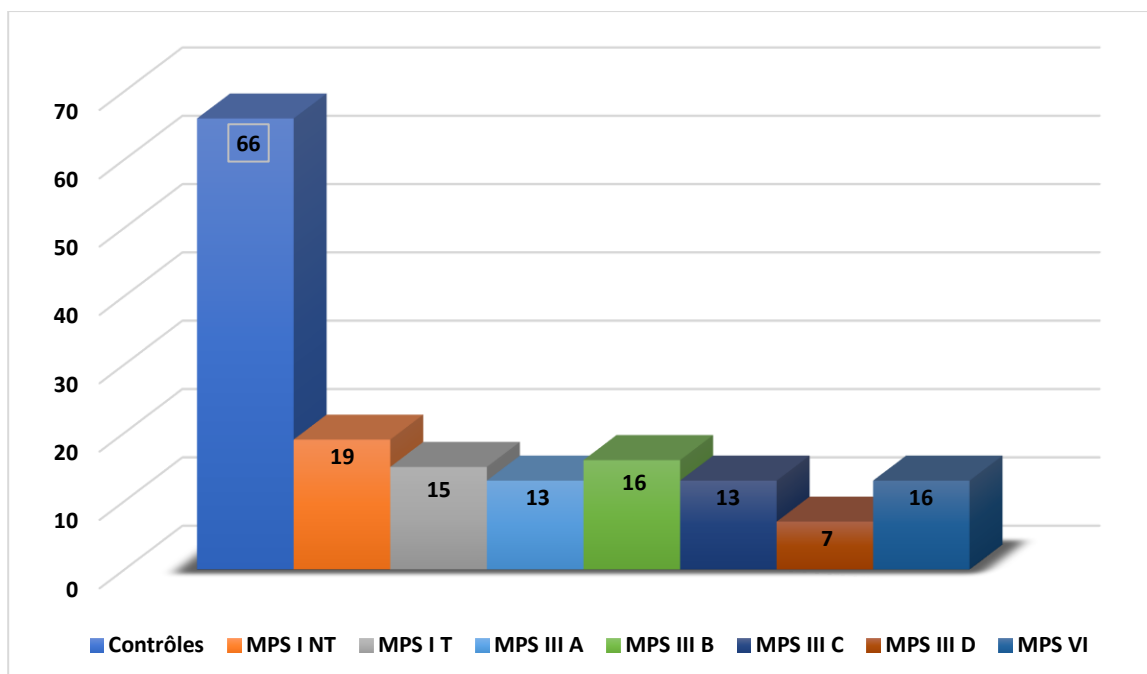


FIGURE 23 : REPARTITION DU NOMBRE D'ÉCHANTILLONS D'URINES SELON LE TYPE DE MUCOPOLYSACCHARIDOSE.

MPS I NT= MPS TYPE I NON TRAITES, MPS I T= MPS I TRAITES.

1.2.2 Aspects éthiques de l'étude

Le projet a été soumis à l'avis du CERNI du CHU de Rouen (CERNI E2016-21) et a reçu un avis favorable. La recherche a été effectuée de façon non interventionnelle à partir de collections biologiques déjà constituées et déclarées auprès du ministère de la recherche.

1.2.3 Partenaires

Ce travail a bénéficié de l'expertise du Service de Biochimie Métabolique - CHU de Rouen, du Laboratoire COBRA - UMR 6014 -IRCOF, Mont-Saint-Aignan, et de l'équipe "Genetics and Physiopathology of Neurodevelopmental Disorders" INSERM U1245 Equipe 4 de l'université de Rouen. Les échantillons ont été recueillis dans les centres de référence et de compétence des maladies lysosomales partenaires du projet (Amiens, Lyon, Lille, Paris Necker-Enfants Malades et Trousseau).

1.2.4 Résultats et publications

L'analyse métabolomique non ciblée de ces échantillons a été réalisée par spectrométrie de masse haute résolution associée à la chromatographie liquide haute performance et la mobilité ionique (SYNAPT G2 HDMS, Waters MS Technologies, Manchester, Royaume-Uni). Cette technique a été mise au point par notre équipe et a fait l'objet d'une publication spécifique (306). L'analyse des données brutes a été réalisée via Progenesis Q1 (Waters MS Technologies, Manchester, UK).

L'approche métabolomique ciblée a consisté à analyser 24 acides aminés et les glycosaminoglycanes (métabolites accumulés dans les mucopolysaccharidoses) par spectrométrie de masse en tandem.

La quantification ciblée de certains métabolites a été réalisée en utilisant la chromatographie phase liquide couplée à la spectrométrie de masse en tandem ; et parfois avec l'aide de kits spécifiques d'un type de métabolite comme pour les acides aminés par exemple. L'analyse des voies métaboliques a été réalisée via le logiciel Mummichog pour chaque groupe d'échantillons afin d'identifier les mécanismes métaboliques impliqués.

Les résultats sont détaillés dans les articles suivants (trois publications originales et une revue) :

1. Tebani A, Schmitz-Afonso I, **Abily-Donval L**, Héron B, Piraud M, Ausseil J, Brassier A, De Lonlay P, Zerimech F, Vaz FM, Gonzalez BJ, Marret S, Afonso C, Bekri S. **Urinary metabolic phenotyping of mucopolysaccharidosis type I combining untargeted**

and targeted strategies with data modeling. *Clin Chim Acta.* 2017 Dec;475 :7-14.(307)

2. Tebani A, **Abily-Donval L**, Schmitz-Afonso I, Héron B, Piraud M, Ausseil J, Zerimech F, Gonzalez B, Marret S, Afonso C, Bekri S. **Unveiling metabolic remodeling in mucopolysaccharidosis type III through integrative metabolomics and pathway analysis.** *J Transl Med.* 2018 Sep 4;16(1) :248.(308)
3. Tebani A, **Abily-Donval L**, Schmitz-Afonso I, Piraud M, Ausseil J, Zerimech F, Pilon C, Pereira T, Marret S, Afonso C, Bekri S. **Analysis of Mucopolysaccharidosis Type VI through Integrative Functional Metabolomics.** *Int J Mol Sci.* 2019 Jan 21;20(2).(309)
4. Tebani A, **Abily-Donval L**, Afonso C, Marret S, Bekri S. Clinical Metabolomics : The New Metabolic Window for Inborn Errors of Metabolism Investigations in the Post-Genomic Era. *Int J Mol Sci.* 2016 Jul 20;17(7).

1.2.5 Conclusion

L'analyse métabolomique a permis de mettre en évidence des profils métabolomiques spécifiques des différents types MPS étudiés *versus* contrôle. L'analyse des voies métaboliques relatives à ces biomolécules a permis de mise en évidence la contribution du stress oxydant et d'une altération de l'autophagie aux bases physiopathologiques des maladie lysosomales.



Review

Clinical Metabolomics : The New Metabolic Window for Inborn Errors of Metabolism Investigations in the Post-Genomic Era

Abdellah Tebani ^{1,2,3}, Lenaig Abily-Donval ^{2,4}, Carlos Afonso ³, Stéphane Marret ^{2,4} and Soumeya Bekri ^{1,2,*}

¹ Department of Metabolic Biochemistry, Rouen University Hospital, Rouen 76031, France; abdellah.tebani@chu-rouen.fr

² Normandie Univ, UNIROUEN, INSERM, CHU Rouen, IRIB, Laboratoire NeoVasc ERI28, Rouen 76000, France; lenaig.abily-donval@chu-rouen.fr (L.A.-D.); stephane.marret@chu-rouen.fr (S.M.)

³ Normandie Univ, UNIROUEN, INSA Rouen, CNRS, COBRA, Rouen 76000, France; carlos.afonso@univ-rouen.fr

⁴ Department of Neonatal Pediatrics and Intensive Care, Rouen University Hospital, Rouen 76031, France

* Correspondence : soumeya.bekri@chu-rouen.fr; Tel. : +33-2-3288-8124; Fax : +33-2-3288-8341

Academic Editor : Ting-Li (Morgan) Han

Received : 11 June 2016; Accepted : 15 July 2016; Published : 20 July 2016

Abstract : Inborn errors of metabolism (IEM) represent a group of about 500 rare genetic diseases with an overall estimated incidence of 1/2500. The diversity of metabolic pathways involved explains the difficulties in establishing their diagnosis. However, early diagnosis is usually mandatory for successful treatment. Given the considerable clinical overlap between some inborn errors, biochemical and molecular tests are crucial in making a diagnosis. Conventional biological diagnosis procedures are based on a time-consuming series of sequential and segmented biochemical tests. The rise of “omic” technologies offers holistic views of the basic molecules that build a biological system at different levels. Metabolomics is the most recent “omic” technology based on biochemical characterization of metabolites and their changes related to genetic and environmental factors. This review addresses the principles underlying metabolomics technologies that allow them to comprehensively assess an individual biochemical profile and their reported applications for IEM investigations in the precision medicine era.

Keywords : metabolomics; inborn errors of metabolism; screening; diagnosis; systems medicine; precision medicine

1. Introduction

The new field of precision medicine is revolutionizing current medical practice and reshaping future medicine. Precision medicine aspires to put the patient as the central driver of healthcare by broadening biological knowledge and acknowledging the greater diversity of individuals [1]. It is well established that complex gene–environment interactions shape normal physiological and disease processes at both the individual and population scale. Predicting normal and pathological states in patients requires dynamic and systematic understanding of these interactions. Systems medicine is a new concept based on holistic approaches for disease diagnosis and monitoring. The basic idea of these approaches is that a complex system is more comprehensively understood if considered as a whole at both the spatial and temporal scales.

Inborn errors of metabolism (IEM) are an appropriate model for systems medicine studies because the biological basis underlying these diseases has been, at least partly, revealed. IEM represent a group of about 500 rare genetic diseases with an overall estimated incidence of 1/2500. Even though these disorders are individually rare, they are collectively more common and cause a significant childhood morbidity and mortality. IEM are genetic disorders resulting from defects in a given biochemical pathway due to the deficiency or abnormality of an enzyme, its cofactor, or a transporter, leading to an accumulation of a substrate or lack of the product. Hence, the diversity of metabolic pathways involved explains the difficulties in establishing a diagnosis. Autosomal recessive transmission is most frequent, but autosomal dominant and X-linked disorders have also been described. IEM may involve mutations in mitochondrial DNA. The pathogenesis of IEM can be explained by mechanisms such as deficiency of an essential product or enzyme, systemic toxic effects of circulating metabolites, and activation or inhibition of alternative metabolism [2]. Based on these pathophysiological traits, several IEM therapies have been developed, including dietary restriction, toxic product clearance, or biotherapies (enzyme replacement and gene therapy) [3]. Initiating these treatments at birth or at early stages is usually mandatory for optimal patient management. The first description of these disorders was made by Sir Archibald Garrod [4], who initiated the “one gene–one disease” paradigm. However, there is a lack of genotype–phenotype correlation in IEM. Furthermore, for the same genetic variation, different phenotypes have been observed in the same family [2]. These observations challenge Garrod’s paradigm and suggest the influence of either genetic or environmental modifying factors. Thus, IEM are more than monogenic diseases, which adds another layer of complexity to disease characterization and diagnosis.

The rise of “omic” approaches, enabled by the tremendous technological shift in both multiscale biological information capture and data management, offers an amazing opportunity to provide new effective tools for screening, diagnosis, treatment, and monitoring of these diseases. Omic technologies offer global views on the basic molecules that build a biological system at the cell, tissue, or organism level. Primarily, they aim to recover, in an untargeted, unbiased, and hypothesis-free fashion, the biological information carried by genes (genomics), mRNA (transcriptomics), proteins (proteomics), and metabolites (metabolomics). These holistic strategies clearly contrast with conventional studies,

which are mainly hypothesis-driven and reductionist. To truly understand disease processes, a global investigative approach needs to be applied at multiple biological informational levels.

Since the early days of medicine, the human body is viewed as a collection of separate and independent components, and thus, physicians typically treated disease by trying to identify the single abnormality related to a single component. This approach lacks contextual information which is vital for mechanistic understanding of pathophysiology and, thus, for designing treatment strategies [5,6]. Indeed, the complete characterization of a biological system should include a structural, an organizational pattern and a functional description [7]. The structure comprises the fundamental actor components (genes, proteins and metabolites). The organization pattern denotes how these actors are linked to each other and how they are organized topologically (e.g., linear or branched sequence of reactions) and morphologically (membrane-bound or functional compartmentalization). The function describes how the whole system behaves in space and time with regard to metabolic fluxes and response to stimuli [8–10].

Systems biology is a new scientific field that tries to achieve this systemic understanding of biology and to fill in the gap between information and context from a biological standpoint. Systems biology can be defined as a holistic and systemic analysis of complex system inter-connections and their functional interrelationships [11–14]. Two vital pillars supported the emergence of systems biology : data generation and data modeling. On the one hand, the surge of high-throughput omics technologies allowed the retrieval of a global and comprehensive biological information. On the other hand, the amazing development of computational capabilities allowed complicated systems modeling and convenient and intuitive visualization. Furthermore, these informatics advancements are crucial for comprehensive integration and insightful interpretation of the complex biological information [15–17]. The patient-centric approach is essential to achieve the promise of personal and stratified medicine. Indeed, unlike conventional medical biology practice based primarily on

sequential studies of genes, proteins, and metabolites, the great challenge of modern biology is to apprehend a disease as a complex, integrated, and dynamic network. The dynamic view refers to the quantitative and qualitative assessment of changes and interactions between the different layers of the biological information [7,18–21]. The genetic classifications of disease are now well established, given the modern genomic tools that can provide rich information about large patient cohorts. However, other highly complementary approaches based on proteomic and metabolic information can help researchers to biochemically or physiologically contextualize the underlying genetic information, thus helping to get closer to the phenotype and allowing patient stratification [22]. Thanks to disruptive technological jumps, a revolutionary vision was pioneered by Lee Hood, who coined the term P4 medicine [19],

which is aimed to be predictive, preventive, personalized, and participatory. This new shift defines a new healthcare strategy in which each person serves as his or her own control over time [23].

The omics surge presents an amazing opportunity to provide new innovative tools for rapid diagnosis of IEM. Furthermore, metabolomics approaches are relevant for IEM because their basic pathophysiology is tightly related to metabolism. These diseases present with nonspecific clinical symptoms and appropriate laboratory tests are crucial in making a diagnosis. However, conventional biological diagnosis procedures are based on a series of sequential and segmented biochemical tests on various separated analytical platforms. This approach is slow, time-consuming, and complex, whereas optimal patient management requires improved speed of biochemical tests to allow early diagnosis and better monitoring of IEM. To address this need of faster screening and diagnosis strategies, metabolic profiling is a promising candidate.

In this review, we describe basic principles underlying metabolic phenotyping and metabolomic approaches that can be used to comprehensively assess an individual biochemical profile and their reported applications in IEM. Data for this review were identified by searches of PubMed and references from relevant articles using the search terms “metabolomics”, “metabonomics”, “metabolic profiling”, “inborn errors of metabolism”, and “inherited metabolic diseases”.

2. Metabolomics

2.1. Metabolites and Metabolome

The idea behind metabolomics goes back to ancient Greece, where doctors used the organoleptic characteristics of urine to link them to different medical conditions. Urine sweetness has been used to detect high glucose in diabetes [24]. Such organoleptic features are, of course, metabolic in origin. The word metabolome was coined by Olivier et al. in 1998 and defined as the set of metabolites synthesized by an organism [25]. Metabolome refers to the comprehensive complement of all metabolites present in a given biological system, fluid, cell, or tissue [26]. Metabolites can be defined as organic small molecules involved in enzymatic reactions. Thus, metabolomics is one of the “omic” technologies based on biochemical and molecular characterizations of the metabolome and the changes in metabolites related to genetic, environmental, drug or dietary, and other factors.

Metabolomics allows researchers to characterize these interactions and to evaluate the biochemical mechanisms involved in such changes in a systematic fashion. Indeed, metabolites fulfill the key criterion in that they change rapidly in response to physiological changes and may generate vital information about biochemical pathways that are modified in patients and in treated patients. Hence, metabolic profiling is highly informative since metabolites act as substrates or products in biochemical metabolic pathways [22,27–29].

Metabolomics has found applications in many disease studies and in complex interacting systems [22]. The possibility of predicting drug effects from baseline metabolic profiles has been demonstrated and gave rise to pharmacometabonomics as a potential effector for patient stratification and personalized medicine [30–35]. It is possible that the future of IEM diagnosis may be found in the developing area of metabolomics by doing simultaneous quantitative metabolic profiling of many metabolites in biological fluids.

2.2. Analytical Strategies and Chemical Information Retrieval

2.2.1. Biological Samples

For biological information recovery, metabolomics generally uses biofluids, cells, or tissue extracts as primary sources of metabolic fingerprint data. Compared with intact or extracted tissues, urine and serum or plasma are the most commonly studied biofluids in clinical practice, because they are easily obtained and prepared [36–39]. However, other specialized fluids could be used, including cerebrospinal fluid [40,41] or saliva [42–44] and even breath [45,46]. Dried blood (and other biofluids) spots samples (DBS) have also been investigated [47–50] and were shown to be an interesting alternative to conventional liquid samples for generating metabolite profiles. Given their practical advantages such as low volume, low cost, and handling convenience, DBS is gaining interest as a sampling support for metabolic profiling in IEM [47,51–53]. Of note, most metabolomics studies, particularly in clinical metabolomics, include data from a single biofluid, most often blood or urine. However, biochemical signature in a biofluid denotes complex interrelationships from different organs, which add another complexity layer for metabolomics data interpretation. This could be only understood by investigating pathophysiological states from a metabolic interactions perspective taking into account local metabolome specificities and their contribution to systemic metabolome. Different data-driven approaches have been described to handle these issues using multiple biofluids sampling and metabolomics data modeling [54,55].

2.2.2. Analytical Technologies

The human metabolome is a complex, highly responsive, and dynamic system. Thus, it raises different analytical challenges compared to other omics analysis approaches that are based on profiling large molecules built with a simple and limited set of subunits, such as nucleotides for genomics and transcriptomics and amino acids for proteomics. Thus, for identification and functional analysis of DNA, RNAs and proteins, the order combination of the subunits is what matters. It is the order of subunits that embodies the observed complexity that carries the biological information. Sequencing technologies rely basically on an incremental detection of these subunits. Researchers must figure out the order of the subunits to decode the carried biological information [56]. However, the same sequencing approach cannot be used to analyze metabolites in complex biofluids, because the analytical challenge is not simply to crack the order code, as there is no obvious order. To retrieve the metabolic information, the metabolome requires a more complex analysis of chemical mixtures that allows components to be individually and selectively differentiated, identified and measured across a wide qualitative and quantitative chemical space.

The diversity of the physicochemical properties of the various metabolites groups adds another layer of complexity to metabolomics studies. This supplemental challenge has been the key driver for the development of various analytical protocols and platforms. Indeed, scientists tackled this analytical challenge even before the term metabolomics was coined. The first scientific article about metabolomics was published by Pauling and colleagues, in which they described a method using gas chromatographic separation with flame ionization detection to analyze the breath [57]. The authors referred to orthomolecular medicine linking the detected biochemical signature to phenotypes.

Since then, huge development has been made. The mainly used metabolic profiling technologies are nuclear magnetic resonance (NMR) spectroscopy [58–60] and mass spectrometry (MS), either combined or not to a gas phase or liquid phase separation method [27,51]. These technologies are suitable for metabolomics studies because they deliver global, unbiased, and comprehensive chemical information from complex mixtures. For information recovery, the multivariate spectroscopic data produced are typically analyzed using chemometric techniques to identify informative metabolic combinations that can be used for either sample classification or global biomarker discovery [51,61,62]. NMR spectroscopy is rapid and nondestructive and has the advantage of being highly reproducible.

It is a powerful spectroscopic technology that offers atom-centered information that is crucial for molecular structure elucidation [63]. High-resolution NMR using stronger magnetic fields or two-dimensional NMR allows higher information recovery. The major drawback of NMR is its lack of sensitivity. However, MS offers

complementary molecular information and is, by far, more sensitive than NMR. Hence, it allows higher metabolome coverage. The use of separation methods coupled to MS, such as liquid chromatography [38,39], gas chromatography [64], or capillary electrophoresis [65], allows a molecular separation step before MS detection. This enhances sensitivity and the dynamic range and provides complementary molecular information using the separation dimension.

Recently, approaches using another gas phase separation, ion mobility spectrometry (IMS) [66], has been gaining interest in metabolomics [67–73]. Indeed, IMS is a well-established post-ionization separation method based on size, shape, and charge performed on a millisecond timescale, which represents an intermediate timescale between chromatography (seconds) and high-resolution MS detection (microseconds). Coupled with high-resolution mass spectrometry and chromatography (LC-IM-MS), IMS provides additional analyte selectivity without significantly compromising the speed of MS-based measurements. The MS dimension affords accurate mass information, while the IMS dimension provides molecular, structural, and conformational information through the determination of the ion collision cross section. Indeed, ion mobility spectrometry adds a separation dimension to the hybrid MS instruments allowing, thus, a more comprehensive analysis of complex biological mixtures [69,74–77]. Furthermore, accessing retention time, accurate mass, and collision cross section obtained by the combination of LC-IM-MS allows measurement integration, which enhances molecular identification and consequently biomarker discovery [78,79].

Fourier transform mass spectrometry is another group of ultra-high-resolution methods that offer the highest resolving power, resolution, and mass-to-charge ratio (m/z) measurement accuracy and, hence, better metabolome coverage [80]. However, given their high cost, these methods are limited to only a few research groups.

Recently, to increase the high-throughput of global metabolic profiling analysis, ambient ionization sources were introduced. They are capable of direct sampling for complex matrices under ambient conditions. For example, atmospheric solids analysis probe [81], desorption electrospray ionization (DESI) [82–84], and rapid evaporative ionization MS methods [85,86] have been demonstrated to provide real-time, interpretable MS data on biofluids and tissues, *in vivo* and *ex vivo*, and will certainly reshape the future for high-throughput real-time metabolome analysis. In many surgeries, it is often difficult to distinct visually between the healthy and diseased tissues, and this requires time-consuming biopsies and immuno-staining procedures to be performed by histopathologists during surgery. By eliminating this need for external tissue histotyping, the iKnife could open the way to true real-time precision surgery. For more details about the use of ambient MS in clinical diagnosis, refer to a recent and detailed review by Ifa et al. [87].

Table 1 presents a comparison between different analytical strategies used in metabolomics with potential interest for IEM. Given the already existing chemical biomarker infrastructure and growing adoption of MS in clinical laboratories, its relatively low cost compared to NMR instruments, and the analytical performance of current mass spectrometers in terms of sensitivity and resolution in particular, MS-based metabolomics is a very promising tool in clinical biochemistry in the near future [88].

2.3. Metabolomics Workflows : Targeted vs. Untargeted

Metabolomics analysis is typically described as two complementary analytical approaches : targeted and untargeted. The first one aims to define the metabolic profile of the groups to study; subsequently, multivariate statistical analysis is undertaken to define the discriminating metabolites (potential biomarkers) between groups. Second, predictive mathematical models based on multivariate statistical analysis can be built. These models predict a subsequent classification of unknown biological samples (e.g., healthy versus diseased, treated versus untreated). The targeted approach focuses on identifying and quantifying selected metabolites according to their involvement in a metabolic pathway or their specific chemical or biochemical proprieties.

Table 1. Comparison of main metabolomics analytical technologies with particular potential in inborn errors of metabolism.

Platform	Technique	Identification Dimensions	Principle	Advantages	Limits
Nuclear Magnetic Resonance based methods (NMR)	1Dimension	Chemical shift	Uses interaction of spin active nuclei (^1H , ^{13}C , ^{31}P) with electromagnetic fields which gives structural, chemical and molecular environment information	Nondestructive Highly reproducible Exact quantification possible Minimal sample preparation Molecular dynamic and compartmental information using diffusional methods Relative high throughput Availability of databases for identification	High instrumentation cost Overlap of metabolites Low sensitivity
	2Dimensions	Chemical shift x Chemical shift			
Mass spectrometry based methods	Direct Injection (DI-MS)	m/z	Uses a nanospray source directly coupled to MS detector. It does not require chromatographic separation	Very high throughput High sensitivity No cross sample contamination No column carryover Low cost analysis Automated analysis Low sample volume requirement	Samples not recoverable (destructive) No retention time information which gives limited specificity Inability to separate isomers and isobaric species Subjected to significant ion suppression phenomenon High ionization discrimination (ESI)
	Liquid chromatography (LC-MS)	Time x m/z	Uses chromatographic columns that enables liquid phase chromatographic separation of molecules followed by MS detection (Suitable for polar to hydrophobic compounds)	Minimal sample preparation (protein precipitation or dilution of biological sample) High throughput capability UPLC can be coupled to any type of MS Flexibility in column chemistry widening the range of detectable compounds High sensitivity	Samples not recoverable (destructive) Very polar molecules need specific chromatographic conditions Retention times are highly dependent of exact chromatographic conditions Batch analysis Lack of large metabolite databases High ionization discrimination (ESI)
	Gas Chromatography (GC-MS)	Time x m/z	Uses chromatographic columns that enables gas phase chromatographic separation of molecules followed by MS detection (Suited for apolar and volatiles compounds)	Structure information obtained through in-source fragmentation Availability of universal databases for identification High sensitive Reproducible	Samples not recoverable (destructive) Requires higher sample preparation Only volatile compounds are detected Polar compounds need derivatization Low ionization discrimination
	Capillary Electrophoresis (CE-MS)	Time x m/z	Uses electrokinetic separation of polar molecules hyphenated with a mass spectrometry detector	Excellent for polar analysis in aqueous samples Measures inorganic and organic anions Low running costs	Samples not recoverable (destructive) Relatively low throughput profiling

Ion Mobility (IM-MS)	Time x m/z (CCS x m/z)	Uses a uniform or periodic electric field and a buffer gas, to separate ions based on size and shape which is hyphenated with mass spectrometry	<p>Very robust and reproducible (ability to determine Collision Cross Section which is a robust chemical descriptor)</p> <p>High peak capacity</p> <p>High selectivity</p> <p>Separation of isomeric and isobaric compounds</p> <p>Very high throughput</p>	Samples not recoverable (destructive) CCS and mass are highly correlated parameters which limits the orthogonality of the method
----------------------	-----------------------------	---	---	--

In general, a metabolomic analysis involves mainly four steps. Step 1 is a preparatory step on both analytical and conceptual aspects. It is initiated by the biological question to consider and the definition of the study aim. It also defines the most informative biological matrix and the experimental design to implement. In addition, this step defines the appropriate sample preparation according to the considered analytical study. Step 2 includes analytical and instrumental strategy choices. During this step, data are collected and processed and then statistical analysis is performed. Step 3 involves the putative annotation, identification, and confirmation of the potential biomarkers generated by the data analysis. Chemical, biochemical, and spectral databases are queried. Step 4 aims to build a predictive mathematical model based on the identified biomarkers. This model is then validated analytically and clinically. This final step involves the integration of experimental data and their interpretation in the studied biological or clinical context [89]. Figure 1 illustrates the general workflow of translational metabolomics.

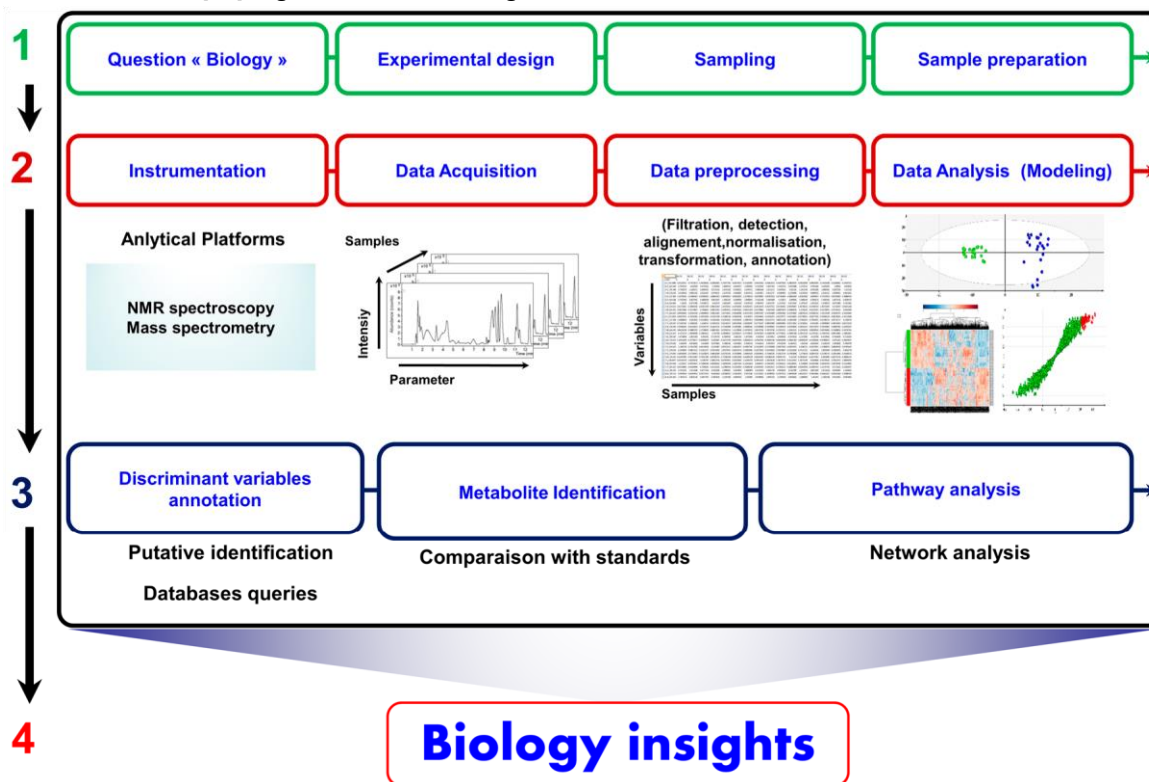


Figure 1. Translational metabolomics workflow.

2.4. Data Analysis, Information Recovery, and the Curse of Dimensionality

Few highly reliable metabolites could be, at some extent, sufficient for diagnostic or monitoring purposes. However, a broader overview using more metabolites is more appropriate to assess, for example, a biochemical pathway. Thus, the choice of the most appropriate data modeling strategy is an important issue and is dependent on the underlying question to be addressed. In mechanistic studies, the structural data descriptions and the underlying extracted information yielded by the built model are more important than its predictive ability to classify new samples. However, in diagnosis applications, the predictive performances of the model are vital regarding samples classification. Hence, the clear and precise definition of the study aims has to be intelligible and purpose driven.

The analytical performance improvements associated with metabolomics platforms led to the generation of complex and high-dimensional datasets. Handling, in a smoothly high-throughput fashion, the huge amount of generated data is a very important issue for transforming the data into clinically actionable knowledge.

2.4.1. Univariate Data Analysis

Metabolomics data analysis can be approached from a univariate perspective using traditional statistical methods that consider only one variable at a time. Univariate methods are common statistical analysis tools and their main advantage is the convenient use and interpretation. To assess the differences between two or more groups, parametric tests such as Student's *t*-test and ANOVA are commonly applied, respectively. However, normality assumptions should be verified for consistent conclusions [90]. Otherwise, non-parametric test such as Mann–Whitney *U* test or Kruskal–Wallis one-way analysis of variance could be used if normality is not assumed. Another important issue is that applying multiple

univariate tests in parallel to a high dimensional dataset raises the multiple testing problem. In metabolomics studies, a large number of features are simultaneously analyzed. Thus, the probability to find a statistically significant difference accidentally (i.e., true positive) is high. In order to handle this multiple testing issue, different correction methods could be used. Each method tries to balance between avoiding false metabolite associations (i.e., false positives) and discarding true associations (i.e., false negatives). In the Bonferroni correction, the significance level for a hypothesis is divided by the number of hypotheses tested simultaneously [90]. Hence, the Bonferroni correction is considered a stringent correction method. Other less conservative methods are available and are mostly based on the minimization of the false positives or false-discovery rate (FDR). FDR-based methods minimize the expected proportion of false positives on the total number of positives [91]. Gene expression microarray data analysis has matured most of these methods, where thousands of genes are simultaneously tested. Similarly, in untargeted metabolomics studies large sets of metabolites are measured in parallel. The use of less restrictive approaches such as FDR methods seems to be more useful.

Furthermore, it should be noted that potential confounding factors like gender, age or diet may affect the output results if not properly addressed. Furthermore, the main limit of these approaches is their lack of handling the correlations and interactions between the different metabolic features. Hence, advanced multivariate approaches are more suitable.

2.4.2. Multivariate Data Analysis

Translating biological data into knowledge requires addressing biology as an informational science using tools that allow to track the information at large scales. To do so, an entire field was born "Bioinformatics" [92]. Bioinformatics can be defined as mean of conceptualizing biology in terms of molecules and by applying "informatics techniques" borrowed to disciplines such as applied mathematics, computer science and statistics to understand and organize the information related to these molecules, on a large scale. In short, bioinformatics is a management information system for a biological system [93].

The high-dimensionality of metabolic data requires adapted statistical tools to retrieve as much as possible chemical information from the data to translate it into biological knowledge. The major challenge is to reduce the dimensionality by selecting relevant signals from the noisy raw data. To achieve this goal, chemometric tools are widely used. Chemometrics is the science of extracting useful information from chemical systems using data-driven means [94]. It is inherently interdisciplinary, borrowing methods from data-analytic disciplines such as multivariate statistics, applied mathematics, and computer science. Thus, chemometrics is applied to solve both descriptive and predictive problems using biochemical data.

The data analysis methods are mainly divided into two types : unsupervised and supervised methods. The former are mainly exploratory, whereas the latter are explanatory and predictive. Unsupervised methods are used to analyze the behavior of the observations in the data set without taking into account any related outcome. Because there is no class labeling or response, the data set is considered as a collection of analogous objects. Unsupervised learning methods track patterns or clustering trends in the data to understand any spontaneous relationships between the samples. It can also highlight the variables that are responsible for these relationships. Based on effective visualization means, unsupervised learning helps to reveal categories of samples or variables that naturally cluster together based on their underlying similarities. In metabolomics data, it is the metabolic similarity that shapes the clustering. Principal component analysis [95] is a widely used pattern recognition method; it is a projection-based method that reduces the dimensionality of the data by creating components or latent variables. Principal component analysis allows a two- or three-dimensional visualization of the data. However, clustering methods aim to identify clusters in the dataset using similarity measures. A dendrogram or a heat map can be then formed to visualize the samples similarities. The commonly used clustering methods are k-means clustering [96], hierarchical cluster analysis [97], and self-organizing maps [98]. Correlation matrix could also be used to get an overview of the data. Because the main goal in metabolomics, especially in clinical context, is to differentiate between groups (healthy versus diseased, treated versus control), a sample can be classified according to its spectral patterns. The metabolic features responsible for the classification can then be identified. The metabolic features intensities in the dataset matrix can be considered as a multidimensional space of metabolites coordinates. Thus, each spectrum is a point in a multidimensional metabolic hyperspace.

In supervised methods, the multivariate datasets can be modeled so that the class label of separate samples known as a validation set can be predicted based on a series of mathematical models derived from the original data, namely the training set. Various supervised methods could be used in metabolomics, including partial least squares (PLS) methods such as PLS-Discriminant Analysis (PLS-DA) [99] and Orthogonal-PLS-DA (OPLS-DA) [100], as well as support vector machines [101]. Methods based on topology data analysis are gaining great interests and seem promising for data

analysis because of their intrinsic flexibility and exploratory and predictive abilities [102]. It must be noted that the retrieved information from the raw data and the generated outputs are highly dependent on the chosen data analysis strategy. Hence, the aim of metabolomics research and the data analysis step are mutually dependent.

Of note, multivariate and univariate data analysis pipelines are not mutually exclusive and it is often recommended to use both to maximize the quality of the information extraction from metabolomics data.

For further details on data analysis techniques and tools in metabolomics, refer to recent reviews on this issue [103–105].

2.5. Pathway and Network Analysis : From Information to Knowledge

The integration of experimental data and computational tools is mandatory to understand complex biological systems. This gave birth to computational biology which could be divided into two distinct branches : knowledge discovery or data-mining, and simulation-based analysis. The former extracts the hidden patterns from huge amount of experimental data, generating hypotheses. However, the latter tests hypotheses with in silico experiments, providing predictions to be confirmed by in vitro and in vivo studies [9].

One of the biggest challenges of any metabolomics study is linking the identified metabolites to biology, which is a crucial step to move from biomarkers towards more mechanistic insights. To achieve this purpose, pathway and network analysis approaches aim to capitalize on the information generated by metabolomics studies to get insightful inference [106,107]. Both approaches exploit the interrelationships properties contained in the metabolomic data. Network modeling and pathway-mapping tools help to decipher metabolites interactions roles in a biological disturbance [107].

Metabolic pathways are sets of metabolites that are connected to the same biological process, and that are linked by one or multiple enzymatic reactions directly or indirectly. Biological databases are therefore seminal enablers providing rich information of different of metabolic pathways (Table 2). Indeed, pathway analysis (PA) uses prior biological knowledge to analyze metabolic patterns from an integrative point of view. Pathway-based methods are currently known as metabolite set enrichment analysis (MSEA), and are methodologically based on the gene set enrichment analysis (GSEA) approach, previously developed for pathway analysis of gene-expression data [108,109].

Table 2. Biological databases.

Databases	Websites	Ref.
KEGG (Kyoto Encyclopedia of Genes and Genomes)	http://www.genome.jp/kegg	[110]
HumanCyc (Encyclopedia of Human Metabolic Pathways)	http://humancyc.org	[111]
MetaCyc (Encyclopedia of Metabolic Pathways)	http://metacyc.org	[112]
Reactome (A Curated Knowledgebase of Pathways)	http://www.reactome.org	[113]
SMPDB (Small Molecule Pathway Database)	http://www.smpdb.ca	[114]
Virtual Metabolic Human Database	https://vmh.uni.lu	[106]
Wikipathways	http://www.wikipathways.org	[115]

There are mainly three distinct methods to perform MSEA [108,116] :

Overrepresentation analysis (ORA) : The basic hypothesis in this method is that relevant pathways can be detected if the proportion of differential expressed metabolites, within a given pathway, exceeds the proportion of metabolites that could be randomly expected. A hypergeometric test or a Fisher's Exact test is used to evaluate the statistical significance of whether the metabolite belongs to the pathway. The final result from an ORA method consists in a list of the most relevant pathways, ranked by *p*-value and/or a multiple-hypothesis-test-corrected *p*-value. The ORA main advantage over non-knowledge-driven (i.e., purely data-driven) analysis is that it gives metabolomic data a biological context. This allows formulating a hypothesis that could subsequently be test experimentally. Hence, ORA turns data analysis into a knowledge generation cycle, proper of the Systems Biology approach. However, PA exhibits some limits. Due to the selected cut-off method for statistical significance potentially important components could be omitted in the analysis. Furthermore, PA assume that pathways are independent from each other, which is contrary to the admitted interaction and overlapping between pathways [108]. Other methods have been developed to overcome these limits.

Quantitative enrichment analysis (QEA) : In this approach, the input data are a set of quantified metabolite from multiple samples. Thus, absolute concentrations are used. Enriched pathways can be identified using different approaches like the Wilcoxon-based test [117], globaltest [118] or globalAncova [119]. Enriched pathways include pathways where a set of metabolites that are significantly changed or pathways where a large number of metabolites that significantly changed [116,120].

Single-sample profiling (SSP) : Unlike the previous methods that are designed for studies involving multiple samples, this method is used at the sample level. In this case, SSP requires a list of metabolite concentrations in biofluids (i.e., urine, blood and CSF), tissue, or cell type and a database with the normal concentration ranges of the chosen metabolites in the analyzed sample. Thus, SSP identifies, from the data, the set of metabolites presenting significantly different levels compared to the normal ranges [116,120].

For better interpretability of pathway analysis outputs, MSEA results could be combined with pathway topological analysis (PTA). PTA measures assess the impact of the disturbed metabolites within the pathway. First, single impacts are evaluated using the degree and betweenness network centrality measures of each metabolite. This represents the number of shortest paths passing through a certain node to estimate its centrality (importance). Subsequently, the overall impact (i.e., pathway impact) is calculated as the sum of the single impact measures of the disturbed metabolites normalized by the sum of the measures of the impact of all the metabolites within the considered pathway [121]. Indeed, changes in the most important nodes within a network generate a more significant impact on the system than changes in bordering or solitary nodes.

From a topological standpoint, a metabolic network can be considered as an interconnected ensemble of nodes presented by metabolites, and edges representing reactions catalyzed by enzymes. Thus, unlike PA, network analysis uses the high degree of correlation in metabolomics data to build metabolic networks that characterize the complex relationships the measured metabolites.

Biological data exhibit a high level of correlation that exists between the different biological components (i.e., DNA, mRNAs, proteins and metabolites). Indeed, a given metabolite may be connected to different metabolic pathways and, thus, show correlation patterns. In other cases, the observed correlations may be due to other causes such as global changes (i.e., diurnal variation in time series studies) or specific changes due to the intrinsic variability of metabolomic data [54,122]. These patterns can provide valuable information about the underlying metabolic network associated to a specific biological process [54,123].

Based on the observed relationship patterns present in the experimental data, correlation-based methods allow building metabolic networks in which each metabolite represents a node. However, unlike the pathway analysis, the links between nodes denotes the level of mathematical correlation between each metabolites pair and called edge. High correlation coefficients are frequent in metabolomics data which is due to the presence of systemic associations

[123]. Hence, using classical correlation coefficients leads to overcrowded networks. In addition, direct and indirect associations are not distinguished. To overcome this problem partial correlation could be used [54,123,124]. In partial correlation approach, the correlation between two metabolites is conditioned against the correlation with the remaining metabolites. Consequently, partial correlation allows discriminating between direct and indirect metabolite correlations. In this method, the link between two metabolites is scored according to the ratios differences between the corresponding metabolites in the two sample groups. Therefore, the related network topology is based on the metabolic differences between the two studied phenotypes. These data-driven strategies have been successfully applied for reconstruction of metabolic networks from metabolomics data [123,125,126].

Metabolite identification is a challenging and time consuming task. Thus, a novel approach, named Mummichog, has been proposed by Li et al. for network analysis. This method predicts biological activity directly from mass spectrometry based untargeted metabolomics data without a priori identification of metabolites. The idea behind this strategy is combining network analysis and metabolite prediction under the same computational framework reducing significantly the metabolomics workflow time. This method has been elegantly illustrated by exploring the activation of innate immune cells. It yielded that glutathione metabolism is modified by viral infection driven by constitutive nitric oxide synthases [127].

A wide variety of software tools are available to analyze metabolomic data at the pathway and network level. Table 3 presents different functional analysis tools for both pathway analysis and visualization.

Table 3. Functional analysis and biological interpretation tools.

Tools	Websites	Ref.
Pathway and Networks Analysis and Visualization		
BioCyc—Omics Viewer	http://biocyc.org	[128]
iPath	http://pathways.embl.de	[129]
Metscape	http://metscape.ncibi.org	[130]
Paintomics	http://www.paintomics.org	[131]
Pathos	http://motif.gla.ac.uk/Pathos	[132]
Pathvisio	http://www.pathvisio.org	[133]
VANTED	http://vanted.ipk-gatersleben.de	[134]
IMPALA	http://impala.molgen.mpg.de	[135]
MBROLE 2.0	http://csbg.cnb.csic.es/mbrole2	[136]
MPEA	http://ekhidna.biocenter.helsinki.fi/poxo/mpea	[137]
Mummichog	http://clinicalmetabolomics.org/init/default/software	[127]
Multifunctional Tools		
MetaboAnalyst	http://www.metaboanalyst.com	[120]
XCMS online	https://xcmsonline.scripps.edu	[138]
MASSPup	http://www.bioprocess.org/massypup	[139]
Workflow4Metabolomics	http://workflow4metabolomics.org	[140]
MetaboLyzr	https://sites.google.com/a/georgetown.edu/fornace-labinformatics/home/metabolizer	[141]

Contextual interpretation is crucial to fully embrace the potential of metabolomics. Indeed, metabolites carry out precious contextual biological information. In a metabolic network, flux is defined as the rate (i.e., quantity per unit time) at which metabolites are converted or transported between different compartments [10]. Thus, metabolic fluxes, or fluxome, represent a unique and functional readout of the phenotype. The fluxome captures the metabolome in its ultimate functional interactions with the environment and the genome [10,142]. As such, the fluxome integrates information on different cellular processes, and hence it is a unique spatiotemporal phenotypic signature of cells. Thus, one or more metabolic fluxes could be altered in a metabolic disorder depending on the complexity of the disease [2]. Different strategies are used to translate metabolomics data into fluxomic insights by modeling of metabolic networks. The network modeling can be achieved using constraints of mass and charge conservation along with stoichiometric and thermodynamic ones [34,143–145]. Based on the stoichiometry of the reactants and products of biochemical reactions, flux balance analysis (FBA) can estimate metabolic fluxes without knowledge about the kinetics of the participating enzymes [10,142]. Recently, Cortassa et al. suggested a new approach, distinct from FBA or metabolic flux analysis, which takes into account kinetic mechanisms and regulatory interactions [146].

3. Potential Integration of Metabolomics in Laboratory Medicine Frameworks

Metabolites embody physiological end-points and regulatory processes directly connected to the fluxome. Hence, the metabolome is very time sensitive and is constantly changing. Therefore, changes in metabolite concentrations are

usually more suitable to describe the biochemical state of a biological system. Because metabolomics is the ultimate expression of the genes' influences and proteins' use of metabolites, it offers a rich and tremendous view on the phenotype. Indeed, metabolites carry out precious contextual biological information that could be used to assess pathophysiological states.

Metabolic profiling as a diagnostic tool opens an informative metabolic window into disease, which makes metabolomics an appealing ally in disease diagnosis.

What makes metabolomics a key driver in the post-genomic era is its tight relationship with the phenotype, whether the phenotype is driven by a monogenic or a multifactorial complex condition. Linking metabolic profile modulation with particular genetic variation [126] and/or environmental factors such as the microbiome [147], diet [148], toxics [56], or therapies [34] offers an exciting opportunity to rationalize diagnostics and translate a more comprehensive information into clinical actionable knowledge. The above-cited factors that influence the metabolome, and then phenotype(s), remind us that assessing metabolites, as chemical supporters of life, is the core for knowledge building that will shape clinical decisions.

Early biochemists such as Cori, Warburg, Meyerhof, and Krebs made seminal contributions to map most fundamental aspects of metabolic pathways and physiology. Therefore, urine chemical properties guided early physicians in founding the concept of IEM [4]. Sir Garrod's idea suggested that a biochemical fingerprint within biofluids was a product of human variation and, hence, could be a surrogate for distinct diseases. Garrod argued that the IEM that he was able to observe "were merely extreme examples of variations of chemical behavior which are probably everywhere present in minor degrees" [149]. In other words, he believed that there were phenotypes that could be associated with specific biochemicals. However, given the limited technical sensitivity back then, he was not able to affirm this idea. Recently, his hypothesis was elegantly confirmed with metabolomics approaches and metabolic modeling [126].

With the expected improvements in the metabolic profiling scope and data quality, metabolomics is destined to play a major and disruptive role in the near future as an efficient screening and diagnostic tool [150]. There are mainly two ways that metabolomics could be implemented in clinical context and laboratory medicine : chemometrics or a quantitative approach. For the former, direct statistical analysis is applied to spectral patterns and signal intensity data, and identification of metabolites may be performed in the last step if needed. This method captures metabolic snapshots and builds pattern-recognition-based models using machine learning techniques to sort samples (subjects) according to their metabolic patterns. This approach is eloquently embodied by the intelligent scalpel (iKnife) introduced by Takatz et al. [85], which instantaneously classifies, *in vivo* and *ex vivo*, cancerous and noncancerous tissues. This compelling technology aims to help surgeons during cancer surgery [85,86]. In contrast, the quantitative approach targets a set of metabolites and then analyzes the quantitative data directly. This approach affords an absolute quantitation of a set of chosen metabolites (e.g., amino acids, carnitines and acylcarnitines, or organic acids). A multivariate predictive model can be built based on the absolute concentration of these metabolites to predict clinical status or intervention outcomes. Compared to quantitative metabolomics, the key advantage of chemometric profiling is its capability of automated and unbiased assessment of metabolomics data. However, it requires a large number of spectra and sample uniformity, which are less of a concern in quantitative metabolomics. Nevertheless, the multivariate data analysis strategies underlying the two strategies are quite similar. Figure 2 illustrates the two clinical workflows.

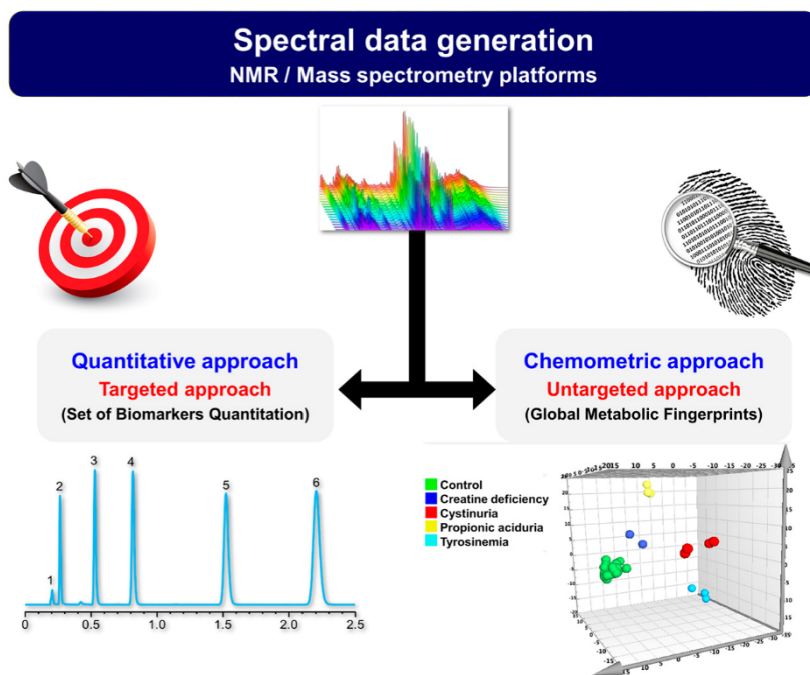


Figure 2. Clinical metabolomics implementation strategies.

4. Applications of Metabolomics in Inborn Errors of Metabolism (IEM) Investigations

IEM being tightly connected with metabolism, the inherent pathophysiological changes are the main determinant of the metabolome and functional understanding of the disease. Hence, due to its intrinsic multidisciplinary nature, integrating biochemistry, analytical chemistry, advanced statistics and bioinformatics, metabolomics analysis represents a promising tool to achieve improved understanding and better diagnosis of IEM in the post-genomic and precision medicine era. For years, MS has been used in the assessment of inherited metabolic diseases. Several IEM are currently diagnosed using targeted MS-based metabolomics methods such as aminoacidopathies, organic acidurias, and fatty acid oxidation disorders [151–155]. Furthermore, MS is now widely implemented in IEM newborn screening national programs worldwide [156]. However, the combination of the already existing tools with data analysis strategies is compelling for better biological information recovery.

To assess different IEM including aminoacidopathies, organic aciduria, and mitochondrial disorders, Janeckova et al. used targeted analysis combined with multivariate data analysis. Their work showed how combining chemometrics modeling and absolute quantification are eloquently complementary in the assessment of IEMs [157]. Drecksen et al. used a similar approach to assess isovaleric aciduria (IVA) based on 86 urine samples : 10 untreated and 10 treated IVA cases, 12 heterozygotes, 22 children controls, and 32 adult controls. The work succeeded in producing a comprehensive profile of metabolites of practical significance in IVA [158]. Osterman et al. described a matrix-assisted laser desorption/ionization MS-based method for acylcarnitine and organic acid analysis on DBS. The method enabled the identification and quantification of metabolites involved in different organic aciduria and beta oxidation deficiencies [159]. Using targeted metabolomics addressing complex lipids, Fan and colleagues showed that some sphingolipids species were elevated in Niemann–Pick Type C1 subjects. These lipid biomarkers may be used for monitoring the efficacy of specific therapy [160].

Given the increasing potential of metabolomics in IEM, different groups published work regarding the usefulness of untargeted-metabolomics-based approaches in IEM in disease characterization, diagnosis, and biomarker discovery. For characterization of disease biosignatures, respiratory chain deficiencies have been investigated by several research groups to track specific metabolic signatures using metabolomics [161–163]. Wikoff et al. used MS-based untargeted metabolomics in plasma to characterize methylmalonic acidemia and propionic aciduria. Propionylcarnitine, a known biomarker,

was retrieved using untargeted strategy which illustrates the potential of metabolic profiling in biomarker detection. Five additional plasma acylcarnitine metabolites presented significant differences between patients and control individuals. In addition, γ -butyrobetaine was highly increased in a subset of patients. This demonstrates that metabolomics can widen the range of metabolites associated with IEM [164]. Auray-Blais et al. used MS-based untargeted metabolomics for biomarker discovery in Fabry disease which led to the discovery of seven

globotriaosylceramide (Gb3) analogues as biomarkers that are now suggested biomarkers for the screening and the follow-up of Fabry disease [61,62,165].

Sholmi et al. presented an elegant computational approach for assessing metabolic profiles of red blood cells enzyme deficiencies. The developed predictive method yielded biomarkers for red blood cells alterations and revealed a strong correlation with disrupted metabolic concentrations. Over 200 metabolites were identified as potential biomarkers due to 176 enzyme deficiencies. Furthermore, already known disease indicators were retrieved by the developed prediction method. Importantly, potential novel biomarkers were also predicted. This approach proved to dramatically increase biomarker discovery performance [166].

Because the metabolome is highly influenced by nutritional factors, diet monitoring also has been investigated using metabolomics. Phenylketonuria is an interesting example of diet monitoring in IEM. Using metabolomics, Mutze et al. showed that a long-term dietary fatty acid restriction influences mitochondrial beta-oxidation intermediates. No functional influence on unsaturated fatty acid metabolism and platelet aggregation in patients with phenylketonuria was detected [167].

Regarding the use of metabolomics platforms as a diagnosis tool, several teams proposed metabolomics workflows. Using NMR and DESI-MS methods, Pan et al. clearly discriminated six patients with IEMs from six controls based on their respective urine metabolic profiles, identifying argininosuccinic aciduria, classic homocystinuria, classic methylmalonic acidemia, maple syrup urine disease, phenylketonuria, and type II tyrosinemia [168]. Later, Denes et al. proposed a method based on high-resolution MS with high throughput using DBS and direct flow injection analysis. Their method has been tested on 500 controls and 66 abnormal samples and showed a clear discrimination of the various assessed metabolic diseases [51]. Ilya et al. also proposed another method based on high-resolution MS coupled to liquid chromatography for the assessment of IEM. Their method resolved highly polar as well as hydrophobic analytes under reverse-phase conditions, enabling analysis of a wide range of chemicals in an untargeted fashion. Their work provides a tailored high-resolution MS platform for IEM and covers various metabolites usually quantified by a combination of different separate instrumentation [169]. Miller et al. described a comprehensive global strategy to assess IEM using liquid chromatography and gas chromatography MS-based metabolomics platforms combining both targeted and untargeted analysis. In total, 120 plasma samples from patients with a confirmed IEM and those of 70 controls were assessed. This strategy allowed, elegantly, comprehensive pathway analysis that provides useful diagnostic information of IEM [170].

Regarding NMR-based platforms, Aygen and colleagues conducted a multi-center clinical study in 14 clinical centers in Turkey. Urine samples from 989 neonates were collected and investigated using NMR spectroscopy in two different laboratories to assess reproducibility. The objectives of their study were twofold : (1) to explore the metabolite variations to set pathological thresholds of specific metabolites in comparison with healthy neonates to develop predictive models; and (2) to build a NMR database from a healthy population of neonates for IEM metabolite identification [171].

5. Clinical and Translational Metabolomics Challenges

5.1. Metabolite Identification

Metabolite identification is the main bottleneck of metabolomics for large adoption in both translational and clinical context. Despite spectral information becomes available in the literature or in spectral databases, metabolites identification is still a challenging task [172]. To the best of our knowledge, there is no software currently available to fully and smoothly facilitate the identification process. Especially, the integration of NMR and MS data, which is essential for the reliable identification of metabolites. Furthermore, metabolite identification is mandatory for absolute quantitation especially in MS based methods requiring the use of labeled isotope. Thus, more efforts are needed to enhance this drawback of metabolomics.

5.2. Standardization and Harmonization

Standardization is a vital aspect for a wide spread of any new technology. Thus, for clinical metabolomics, harmonization of the sample preparation, processing, analysis and reporting using validated and standardized protocols is mandatory [173,174]. This is important since biological samples change over time. The lack of harmonization in protocols for sample handling, MS and NMR data generation and data reporting could lead to poor reproducibility and, thus, to data misinterpretation, particularly in population metabolic profiling. This is a fundamental obstacle for clinical translation. A definition of normal or reference samples is also important to build reference databases. This will rely on signatures derived from the complete characterization of the considered disease, IEM in the scope of this review. Finally, addressing these standardization issues is essential for regulatory compliance,

which is a prerequisite for clinical implementation and adoption.

5.3. Automation, Data Visualization and Clinical Actionability

Automation at different stages, instrument-, pre- and post-analytic levels are a very important issue for large clinical adoption of any diagnostic innovation. Metabolomics workflow automation is a key enabler regarding high-throughput, reproducibility and reliability which are pillars of modern laboratory medicine practice. To address this limit, current efforts are promising like the iKnife, which would allow real-time cancer diagnosis [85] and breathomics strategies for lung and respiratory diseases based on breath signatures [72]. Data fusion and integration of omics and other biological and clinical data is another great challenge to fully unveil the potential of metabolomics [17,175]. With this regard, combining genomic and metabolic profiling information to enhance clinical diagnostics and to enable patient stratification and monitoring of interventional pathways patient journeys is a promising field [22,176]. The clinical actionability would involve advanced mathematical modeling of genomic and metabolic data sets in relation to patient clinical data using machine learning and expert systems. Intuitive visualization tools of the data in clinical accessible formats are needed to support effective clinical decision making. Figure 3 presents the main challenges in clinical metabolomics.



Figure 3. Metabolomics challenges for effective clinical implementation.

6. Conclusions

It is common to perform early diagnosis of IEM by assessing specific metabolic biomarkers related to a genetic defect. However, the original paradigm of “one gene–one enzyme–one disease” is no longer viewed as a reality for IEM. The impact of an altered protein on metabolic flux is not easily predictable. Indeed, the metabolic pathways are not linear and metabolites are tightly linked with several interactions within a highly organized network [21,177]. Depending on the complexity of the disease, one primary metabolite flux or an entire network of metabolite fluxes might be affected [2,20]. Therefore, a complete contextual, multilayer, network-based functional overview is needed to effectively assess all the actors of a given pathway in a holistic fashion [8]. Systemic approaches are needed to understand IEM complexity and to effectively diagnose and treat them [21]. To achieve such a goal, metabolomics is a key driver in the systems medicine based strategy. The great potential of metabolomics integration with other omics data will allow systems biology and clinical data to be linked. This paves the way for a paradigm shift in medical practice from cohort evidence-based medicine to algorithm-based precision medicine. This will in turn enhance clinicians’ abilities to be more pre-emptive and thus, more efficient in handling IEM.

Metabolomics is still in its infancy with regard to the investigation of IEM, and its great potential has yet to be explored worldwide at both the basic and clinical sides. Improving workflows for high-quality data acquisition, processing, and visualization is an important issue for effectively translating the biological information into actionable knowledge under clinically accessible formats for effective healthcare management. However, this innovative global approach also requires a paradigm shift in our practice at different levels. A complete change is needed in our screening and diagnosis strategies. Thus, a disruptive move from a hypothesis-driven approach to a more data-driven and hypothesis-generating approach is crucial to address the challenges of the post-genomic era. The core idea of the paradigm shift in IEM laboratory investigation is presented in Figure 4.

Furthermore, totally new investigative thinking is needed to transform all aspects of the laboratory medicine enterprise, including education, research, and healthcare. Upgrading medical practitioners’ skill sets on both the clinical and laboratory sides is needed to smoothly achieve the full potential of systems medicine. These skills integrate biology, computing and data analytics to develop common communication channels for more effective medical interactions. This ongoing high digitization of the individual biological and clinical information offers a tremendous and exciting opportunity to fully embrace the promising era of precision medicine.

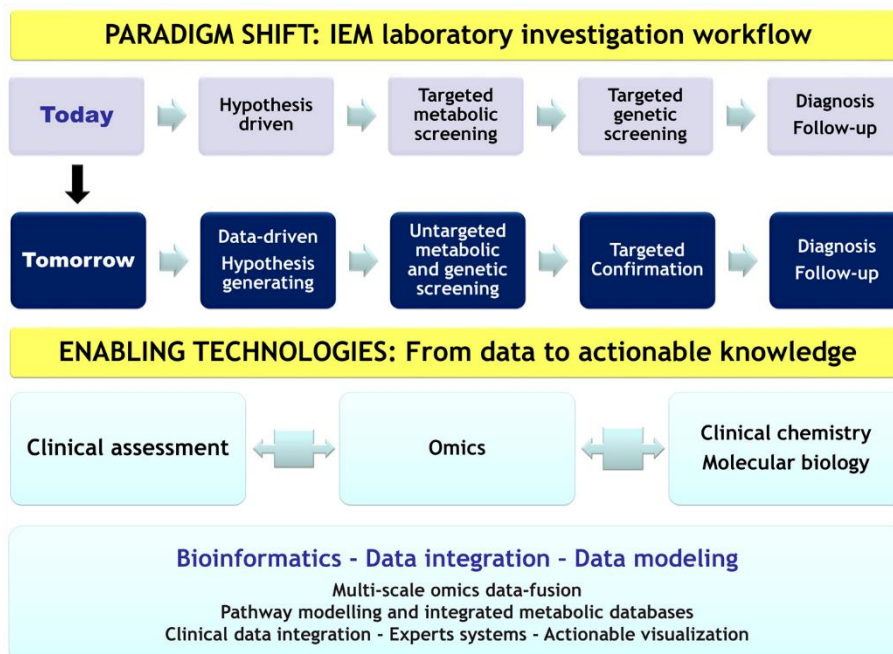


Figure 4. Paradigm shift in Inborn Errors of Metabolism diagnosis workflow. The change in molecular information recovery in laboratory investigation workflow is driven by advancing analytical technologies and bioinformatics systems for a more effective medical practice using an integrative computational framework. IEM : Inborn Errors of Metabolism.

Acknowledgments : This work was supported by the Normandy University, the Institut National de la Santé et de la Recherche Médicale (INSERM), the Conseil Régional de Normandie, Labex SynOrg (ANR-11-LABX-0029) and the European Regional Development Fund (ERDF 31708).

Author Contributions : Abdellah Tebani performed literature review and wrote the manuscript including tables and figures. Lenaig Abily-Donval performed literature search. Carlos Afonso critically revised and edited the manuscript. Stéphane Marret critically revised and edited the manuscript. Soumeiya Bekri conceived the topic of the review and critically revised and edited the manuscript. All authors approved the final manuscript.

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

References

- Collins, F.S.; Varmus, H. A new initiative on precision medicine. *N. Engl. J. Med.* **2015**, *372*, 793–795. [[CrossRef](#)] [[PubMed](#)]
- Lanpher, B.; Brunetti-Pierri, N.; Lee, B. Inborn errors of metabolism : The flux from mendelian to complex diseases. *Nat. Rev. Genet.* **2006**, *7*, 449–460. [[CrossRef](#)] [[PubMed](#)]
- Vernon, H.J. Inborn errors of metabolism : Advances in diagnosis and therapy. *JAMA Pediatr.* **2015**, *169*, 778–782. [[CrossRef](#)] [[PubMed](#)]
- Garrod, A. The incidence of alkaptonuria : A study in chemical individuality. *Lancet* **1902**, *160*, 1616–1620. [[CrossRef](#)]
- Ahn, A.C.; Tewari, M.; Poon, C.S.; Phillips, R.S. The limits of reductionism in medicine : Could systems biology offer an alternative? *PLoS Med.* **2006**, *3*, e208. [[CrossRef](#)] [[PubMed](#)]
- Regenmortel, M.H.V.V. Reductionism and complexity in molecular biology. *EMBO Rep.* **2004**, *5*, 1016–1020. [[CrossRef](#)] [[PubMed](#)]
- Aon, M.A.; Lloyd, D.; Saks, V. From physiology, genomes, systems, and self-organization to systems biology : The historical roots of a twenty-first century approach to complexity. In *Systems Biology of Metabolic and Signaling Networks : Energy, Mass and Information Transfer*; Aon, A.M., Saks, V., Schlattner, U., Eds.; Springer : Berlin & Heidelberg, Germany, 2014; pp. 3–17.
- Aon, M.A. Complex systems biology of networks : The riddle and the challenge. In *Systems Biology of Metabolic and Signaling Networks*; Springer : Berlin & Heidelberg, Germany, 2014; pp. 19–35.
- Kitano, H. Computational systems biology. *Nature* **2002**, *420*, 206–210. [[CrossRef](#)] [[PubMed](#)]
- Aon, M.A.; Cortassa, S. Systems biology of the fluxome. *Processes* **2015**, *3*, 607–618. [[CrossRef](#)]
- Weston, A.D.; Hood, L. Systems biology, proteomics, and the future of health care : Toward predictive, preventative, and personalized medicine. *J. Proteome Res.* **2004**, *3*, 179–196. [[CrossRef](#)] [[PubMed](#)]
- Ehrenberg, M.; Elf, J.; Aurell, E.; Sandberg, R.; Tegner, J. Systems biology is taking off. *Genome Res.* **2003**, *13*, 2377–2380. [[CrossRef](#)] [[PubMed](#)]

13. Kitano, H. Looking beyond the details : A rise in system-oriented approaches in genetics and molecular biology. *Curr. Genet.* **2002**, *41*, 1–10. [[CrossRef](#)] [[PubMed](#)]
14. Kitano, H. Systems biology : A brief overview. *Science* **2002**, *295*, 1662–1664. [[CrossRef](#)] [[PubMed](#)]
15. Tenenbaum, J.D.; Avillach, P.; Benham-Hutchins, M.; Breitenstein, M.K.; Crowgey, E.L.; Hoffman, M.A.; Jiang, X.; Madhavan, S.; Mattison, J.E.; Nagarajan, R.; et al. An informatics research agenda to support precision medicine : Seven key areas. *JAMIA* **2016**, *23*, 791–795. [[CrossRef](#)] [[PubMed](#)]
16. McMurry, J.; Kohler, S.; Balhoff, J.; Borromeo, C.; Brush, M.; Carbon, S.; Conlin, T.; Dunn, N.; Engelstad, M.; Foster, E.; et al. Navigating the phenotype frontier : The monarch initiative. *bioRxiv* **2016**. [[CrossRef](#)]
17. Ritchie, M.D.; Holzinger, E.R.; Li, R.; Pendergrass, S.A.; Kim, D. Methods of integrating data to uncover genotype-phenotype interactions. *Nat. Rev. Genet.* **2015**, *16*, 85–97. [[CrossRef](#)] [[PubMed](#)]
18. Sperisen, P.; Cominetti, O.; Martin, F.-P.J. Longitudinal omics modeling and integration in clinical metabonomics research : Challenges in childhood metabolic health research. *Front. Mol. Biosci.* **2015**. [[CrossRef](#)] [[PubMed](#)]
19. Hood, L.; Balling, R.; Auffray, C. Revolutionizing medicine in the 21st century through systems approaches. *Biotechnol. J.* **2012**, *7*, 992–1001. [[CrossRef](#)] [[PubMed](#)]
20. Cho, D.-Y.; Kim, Y.-A.; Przytycka, T.M. Chapter 5 : Network biology approach to complex diseases. *PLoS Comput. Biol.* **2012**, *8*, e1002820. [[CrossRef](#)] [[PubMed](#)]
21. Argmann, C.A.; Houten, S.M.; Zhu, J.; Schadt, E.E. A next generation multiscale view of inborn errors of metabolism. *Cell Metab.* **2016**, *23*, 13–26. [[CrossRef](#)] [[PubMed](#)]
22. Nicholson, J.K.; Holmes, E.; Kinross, J.M.; Darzi, A.W.; Takats, Z.; Lindon, J.C. Metabolic phenotyping in clinical and surgical environments. *Nature* **2012**, *491*, 384–392. [[CrossRef](#)] [[PubMed](#)]
23. Chen, R.; Mias, G.I.; Li-Pook-Tham, J.; Jiang, L.; Lam, H.Y.K.; Chen, R.; Miriami, E.; Karczewski, K.J.; Hariharan, M.; Dewey, F.E.; et al. Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell* **2012**, *148*, 1293–1307. [[CrossRef](#)] [[PubMed](#)]
24. Nicholson, J.K.; Lindon, J.C. Systems biology : Metabonomics. *Nature* **2008**, *455*, 1054–1056. [[CrossRef](#)] [[PubMed](#)]
25. Oliver, S.G.; Winson, M.K.; Kell, D.B.; Baganz, F. Systematic functional analysis of the yeast genome. *Trends Biotechnol.* **1998**, *16*, 373–378. [[CrossRef](#)]
26. Nicholson, J.K.; Lindon, J.C.; Holmes, E. “Metabonomics” : Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **1999**, *29*, 1181–1189. [[CrossRef](#)] [[PubMed](#)]
27. Dunn, W.B.; Broadhurst, D.I.; Atherton, H.J.; Goodacre, R.; Griffin, J.L. Systems level studies of mammalian metabolomes : The roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chem. Soc. Rev.* **2011**, *40*, 387–426. [[CrossRef](#)] [[PubMed](#)]
28. Fiehn, O. Metabolomics—The link between genotypes and phenotypes. *Plant Mol. Biol.* **2002**, *48*, 155–171. [[CrossRef](#)] [[PubMed](#)]
29. Holmes, E.; Wilson, I.D.; Nicholson, J.K. Metabolic phenotyping in health and disease. *Cell* **2008**, *134*, 714–717. [[CrossRef](#)] [[PubMed](#)]
30. Clayton, T.A.; Lindon, J.C.; Cloarec, O.; Antti, H.; Charuel, C.; Hanton, G.; Provost, J.P.; le Net, J.L.; Baker, D.; Walley, R.J.; et al. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature* **2006**, *440*, 1073–1077. [[CrossRef](#)] [[PubMed](#)]
31. Kaddurah-Daouk, R.; Kristal, B.S.; Weinshilboum, R.M. Metabolomics : A global biochemical approach to drug response and disease. *Annu. Rev. Pharmacol. Toxicol.* **2008**, *48*, 653–683. [[CrossRef](#)] [[PubMed](#)]
32. James, L.P. Metabolomics : Integration of a new “omics” with clinical pharmacology. *Clin. Pharmacol. Ther.* **2013**, *94*, 547–551. [[CrossRef](#)] [[PubMed](#)]
33. Kaddurah-Daouk, R.; Weinshilboum, R.M. Pharmacometabolomics : Implications for clinical pharmacology and systems pharmacology. *Clin. Pharmacol. Ther.* **2014**, *95*, 154–167. [[CrossRef](#)] [[PubMed](#)]
34. Kell, D.B.; Goodacre, R. Metabolomics and systems pharmacology : Why and how to model the human metabolic network for drug discovery. *Drug Discov. Today* **2014**, *19*, 171–182. [[CrossRef](#)] [[PubMed](#)]
35. Everett, J.R. Pharmacometabonomics in humans : A new tool for personalized medicine. *Pharmacogenomics* **2015**, *16*, 737–754. [[CrossRef](#)] [[PubMed](#)]
36. Dunn, W.B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J.D.; Halsall, A.; Haselden, J.N.; et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **2011**, *6*, 1060–1083. [[CrossRef](#)] [[PubMed](#)]
37. Nunes de Paiva, M.J.; Menezes, H.C.; de Lourdes Cardeal, Z. Sampling and analysis of metabolomes in biological fluids. *Analyst* **2014**, *139*, 3683–3694. [[CrossRef](#)] [[PubMed](#)]
38. Want, E.J.; Masson, P.; Michopoulos, F.; Wilson, I.D.; Theodoridis, G.; Plumb, R.S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J.K. Global metabolic profiling of animal and human tissues via UPLC-MS. *Nat. Protoc.* **2013**, *8*, 17–32. [[CrossRef](#)] [[PubMed](#)]
39. Want, E.J.; Wilson, I.D.; Gika, H.; Theodoridis, G.; Plumb, R.S.; Shockcor, J.; Holmes, E.; Nicholson, J.K. Global metabolic profiling procedures for urine using UPLC-MS. *Nat. Protoc.* **2010**, *5*, 1005–1018. [[CrossRef](#)] [[PubMed](#)]

40. Graham, S.F.; Chevallier, O.P.; Roberts, D.; Hölscher, C.; Elliott, C.T.; Green, B.D. Investigation of the human brain metabolome to identify potential markers for early diagnosis and therapeutic targets of Alzheimer's disease. *Anal. Chem.* **2013**, *85*, 1803–1811. [[CrossRef](#)] [[PubMed](#)]
41. Wuolikainen, A.; Hedenstrom, M.; Moritz, T.; Marklund, S.L.; Antti, H.; Andersen, P.M. Optimization of procedures for collecting and storing of CSF for studying the metabolome in ALS. *Amyotroph. Lateral Scler.* **2009**, *10*, 229–236. [[CrossRef](#)] [[PubMed](#)]
42. Dame, Z.; Aziat, F.; Mandal, R.; Krishnamurthy, R.; Bouatra, S.; Borzouie, S.; Guo, A.; Sajed, T.; Deng, L.; Lin, H.; et al. The human saliva metabolome. *Metabolomics* **2015**, *11*, 1864–1883. [[CrossRef](#)]
43. Kawasaki, G.; Ichikawa, Y.; Yoshitomi, I.; Umeda, M. Metabolomics of salivary biomarkers in yusho patients. *Hukuoka Acta Med.* **2015**, *106*, 144–148. [[PubMed](#)]
44. Mikkonen, J.J.; Singh, S.P.; Herrala, M.; Lappalainen, R.; Myllymaa, S.; Kullaa, A.M. Salivary metabolomics in the diagnosis of oral cancer and periodontal diseases. *J. Periodontal Res.* **2015**. [[CrossRef](#)] [[PubMed](#)]
45. Bach, J.-P.; Gold, M.; Mengel, D.; Hattesoehl, A.; Lubbe, D.; Schmid, S.; Tackenberg, B.; Rieke, J.; Maddula, S.; Baumbach, J.I.; et al. Measuring compounds in exhaled air to detect Alzheimer's disease and parkinson? S disease. *PLoS ONE* **2015**, *10*, e0132227. [[CrossRef](#)] [[PubMed](#)]
46. Pijls, K.E.; Smolinska, A.; Jonkers, D.M.A.E.; Dallinga, J.W.; Masclee, A.A.M.; Koek, G.H.; van Schooten, F.-J. A profile of volatile organic compounds in exhaled air as a potential non-invasive biomarker for liver cirrhosis. *Sci. Rep.* **2016**. [[CrossRef](#)] [[PubMed](#)]
47. Koulman, A.; Prentice, P.; Wong, M.C.; Matthews, L.; Bond, N.J.; Eiden, M.; Griffin, J.L.; Dunger, D.B. The development and validation of a fast and robust dried blood spot based lipid profiling method to study infant metabolism. *Metabolomics* **2014**, *10*, 1018–1025. [[CrossRef](#)] [[PubMed](#)]
48. Wilson, I. Global metabolic profiling (metabonomics/metabolomics) using dried blood spots : Advantages and pitfalls. *Bioanalysis* **2011**, *3*, 2255–2257. [[CrossRef](#)] [[PubMed](#)]
49. Michopoulos, F.; Theodoridis, G.; Smith, C.J.; Wilson, I.D. Metabolite profiles from dried blood spots for metabonomic studies using UPLC combined with orthogonal acceleration TOF-MS : Effects of different papers and sample storage stability. *Bioanalysis* **2011**, *3*, 2757–2767. [[CrossRef](#)] [[PubMed](#)]
50. Prentice, P.; Turner, C.; Wong, M.C.Y.; Dalton, R.N. Stability of metabolites in dried blood spots stored at different temperatures over a 2-year period. *Bioanalysis* **2013**, *5*, 1507–1514. [[CrossRef](#)] [[PubMed](#)]
51. Denes, J.; Szabo, E.; Robinette, S.L.; Szatmari, I.; Szonyi, L.; Kreuder, J.G.; Rauterberg, E.W.; Takats, Z. Metabonomics of newborn screening dried blood spot samples : A novel approach in the screening and diagnostics of inborn errors of metabolism. *Anal. Chem.* **2012**, *84*, 10113–10120. [[CrossRef](#)] [[PubMed](#)]
52. Wagner, M.; Tonoli, D.; Varesio, E.; Hopfgartner, G. The use of mass spectrometry to analyze dried blood spots. *Mass Spectrom. Rev.* **2014**, *35*, 361–438. [[CrossRef](#)] [[PubMed](#)]
53. Oliveira, R.V.; Henion, J.; Wickremsinhe, E.R. Automated high-capacity on-line extraction and bioanalysis of dried blood spot samples using liquid chromatography/high-resolution accurate mass spectrometry. *Rapid Commun. Mass Spectrom.* **2014**, *28*, 2415–2426. [[CrossRef](#)] [[PubMed](#)]
54. Do, K.T.; Kastenmüller, G.; Mook-Kanamori, D.O.; Yousri, N.A.; Theis, F.J.; Suhre, K.; Krumsiek, J. Network-based approach for analyzing intra- and interfluid metabolite associations in human blood, urine, and saliva. *J. Proteome Res.* **2015**, *14*, 1183–1194. [[CrossRef](#)] [[PubMed](#)]
55. Torell, F.; Bennett, K.; Cereghini, S.; Rannar, S.; Lundstedt-Enkel, K.; Moritz, T.; Haumaitre, C.; Trygg, J.; Lundstedt, T. Multi-organ contribution to the metabolic plasma profile using hierarchical modelling. *PLoS ONE* **2015**, *10*, e0129260. [[CrossRef](#)] [[PubMed](#)]
56. Athersuch, T. Metabolome analyses in exposome studies : Profiling methods for a vast chemical space. *Arch. Biochem. Biophys.* **2016**, *589*, 177–186. [[CrossRef](#)] [[PubMed](#)]
57. Pauling, L.; Robinson, A.B.; Teranishi, R.; Cary, P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 2374–2376. [[CrossRef](#)] [[PubMed](#)]
58. Jimenez, B.; Montoliu, C.; MacIntyre, D.A.; Serra, M.A.; Wassel, A.; Jover, M.; Romero-Gomez, M.; Rodrigo, J.M.; Pineda-Lucena, A.; Felipe, V. Serum metabolic signature of minimal hepatic encephalopathy by (1) h-nuclear magnetic resonance. *J. Proteome Res.* **2010**, *9*, 5180–5187. [[CrossRef](#)] [[PubMed](#)]
59. Wijeyesekera, A.; Selman, C.; Barton, R.H.; Holmes, E.; Nicholson, J.K.; Withers, D.J. Metabotyping of long-lived mice using 1 h NMR spectroscopy. *J. Proteome Res.* **2012**, *11*, 2224–2235. [[CrossRef](#)] [[PubMed](#)]
60. Larive, C.K.; Barding, G.A., Jr.; Dinges, M.M. NMR spectroscopy for metabolomics and metabolic profiling. *Anal. Chem.* **2015**, *87*, 133–146. [[CrossRef](#)] [[PubMed](#)]
61. Auray-Blais, C.; Boutin, M. Novel GB(3) isoforms detected in urine of fabry disease patients : A metabolomic study. *Curr. Med. Chem.* **2012**, *19*, 3241–3252. [[CrossRef](#)] [[PubMed](#)]
62. Manwaring, V.; Boutin, M.; Auray-Blais, C. A metabolomic study to identify new globotriaosylceramide-related biomarkers in the plasma of fabry disease patients. *Anal. Chem.* **2013**, *85*, 9039–9048. [[CrossRef](#)] [[PubMed](#)]
63. Emwas, A.-H.; Salek, R.; Griffin, J.; Merzaban, J. Nmr-based metabolomics in human disease diagnosis : Applications, limitations, and recommendations. *Metabolomics* **2013**, *9*, 1048–1072. [[CrossRef](#)]

64. Chan, E.C.Y.; Pasikanti, K.K.; Nicholson, J.K. Global urinary metabolic profiling procedures using gas chromatography-mass spectrometry. *Nat. Protoc.* **2011**, *6*, 1483–1499. [[CrossRef](#)] [[PubMed](#)]
65. Ramautar, R.; Somsen, G.W.; de Jong, G.J. CE-MS for metabolomics : Developments and applications in the period 2012–2014. *Electrophoresis* **2015**, *36*, 212–224. [[CrossRef](#)] [[PubMed](#)]
66. Hill, H.H., Jr.; Siems, W.F.; St Louis, R.H.; McMinn, D.G. Ion mobility spectrometry. *Anal. Chem.* **1990**, *62*, 1201A–1209A. [[CrossRef](#)] [[PubMed](#)]
67. Paglia, G.; Angel, P.; Williams, J.P.; Richardson, K.; Olivos, H.J.; Thompson, J.W.; Menikarachchi, L.; Lai, S.; Walsh, C.; Moseley, A.; et al. Ion mobility-derived collision cross section as an additional measure for lipid fingerprinting and identification. *Anal. Chem.* **2015**, *87*, 1137–1144. [[CrossRef](#)] [[PubMed](#)]
68. Maldini, M.; Natella, F.; Baima, S.; Morelli, G.; Scaccini, C.; Langridge, J.; Astarita, G. Untargeted metabolomics reveals predominant alterations in lipid metabolism following light exposure in broccoli sprouts. *Int. J. Mol. Sci.* **2015**, *16*, 13678–13691. [[CrossRef](#)] [[PubMed](#)]
69. Paglia, G.; Williams, J.P.; Menikarachchi, L.; Thompson, J.W.; Tyldesley-Worster, R.; Halldórsson, S.; Rolfsson, O.; Moseley, A.; Grant, D.; Langridge, J.; et al. Ion mobility derived collision cross sections to support metabolomics applications. *Anal. Chem.* **2014**, *86*, 3985–3993. [[CrossRef](#)] [[PubMed](#)]
70. Wickramasekara, S.I.; Zandkarimi, F.; Morre, J.; Kirkwood, J.; Legette, L.; Jiang, Y.; Gombart, A.F.; Stevens, J.F.; Maier, C.S. Electrospray quadrupole travelling wave ion mobility time-of-flight mass spectrometry for the detection of plasma metabolome changes caused by xanthohumol in obese Zucker (fa/fa) rats. *Metabolites* **2013**, *3*, 701–717. [[CrossRef](#)] [[PubMed](#)]
71. Dwivedi, P.; Schultz, A.J.; Hill, H.H. Metabolic profiling of human blood by high resolution ion mobility mass spectrometry (IM-MS). *Int. J. Mass Spectrom.* **2010**, *298*, 78–90. [[CrossRef](#)] [[PubMed](#)]
72. Hauschild, A.C.; Frisch, T.; Baumbach, J.I.; Baumbach, J. Carotia : Revealing hidden confounder markers in metabolic breath profiles. *Metabolites* **2015**, *5*, 344–363. [[CrossRef](#)] [[PubMed](#)]
73. Smolinska, A.; Hauschild, A.C.; Fijten, R.R.; Dallinga, J.W.; Baumbach, J.; van Schooten, F.J. Current breathomics—A review on data pre-processing techniques and machine learning in metabolomics breath analysis. *J. Breath Res.* **2014**. [[CrossRef](#)] [[PubMed](#)]
74. Fenn, L.; Kliman, M.; Mahsut, A.; Zhao, S.; McLean, J. Characterizing ion mobility-mass spectrometry conformation space for the analysis of complex biological samples. *Anal. Bioanal. Chem.* **2009**, *394*, 235–244. [[CrossRef](#)] [[PubMed](#)]
75. Fenn, L.; McLean, J. Biomolecular structural separations by ion mobility–mass spectrometry. *Anal. Bioanal. Chem.* **2008**, *391*, 905–909. [[CrossRef](#)] [[PubMed](#)]
76. Kliman, M.; May, J.C.; McLean, J.A. Lipid analysis and lipidomics by structurally selective ion mobility-mass spectrometry. *Biochim. Biophys. Acta* **2011**, *1811*, 935–945. [[CrossRef](#)] [[PubMed](#)]
77. Tebani, A.; Schmitz-Afonso, I.; Rutledge, D.N.; Gonzalez, B.J.; Bekri, S.; Afonso, C. Optimization of a liquid chromatography ion mobility-mass spectrometry method for untargeted metabolomics using experimental design and multivariate data analysis. *Anal. Chim. Acta* **2016**, *913*, 55–62. [[CrossRef](#)] [[PubMed](#)]
78. May, J.C.; Goodwin, C.R.; McLean, J.A. Ion mobility-mass spectrometry strategies for untargeted systems, synthetic, and chemical biology. *Curr. Opin. Biotechnol.* **2015**, *31*, 117–121. [[CrossRef](#)] [[PubMed](#)]
79. Sherrod, S.D.; McLean, J.A. Systems-wide high-dimensional data acquisition and informatics using structural mass spectrometry strategies. *Clin. Chem.* **2015**, *62*, 77–83. [[CrossRef](#)] [[PubMed](#)]
80. Junot, C.; Madalinski, G.; Tabet, J.C.; Ezan, E. Fourier transform mass spectrometry for metabolome analysis. *Analyst* **2010**, *135*, 2203–2219. [[CrossRef](#)] [[PubMed](#)]
81. Twohig, M.; Shockcor, J.P.; Wilson, I.D.; Nicholson, J.K.; Plumb, R.S. Use of an atmospheric solids analysis probe (ASAP) for high throughput screening of biological fluids : Preliminary applications on urine and bile. *J. Proteome Res.* **2010**, *9*, 3590–3597. [[CrossRef](#)] [[PubMed](#)]
82. Eberlin, L.S.; Norton, I.; Orringer, D.; Dunn, I.F.; Liu, X.; Ide, J.L.; Jarmusch, A.K.; Ligon, K.L.; Jolesz, F.A.; Golby, A.J.; et al. Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1611–1616. [[CrossRef](#)] [[PubMed](#)]
83. Ferreira, C.R.; Jarmusch, A.K.; Pirro, V.; Alfaro, C.M.; Gonzalez-Serrano, A.F.; Niemann, H.; Wheeler, M.B.; Rabel, R.A.; Hallett, J.E.; Houser, R.; et al. Ambient ionisation mass spectrometry for lipid profiling and structural analysis of mammalian oocytes, preimplantation embryos and stem cells. *Reprod. Fertil. Dev.* **2015**, *27*, 621–637. [[CrossRef](#)] [[PubMed](#)]
84. Kerian, K.S.; Jarmusch, A.K.; Pirro, V.; Koch, M.O.; Masterson, T.A.; Cheng, L.; Cooks, R.G. Differentiation of prostate cancer from normal tissue in radical prostatectomy specimens by desorption electrospray ionization and touch spray ionization mass spectrometry. *Analyst* **2015**, *140*, 1090–1098. [[CrossRef](#)] [[PubMed](#)]
85. Balog, J.; Sasi-Szabo, L.; Kinross, J.; Lewis, M.R.; Muirhead, L.J.; Veselkov, K.; Mirnezami, R.; Dezso, B.; Damjanovich, L.; Darzi, A.; et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Sci. Transl. Med.* **2013**. [[CrossRef](#)] [[PubMed](#)]
86. Balog, J.; Kumar, S.; Alexander, J.; Golf, O.; Huang, J.; Wiggins, T.; Abbassi-Ghadi, N.; Enyedi, A.; Kacska, S.; Kinross, J.; et al. In vivo endoscopic tissue identification by rapid evaporative ionization mass spectrometry (REIMS). *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 11059–11062. [[CrossRef](#)] [[PubMed](#)]

87. Ifa, D.R.; Eberlin, L.S. Ambient ionization mass spectrometry for cancer diagnosis and surgical margin evaluation. *Clin. Chem.* **2016**, *62*, 111–123. [[CrossRef](#)] [[PubMed](#)]
88. Annesley, T.; Diamandis, E.; Bachmann, L.; Hanash, S.; Hart, B.; Javahery, R.; Singh, R.; Smith, R. A spectrum of views on clinical mass spectrometry. *Clin. Chem.* **2016**, *62*, 30–36. [[CrossRef](#)] [[PubMed](#)]
89. Wishart, D.S.; Jewison, T.; Guo, A.C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; et al. Hmdb 3.0—The human metabolome database in 2013. *Nucleic Acids Res.* **2013**, *41*, D801–D807. [[CrossRef](#)] [[PubMed](#)]
90. Broadhurst, D.I.; Kell, D.B. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* **2006**, *2*, 171–196. [[CrossRef](#)]
91. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate : A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **1995**, *57*, 289–300.
92. Hogeweg, P. The roots of bioinformatics in theoretical biology. *PLoS Comput. Biol.* **2011**, *7*, e1002021. [[CrossRef](#)] [[PubMed](#)]
93. Luscombe, N.M.; Greenbaum, D.; Gerstein, M. What is bioinformatics? A proposed definition and overview of the field. *Methods Inf. Med.* **2001**, *40*, 346–358. [[PubMed](#)]
94. Brereton, R.G. A short history of chemometrics : A personal view. *J. Chemom.* **2014**, *28*, 749–760. [[CrossRef](#)]
95. Hotelling, H. *Analysis of a Complex of Statistical Variables into Principal Components*; Warwick & York : Oradell, NJ, USA, 1933.
96. Hartigan, J.A.; Wong, M.A. Algorithm as 136 : A k-means clustering algorithm. *J. R. Stat. Soc.* **1979**, *28*, 100–108. [[CrossRef](#)]
97. Johnson, S.C. Hierarchical clustering schemes. *Psychometrika* **1967**, *32*, 241–254. [[CrossRef](#)] [[PubMed](#)]
98. Kohonen, T. The self-organizing map. *Proc. IEEE* **1990**, *78*, 1464–1480. [[CrossRef](#)]
99. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression : A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
100. Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* **2002**, *16*, 119–128. [[CrossRef](#)]
101. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
102. Offroy, M.; Duponchel, L. Topological data analysis : A promising big data exploration tool in biology, analytical chemistry and physical chemistry. *Anal. Chim. Acta* **2016**, *910*, 1–11. [[CrossRef](#)] [[PubMed](#)]
103. Ren, S.; Hinzman, A.; Kang, E.; Szczesniak, R.; Lu, L. Computational and statistical analysis of metabolomics data. *Metabolomics* **2015**, *11*, 1492–1513. [[CrossRef](#)]
104. Gromski, P.S.; Muhamadali, H.; Ellis, D.I.; Xu, Y.; Correa, E.; Turner, M.L.; Goodacre, R. A tutorial review : Metabolomics and partial least squares-discriminant analysis—A marriage of convenience or a shotgun wedding. *Anal. Chim. Acta* **2015**, *879*, 10–23. [[CrossRef](#)] [[PubMed](#)]
105. Misra, B.B.; van der Hooft, J.J. Updates in metabolomics tools and resources : 2014–2015. *Electrophoresis* **2016**, *37*, 86–110. [[CrossRef](#)] [[PubMed](#)]
106. Thiele, I.; Swainston, N.; Fleming, R.M.; Hoppe, A.; Sahoo, S.; Aurich, M.K.; Haraldsdottir, H.; Mo, M.L.; Rolfsson, O.; Stobbe, M.D.; et al. A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* **2013**, *31*, 419–425. [[CrossRef](#)] [[PubMed](#)]
107. Cazzaniga, P.; Damiani, C.; Besozzi, D.; Colombo, R.; Nobile, M.S.; Gaglio, D.; Pescini, D.; Molinari, S.; Mauri, G.; Alberghina, L.; et al. Computational strategies for a system-level understanding of metabolism. *Metabolites* **2014**, *4*, 1034–1087. [[CrossRef](#)] [[PubMed](#)]
108. Garcia-Campos, M.A.; Espinal-Enriquez, J.; Hernandez-Lemus, E. Pathway analysis : State of the art. *Front. Physiol.* **2015**. [[CrossRef](#)] [[PubMed](#)]
109. Khatri, P.; Sirota, M.; Butte, A.J. Ten years of pathway analysis : Current approaches and outstanding challenges. *PLoS Comput. Biol.* **2012**, *8*, e1002375. [[CrossRef](#)] [[PubMed](#)]
110. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. Kegg as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2016**, *44*, D457–D462. [[CrossRef](#)] [[PubMed](#)]
111. Romero, P.; Wagg, J.; Green, M.L.; Kaiser, D.; Krummenacker, M.; Karp, P.D. Computational prediction of human metabolic pathways from the complete human genome. *Genome Biol.* **2005**. [[CrossRef](#)] [[PubMed](#)]
112. Caspi, R.; Foerster, H.; Fulcher, C.A.; Kaipa, P.; Krummenacker, M.; Latendresse, M.; Paley, S.; Rhee, S.Y.; Shearer, A.G.; Tissier, C.; et al. The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway/genome databases. *Nucleic Acids Res.* **2008**, *36*, D623–D631. [[CrossRef](#)] [[PubMed](#)]
113. Vastrik, I.; D'Eustachio, P.; Schmidt, E.; Gopinath, G.; Croft, D.; de Bono, B.; Gillespie, M.; Jassal, B.; Lewis, S.; Matthews, L.; et al. Reactome : A knowledge base of biologic pathways and processes. *Genome Biol.* **2007**. [[CrossRef](#)] [[PubMed](#)]
114. Jewison, T.; Su, Y.; Disfany, F.M.; Liang, Y.; Knox, C.; Maciejewski, A.; Poelzer, J.; Huynh, J.; Zhou, Y.; Arndt, D.; et al. Smpdb 2.0 : Big improvements to the small molecule pathway database. *Nucleic Acids Res.* **2014**, *42*, D478–D484. [[CrossRef](#)] [[PubMed](#)]
115. Kelder, T.; van Iersel, M.P.; Hanspers, K.; Kutmon, M.; Conklin, B.R.; Evelo, C.T.; Pico, A.R. Wikipathways : Building research communities on biological pathways. *Nucleic Acids Res.* **2012**, *40*, D1301–D1307. [[CrossRef](#)] [[PubMed](#)]
116. Xia, J.; Wishart, D.S. MSEA : A web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res.* **2010**, *38*, W71–W77. [[CrossRef](#)] [[PubMed](#)]

117. Adjaye, J.; Huntriss, J.; Herwig, R.; BenKahla, A.; Brink, T.C.; Wierling, C.; Hultschig, C.; Groth, D.; Yaspo, M.L.; Picton, H.M.; et al. Primary differentiation in the human blastocyst : Comparative molecular portraits of inner cell mass and trophectoderm cells. *Stem Cells* **2005**, *23*, 1514–1525. [[CrossRef](#)] [[PubMed](#)]
118. Goeman, J.J.; van de Geer, S.A.; de Kort, F.; van Houwelingen, H.C. A global test for groups of genes : Testing association with a clinical outcome. *Bioinformatics* **2004**, *20*, 93–99. [[CrossRef](#)] [[PubMed](#)]
119. Hummel, M.; Meister, R.; Mansmann, U. Globalancova : Exploration and assessment of gene group effects. *Bioinformatics* **2008**, *24*, 78–85. [[CrossRef](#)] [[PubMed](#)]
120. Xia, J.; Sinelnikov, I.V.; Han, B.; Wishart, D.S. Metaboanalyst 3.0—Making metabolomics more meaningful. *Nucleic Acids Res.* **2015**, *43*, W251–W257. [[CrossRef](#)] [[PubMed](#)]
121. Xia, J.; Wishart, D.S. METPA : A web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* **2010**, *26*, 2342–2344. [[CrossRef](#)] [[PubMed](#)]
122. Steuer, R. Review : On the analysis and interpretation of correlations in metabolomic data. *Brief. Bioinform.* **2006**, *7*, 151–158. [[CrossRef](#)] [[PubMed](#)]
123. Krumsiek, J.; Suhre, K.; Illig, T.; Adamski, J.; Theis, F.J. Gaussian graphical modeling reconstructs pathway reactions from high-throughput metabolomics data. *BMC Syst. Biol.* **2011**. [[CrossRef](#)] [[PubMed](#)]
124. Valcarcel, B.; Wurtz, P.; Seich al Basatena, N.K.; Tukiainen, T.; Kangas, A.J.; Soinen, P.; Jarvelin, M.R.; Ala-Korpela, M.; Ebbels, T.M.; de Iorio, M. A differential network approach to exploring differences between biological states : An application to prediabetes. *PLoS ONE* **2011**, *6*, e24702. [[CrossRef](#)] [[PubMed](#)]
125. Bartel, J.; Krumsiek, J.; Schramm, K.; Adamski, J.; Gieger, C.; Herder, C.; Carstensen, M.; Peters, A.; Rathmann, W.; Roden, M.; et al. The human blood metabolome-transcriptome interface. *PLoS Genet.* **2015**, *11*, e1005274. [[CrossRef](#)] [[PubMed](#)]
126. Shin, S.Y.; Fauman, E.B.; Petersen, A.K.; Krumsiek, J.; Santos, R.; Huang, J.; Arnold, M.; Erte, I.; Forgetta, V.; Yang, T.P.; et al. An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **2014**, *46*, 543–550. [[CrossRef](#)] [[PubMed](#)]
127. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F.H.; Khan, N.; Soltow, Q.A.; Jones, D.P.; Pulendran, B. Predicting network activity from high throughput metabolomics. *PLoS Comput. Biol.* **2013**, *9*, e1003123. [[CrossRef](#)] [[PubMed](#)]
128. Caspi, R.; Billington, R.; Ferrer, L.; Foerster, H.; Fulcher, C.A.; Keseler, I.M.; Kothari, A.; Krummenacker, M.; Latendresse, M.; Mueller, L.A.; et al. The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway/genome databases. *Nucleic Acids Res.* **2016**, *44*, D471–D480. [[CrossRef](#)] [[PubMed](#)]
129. Yamada, T.; Letunic, I.; Okuda, S.; Kanehisa, M.; Bork, P. Ipath2.0 : Interactive pathway explorer. *Nucleic Acids Res.* **2011**, *39*, W412–W415. [[CrossRef](#)] [[PubMed](#)]
130. Karnovsky, A.; Weymouth, T.; Hull, T.; Tarcea, V.G.; Scardoni, G.; Laudanna, C.; Sartor, M.A.; Stringer, K.A.; Jagadish, H.V.; Burant, C.; et al. Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics* **2012**, *28*, 373–380. [[CrossRef](#)] [[PubMed](#)]
131. Garcia-Alcalde, F.; Garcia-Lopez, F.; Dopazo, J.; Conesa, A. Paintomics : A web based tool for the joint visualization of transcriptomics and metabolomics data. *Bioinformatics* **2011**, *27*, 137–139. [[CrossRef](#)] [[PubMed](#)]
132. Leader, D.P.; Burgess, K.; Creek, D.; Barrett, M.P. Pathos : A web facility that uses metabolic maps to display experimental changes in metabolites identified by mass spectrometry. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 3422–3426. [[CrossRef](#)] [[PubMed](#)]
133. Kutmon, M.; van Iersel, M.P.; Bohler, A.; Kelder, T.; Nunes, N.; Pico, A.R.; Evelo, C.T. Pathvisio 3 : An extendable pathway analysis toolbox. *PLoS Comput. Biol.* **2015**, *11*, e1004085. [[CrossRef](#)] [[PubMed](#)]
134. Rohn, H.; Junker, A.; Hartmann, A.; Grafarend-Belau, E.; Treutler, H.; Klapperstück, M.; Czauderna, T.; Klukas, C.; Schreiber, F. Vanted v2 : A framework for systems biology applications. *BMC Syst. Biol.* **2012**, *6*, 1–13. [[CrossRef](#)] [[PubMed](#)]
135. Kamburov, A.; Cavill, R.; Ebbels, T.M.D.; Herwig, R.; Keun, H.C. Integrated pathway-level analysis of transcriptomics and metabolomics data with impala. *Bioinformatics* **2011**, *27*, 2917–2918. [[CrossRef](#)] [[PubMed](#)]
136. Lopez-Ibanez, J.; Pazos, F.; Chagoyen, M. Mbrole 2.0-functional enrichment of chemical compounds. *Nucleic Acids Res.* **2016**, *44*, W201–W204. [[CrossRef](#)] [[PubMed](#)]
137. Kankainen, M.; Gopalacharyulu, P.; Holm, L.; Oresic, M. Mpea—Metabolite pathway enrichment analysis. *Bioinformatics* **2011**, *27*, 1878–1879. [[CrossRef](#)] [[PubMed](#)]
138. Tautenhahn, R.; Patti, G.J.; Rinehart, D.; Siuzdak, G. Xcms online : A web-based platform to process untargeted metabolomic data. *Anal. Chem.* **2012**, *84*, 5035–5039. [[CrossRef](#)] [[PubMed](#)]
139. Winkler, R. An evolving computational platform for biological mass spectrometry : Workflows, statistics and data mining with massypup64. *PeerJ* **2015**, *3*, e1401. [[CrossRef](#)] [[PubMed](#)]
140. 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. Workflow4metabolomics : A collaborative research infrastructure for computational metabolomics. *Bioinformatics* **2015**, *31*, 1493–1495. [[CrossRef](#)] [[PubMed](#)]
141. Mak, T.D.; Laiakis, E.C.; Goudarzi, M.; Fornace, A.J., Jr. Metabolizer : A novel statistical workflow for analyzing postprocessed LC-MS metabolomics data. *Anal. Chem.* **2014**, *86*, 506–513. [[CrossRef](#)] [[PubMed](#)]
142. Cascante, M.; Marin, S. Metabolomics and fluxomics approaches. *Essays Biochem.* **2008**, *45*, 67–82. [[CrossRef](#)] [[PubMed](#)]
143. Cortassa, S.; Aon, M.A. Computational modeling of mitochondrial function. *Methods Mol. Biol.* **2012**, *810*, 311–326. [[PubMed](#)]

144. Winter, G.; Kromer, J.O. Fluxomics—Connecting “omics” analysis and phenotypes. *Environ. Microbiol.* **2013**, *15*, 1901–1916. [[CrossRef](#)] [[PubMed](#)]
145. Aurich, M.K.; Thiele, I. Computational modeling of human metabolism and its application to systems biomedicine. *Methods Mol. Biol.* **2016**, *1386*, 253–281. [[PubMed](#)]
146. Cortassa, S.; Caceres, V.; Bell, L.N.; O'Rourke, B.; Paolocci, N.; Aon, M.A. From metabolomics to fluxomics : A computational procedure to translate metabolite profiles into metabolic fluxes. *Biophys. J.* **2015**, *108*, 163–172. [[CrossRef](#)] [[PubMed](#)]
147. Cho, I.; Blaser, M.J. The human microbiome : At the interface of health and disease. *Nat. Rev. Genet.* **2012**, *13*, 260–270. [[CrossRef](#)] [[PubMed](#)]
148. Holmes, E.; Loo, R.L.; Stamler, J.; Bictash, M.; Yap, I.K.; Chan, Q.; Ebbels, T.; de Iorio, M.; Brown, I.J.; Veselkov, K.A.; et al. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* **2008**, *453*, 396–400. [[CrossRef](#)] [[PubMed](#)]
149. Garrod, A.E. *The Inborn Factors in Disease*; Clarendon Press : Oxford, UK, 1931.
150. Beebe, K.; Kennedy, A.D. Sharpening precision medicine by a thorough interrogation of metabolic individuality. *Comput. Struct. Biotechnol. J.* **2016**, *14*, 97–105. [[CrossRef](#)] [[PubMed](#)]
151. Auray-Blais, C.; Maranda, B.; Lavoie, P. High-throughput tandem mass spectrometry multiplex analysis for newborn urinary screening of creatine synthesis and transport disorders, triple H syndrome and otc deficiency. *Clin. Chim. Acta* **2014**, *436*, 249–255. [[CrossRef](#)] [[PubMed](#)]
152. Pitt, J.J. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin. Biochem. Rev.* **2009**, *30*, 19–34. [[PubMed](#)]
153. Pitt, J.J. Newborn screening. *Clin. Biochem. Rev.* **2010**, *31*, 57–68. [[PubMed](#)]
154. Pitt, J.J.; Eggington, M.; Kahler, S.G. Comprehensive screening of urine samples for inborn errors of metabolism by electrospray tandem mass spectrometry. *Clin. Chem.* **2002**, *48*, 1970–1980. [[PubMed](#)]
155. Spacil, Z.; Tatipaka, H.; Barcenas, M.; Scott, C.R.; Turecek, F.; Gelb, M.H. High-throughput assay of 9 lysosomal enzymes for newborn screening. *Clin. Chem.* **2013**, *59*, 502–511. [[CrossRef](#)] [[PubMed](#)]
156. Therrell, B.L.; Padilla, C.D.; Loeber, J.G.; Kneisser, I.; Saadallah, A.; Borrajo, G.J.; Adams, J. Current status of newborn screening worldwide : 2015. *Semin. Perinatol.* **2015**, *39*, 171–187. [[CrossRef](#)] [[PubMed](#)]
157. Janeckova, H.; Hron, K.; Wojtowicz, P.; Hlidkova, E.; Baresova, A.; Friedecky, D.; Zidkova, L.; Hornik, P.; Behulova, D.; Prochazkova, D.; et al. Targeted metabolomic analysis of plasma samples for the diagnosis of inherited metabolic disorders. *J. Chromatogr. A* **2012**, *1226*, 11–17. [[CrossRef](#)] [[PubMed](#)]
158. Dercksen, M.; Koekemoer, G.; Duran, M.; Wanders, R.J.A.; Mienie, L.J.; Reinecke, C.J. Organic acid profile of isovaleric acidemia : A comprehensive metabolomics approach. *Metabolomics* **2013**, *9*, 765–777. [[CrossRef](#)]
159. Ostermann, K.M.; Dieplinger, R.; Lutsch, N.M.; Strupat, K.; Metz, T.F.; Mechtler, T.P.; Kasper, D.C. Matrix-assisted laser desorption/ionization for simultaneous quantitation of (acyl)-carnitines and organic acids in dried blood spots. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 1497–1504. [[CrossRef](#)] [[PubMed](#)]
160. Fan, M.; Sidhu, R.; Fujiwara, H.; Tortelli, B.; Zhang, J.; Davidson, C.; Walkley, S.U.; Bagel, J.H.; Vite, C.; Yanjanin, N.M.; et al. Identification of niemann-pick c1 disease biomarkers through sphingolipid profiling. *J. Lipid Res.* **2013**, *54*, 2800–2814. [[CrossRef](#)] [[PubMed](#)]
161. Reinecke, C.J.; Koekemoer, G.; Westhuizen, F.H.; Louw, R.; Lindeque, J.Z.; Mienie, L.J.; Smuts, I. Metabolomics of urinary organic acids in respiratory chain deficiencies in children. *Metabolomics* **2011**, *8*, 264–283. [[CrossRef](#)]
162. Smuts, I.; Westhuizen, F.H.; Louw, R.; Mienie, L.J.; Engelke, U.F.H.; Wevers, R.A.; Mason, S.; Koekemoer, G.; Reinecke, C.J. Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. *Metabolomics* **2012**, *9*, 379–391. [[CrossRef](#)]
163. Venter, L.; Lindeque, Z.; Jansen van Rensburg, P.; van der Westhuizen, F.; Smuts, I.; Louw, R. Untargeted urine metabolomics reveals a biosignature for muscle respiratory chain deficiencies. *Metabolomics* **2014**, *11*, 111–121. [[CrossRef](#)]
164. Wikoff, W.R.; Gangoiti, J.A.; Barshop, B.A.; Siuzdak, G. Metabolomics identifies perturbations in human disorders of propionate metabolism. *Clin. Chem.* **2007**, *53*, 2169–2176. [[CrossRef](#)] [[PubMed](#)]
165. Auray-Blais, C.; Boutin, M.; Gagnon, R.; Dupont, F.O.; Lavoie, P.; Clarke, J.T. Urinary globotriaosylsphingosine-related biomarkers for fabry disease targeted by metabolomics. *Anal. Chem.* **2012**, *84*, 2745–2753. [[CrossRef](#)] [[PubMed](#)]
166. Shlomi, T.; Cabili, M.N.; Ruppin, E. Predicting metabolic biomarkers of human inborn errors of metabolism. *Mol. Syst. Biol.* **2009**. [[CrossRef](#)] [[PubMed](#)]
167. Mutze, U.; Beblo, S.; Kortz, L.; Matthies, C.; Koletzko, B.; Bruegel, M.; Rohde, C.; Thiery, J.; Kiess, W.; Ceglarek, U. Metabolomics of dietary fatty acid restriction in patients with phenylketonuria. *PLoS ONE* **2012**, *7*, e43021. [[CrossRef](#)] [[PubMed](#)]
168. Pan, Z.; Gu, H.; Talaty, N.; Chen, H.; Shanaiah, N.; Hainline, B.E.; Cooks, R.G.; Raftery, D. Principal component analysis of urine metabolites detected by nmr and desi-ms in patients with inborn errors of metabolism. *Anal. Bioanal. Chem.* **2007**, *387*, 539–549. [[CrossRef](#)] [[PubMed](#)]
169. Gertsman, I.; Gangoiti, J.A.; Barshop, B.A. Validation of a dual LC-HRMS platform for clinical metabolic diagnosis in serum, bridging quantitative analysis and untargeted metabolomics. *Metabolomics* **2014**, *10*, 312–323. [[CrossRef](#)] [[PubMed](#)]

170. Miller, M.; Kennedy, A.; Eckhart, A.; Burrage, L.; Wulff, J.; Miller, L.D.; Milburn, M.; Ryals, J.; Beaudet, A.; Sun, Q.; et al. Untargeted metabolomic analysis for the clinical screening of inborn errors of metabolism. *J. Inherit. Metab. Dis.* **2015**, *1–11*. [[CrossRef](#)] [[PubMed](#)]
171. Aygen, S.; Dürr, U.; Hegele, P.; Kunig, J.; Spraul, M.; Schäfer, H.; Krings, D.; Cannet, C.; Fang, F.; Schütz, B.; et al. NMR-based screening for inborn errors of metabolism : Initial results from a study on turkish neonates. *JIMD Rep.* **2014**, *16*, 101–111. [[PubMed](#)]
172. Goodacre, R.; Broadhurst, D.; Smilde, A.K.; Kristal, B.S.; Baker, J.D.; Beger, R.; Bessant, C.; Connor, S.; Capuani, G.; Craig, A. Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics* **2007**, *3*, 231–241. [[CrossRef](#)]
173. Chitayat, S.; Rudan, J.F. Chapter 10—Phenome centers and global harmonization. In *Metabolic Phenotyping in Personalized and Public Healthcare*; Academic Press : Boston, MA, USA, 2016; pp. 291–315.
174. Kohler, I.; Verhoeven, A.; Derks, R.J.; Giera, M. Analytical pitfalls and challenges in clinical metabolomics. *Bioanalysis* **2016**, *8*, 1509–1532. [[CrossRef](#)] [[PubMed](#)]
175. Alyass, A.; Turcotte, M.; Meyre, D. From big data analysis to personalized medicine for all : Challenges and opportunities. *BMC Med. Genom.* **2015**, *8*, 1–12. [[CrossRef](#)] [[PubMed](#)]
176. Tarailo-Graovac, M.; Shyr, C.; Ross, C.J.; Horvath, G.A.; Salvarinova, R.; Ye, X.C.; Zhang, L.H.; Bhavsar, A.P.; Lee, J.J.; Drogemoller, B.I.; et al. Exome sequencing and the management of neurometabolic disorders. *N. Engl. J. Med.* **2016**, *374*, 2246–2255. [[CrossRef](#)] [[PubMed](#)]
177. Sahoo, S.; Franzson, L.; Jonsson, J.J.; Thiele, I. A compendium of inborn errors of metabolism mapped onto the human metabolic network. *Mol. Biosyst.* **2012**, *8*, 2545–2558. [[CrossRef](#)] [[PubMed](#)]



© 2016 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 (<http://creativecommons.org/licenses/by/4.0/>).

Urinary metabolic phenotyping of mucopolysaccharidosis type I combining untargeted and targeted strategies with data modeling

Abdellah TEBANI^{1,2,3}, Isabelle SCHMITZ-AFONSO³, Lenaig ABILY-DONVAL^{2,4}, Bénédicte HERON⁵, Monique PIRAUD⁶, Jérôme AUSSEIL⁷, Anais BRASSIER⁸, Pascale De LONLAY⁸, Farid ZERMICHE⁹, Frédéric M VAZ¹⁰, Bruno J. GONZALEZ², Stephane MARRET^{2,4}, Carlos AFONSO³, Soumeya BEKRI^{1,2,*}

¹ Department of Metabolic Biochemistry, Rouen University Hospital, Rouen, 76000, France

² Normandie Univ, UNIROUEN, CHU Rouen, INSERM U1245, 76000 Rouen, France

³ Normandie Univ, UNIROUEN, INSA Rouen, CNRS, COBRA, 76000 Rouen, France

⁴ Department of Neonatal Pediatrics and Intensive Care, Rouen University Hospital, Rouen, 76031, France

⁵ Centre de Référence des Maladies Lysosomales (CRML), Neuropédiatrie, CHU Trousseau, APHP, Paris, France.

⁶ Service Maladies Héritaires du Métabolisme et Dépistage Néonatal, Centre de Biologie et de Pathologie Est CHU de Lyon, Lyon, France.

⁷ INSERM U1088, Laboratoire de Biochimie Spécialisée, Centre de Biologie Humaine, CHU Sud, 80054, Amiens Cedex, France.

⁸ Reference Center of Inherited Metabolic Diseases, Imagine Institute, Hospital Necker Enfants Malades, APHP, University Paris Descartes, Paris, France.

⁹ Laboratoire de Biochimie et Biologie Moléculaire, Université de Lille et Pôle de Biologie Pathologie Génétique du CHRU de Lille, 59000, Lille, France.

¹⁰ Department of Clinical Chemistry and Pediatrics, Laboratory Genetic Metabolic Diseases, Academic Medical Center, Amsterdam, The Netherlands

Abstract

BACKGROUND: Application of metabolic phenotyping could expand the pathophysiological knowledge of mucopolysaccharidoses (MPS). However, few MPS studies applied broad untargeted metabolite profiling to reveal the comprehensive metabolic impairments in MPS.

METHODS: We applied targeted and untargeted metabolic profiling in urine samples obtained from a French cohort comprising 19 MPSI patients and 15 treated MPSI patients along with 66 controls. For that purpose, we used ultra-high performance liquid chromatography combined with ion mobility and high resolution mass spectrometry using a protocol designed for large-scale metabolomics studies regarding robustness and reproducibility. Furthermore, 24 amino acids have been quantified using a liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). Keratan sulfate, Heparane sulfate and Dermatan sulfate concentrations have been measured using an LC-MS/MS method. Univariate and multivariate data analysis have been used to select discriminant metabolites. The mummichog algorithm has been used for pathway analysis.

RESULTS: The studied groups yielded distinct biochemical phenotypes using multivariate data analysis. Univariate statistics also revealed metabolites that differentiated the groups. Specifically, metabolites related to the amino acids metabolism. Pathway analysis revealed that several major amino acid pathways were dysregulated in MPS with arginine, proline and glutathione metabolisms being the most affected.

CONCLUSION: This study constitutes one of the first metabolic phenotyping studies of MPSI. The findings might help to generate new hypotheses about MPS pathophysiology and to develop further targeted studies of a smaller number of potentially key metabolites.

Keywords: Metabolomics; mucopolysaccharidosis type I; lysosomal storage diseases; mass spectrometry; ion mobility; inborn errors of metabolism

1. Introduction

Inborn errors of metabolism (IEM) represent a group of about 500 rare diseases with an overall estimated incidence of 1/2500. The diversity of involved metabolisms explains the difficulties in establishing their diagnosis. Optimal management of these patients requires then improved speed of biochemical investigation to allow early diagnosis and better monitoring. The rise of "omic" approaches offered a growing hope to provide new effective tools for screening, diagnosis, treatment and monitoring of these diseases. Indeed, unlike the conventional medical biology practice based primarily on the sequential study of genes, proteins and metabolites, the great challenge of modern biology is to understand disease as a complex, integrated and dynamic network [1]. The concept of "metabolome" refers to the comprehensive complement of all metabolites present in a given biological system, fluid, cell or tissue [2]. So, metabolomics is one of the "omic" technologies based on biochemical characterizations of the metabolome and its changes related to genetic and environment factors. Metabolomics allows to characterize these interactions and to evaluate the biochemical mechanisms involved in such changes in a systematic fashion [3,4]. Given the strong link between IEM and metabolism, metabolomics is a very appealing tool to explore these diseases [5]. For years, mass spectrometry has been used in the assessment of inherited metabolic diseases [6-8]. However, few metabolomic research has been published in lysosomal storage diseases field. Lysosomal storage diseases (LSD) represent a group of about 50 inherited disorders due to lysosomal proteins deficiencies. This leads to a progressive accumulation of compounds within the lysosome. This metabolite storage causes various organ failures and premature death [9]. Mucopolysaccharidoses (MPS) belong to the LSD group. They are caused by impaired catabolism of glycosaminoglycans (GAGs), leading to their accumulation in lysosomes and extracellular matrix [10]. Then they are secreted into the bloodstream to be eliminated urine. Accumulated GAGs causes progressively multiple tissues and organs damage [11]. There are 11 known enzyme deficiencies, resulting in seven distinct forms of MPS [9]. Overall incidence is more than 1 in 7,500 live births. Most MPS patients are asymptomatic in newborn. Despite, acute neonatal symptoms may be observed in MPS I and MPS VII. MPS symptoms and severity vary with patients and MPS subtype. Several treatments are in clinical use or being investigated under clinical trials for patients with some types of MPS [12]. Mucopolysaccharidosis type I (MPS I) is a rare autosomal recessive disorder caused by α -L-Iduronidase (IDUA, EC 3.2.1.76) deficiency. IDUA degrades complex polysaccharides by removing a single α -L-iduronyl residue from heparan sulfate and dermatan sulfate. The symptoms range from the severe Hurler form [MPS IH - OMIM #67014] to the more attenuated Hurler-Scheie (MPS IH/S - OMIM #607015) and Scheie (MPS IS - OMIM #67016) phenotypes. The classification is mainly based on the age at first symptoms and the presence or not of mental retardation [13]. The average survival age is of 28 years which imply patient's shift from pediatrics to adults [14]. The aim of this study is to apply both targeted and untargeted metabolic profiling on mucopolysaccharidosis type I patients compared to controls and to treated patients to assess metabolic changes in this condition.

2. Materials and Methods

2.1. Urine samples

Random urine samples were collected from MPS patients in whom the diagnosis had been confirmed by demonstrating marked enzyme deficiency in leucocytes or by molecular analysis. Urine samples were collected within seven reference centers for inherited metabolic diseases in France. Nineteen untreated MPSI patients were evaluated: 18 males (age range from 1 to 43.6 years, mean age: 22 years) and 1 female (age 5.5 years). Control urine samples from 66 healthy subjects, 27 males and 40 females (age range from 5.5 to 70 years, mean age: 40.8 years). Fifteen samples from MPSI treated patients (MPSIT) with enzyme replacement therapy, 11 males and 4 females (age range from 1.3 to 39.3 years, mean age: 11.5 years) were analyzed for comparison. All samples were frozen at -80°C until analysis. This project was approved by the Research Ethics Board of Rouen University Hospital (CERNI E2016-21).

2.2. Metabolic phenotyping

2.2.1. Sample preparation

For untargeted metabolomics, urine samples were processed by transferring 200 μ L of urines to 1.5 mL tubes and centrifuged at 4°C for 10 min at 13,000 g then 100 μ L ultrapure water were added to 100 μ L of supernatant and mixed. For amino acids and GAGs analysis, detailed protocols are presented in supporting information.

2.2.2. Untargeted analysis

Ultraperformance liquid chromatography-ion mobility mass spectrometry (UPLC-IM-MS) and data-independent MS acquisition with simultaneous analysis of molecular fragmentation (MS^E) were performed on a Waters Synapt G2 HDMS (Saint Quentin en Yvelines, France) mass spectrometer equipped with a Waters nanoAcquity UPLC system and autosampler (Saint Quentin en Yvelines, France), as previously described [15]. Metabolites separation was carried out at 40 °C using a 1.0 x 100 mm, 1.8 μ m Acquity UPLC HSS T3 column (Waters), with a particle size of 1.8 μ m, equipped with a 0.2 μ m prefilter. Urine was eluted from the LC column using the following linear gradient (curve number 6): 0–1 min: 99% A; 1–3 min, 99–85% A; 3–6 min, 85–50% A; 6–9 min, 50–0% A; 9–12 min, 100% 12–16 min, 99% A for re-equilibration. Solvent A was water and solvent B was acetonitrile, both solvents contained 0.1% formic acid. For ion mobility, the helium cell gas flow, wave height, Trap Bias and IMS wave delay were set at 180 mL/min, 40 V, 45 V and 450 μ s respectively. The TOF analyzer was operated in the V resolution mode with an average mass resolution of $m/\Delta m$ 20,000 (full-width at half-maximum definition). Data acquisition of an ion mobility

experiment consisted of 200 bins. Collision Cross Section (CCS) values, obtained in nitrogen, were experimentally determined using singly charged Poly-DL-alanine oligomers as the TWIM calibrant species for ESI+. CCS values were derived according to previously reported procedures [16]. The ion mobility resolution was $\sim 40 \text{ } \Omega/\Delta\Omega$ (fwhm). The he CCS values reported were determined at the apex of the ion-mobility peak. Detailed instrument settings are presented in Table S-1 (Supporting Information). The mass spectrometer was operated in positive electrospray ionization mode. A mass range of m/z 50–1200 was used in both modes. The sample cone voltage, extraction cone voltage, source temperature, desolvation temperature, desolvation gas flow and cone gas flow were optimized as previously described [15]. Leucine enkephalin was used as the lock mass $[M-H]^+$ at m/z 556.2771. Sodium formate solution was used for external instrument calibration.

2.2.3. Raw data preprocessing

All LC-IM-MS raw data files data processing, peak detection and peak matching across samples using retention time (t_R) correction and chromatographic alignment along with drift time and CCS calculation were performed using Progenesis QI (Waters MS Technologies, Manchester, UK) to yield a data matrix containing retention times, accurate masses, CCS and peak intensities. The preprocessing step resulted in an X-matrix where t_R , CCS and m/z values were concatenated into “ t_R _m/z_CCS” features (in columns) present in each sample (in rows) with corresponding peak areas.

2.2.4. Quality Control

Ten μL of each urine sample are mixed together to generate a pooled quality control sample (QCs). QCs and solvent blank samples (mobile phase) were injected sequentially in-between the urine samples. In addition, a dilution series of QC samples (6%, 12.5%, 25%, 50% and 100% original concentration) are used to assess the quality of the extracted features. An analysis sequence is presented in Figure S1. In this study, we used a filter strategy in which the features intensity must be correlated to the matrix concentration in a series of diluted QC samples. Furthermore, RSD values derived from repeated measurements of a pooled QC sample were used. The threshold was set to $\text{RSD} < 25\%$. More details are presented as supporting information.

2.2.5. Targeted analysis

2.2.5.1. Amino acids quantification

The analysis of free amino acid profiles in urine was based on a liquid chromatography coupled to tandem mass spectrometry method and the aTRAQ reagent. The aTRAQ kit allows to quantify 24 free amino acids, both proteinogenic and non-proteinogenic, in a range of biological fluids. The analyses were performed using the liquid chromatography instrument Prominence Shimadzu LC system consisting of a DGU-20A3 degasser, a LC-20AB pump, a SIL-20A8HT autosampler, and a CTO-20AC oven (Shimadzu, Prominence, Kyoto, Japan) coupled to the 4000 QTRAP mass spectrometer (Sciex, Framingham, MA, USA) with an electrospray ion source. The detailed description of the applied LC-ESI-MS/MS methodology for amino acid

determination was presented in supplementary material. The amino acids concentrations were normalized by dividing them by creatininuria of the same urine sample.

2.2.5.2. Glycosaminoglycan quantification (HS, DS and KS)

Total urinary GAGs were measured with the dimethylmethylene blue-binding assay [17]. GAG-derived disaccharides in urine were analyzed using a multiplex assay as previously described by de Eveline et al [18]. Briefly, Heparan sulfate, dermatan sulfate and keratin sulfate were enzymatically digested into disaccharides. The disaccharides were quantified on a Waters Quattro Premier XE (tandem) mass spectrometer (Waters Corporation, Milford, MA, USA) coupled to an Acquity UPLC system (UPLC-MS/MS) with a ESI source. The disaccharides were separated on a Thermo Hypercarb HPLC column (100×2.1 mm, 5 µm). All disaccharides were detected and quantified in the MRM acquisition mode. The detailed description of the protocol is presented in [supplementary material](#). The GAG concentrations were normalized by dividing them by creatininuria determined in the same urine sample.

2.3. Statistical Analyses

A one-way ANOVA test was applied to each selected variable in order to confirm their actual difference between the three groups. A t-test is used when binary comparison is applied. Furthermore, the Benjamini and Hochberg false discovery rate (FDR) method was used for calculating the false-positive rate associated with multiple comparisons, and provides corrected q-values with a 0.05 significance level (FDR 5%). A Receiver operating characteristic curve (ROC) has been used to assess the diagnostic performance of the chosen classifiers.

2.4. Data analysis and modeling

Support vector regression normalization method was applied using the MetNormalizer R package [19] before any data analysis, to remove the unwanted intra- and inter-batch measurements analytical variations. The effect of this normalization step on the raw data is shown in Figures S-2, S-3 and S-4 (Supporting information). Then normalized data matrix has been log-transformed and pareto-scaled. All data analyzes and modeling were done using SIMCA 14.0 (MKS DAS, Umeå, Sweden) and R software. First, hierarchical cluster analysis and Principal Component Analysis (PCA) were used as exploratory unsupervised methods [20]. Orthogonal Partial Least-Squares-Discriminant Analysis (OPLSDA) was used a supervised method for predictive modelling purposes. Details regarding data modeling, characteristics and validation results from all OPLS-DA models are provided in supplemental material.

2.5. Feature selection and annotation

To select the most discriminant variables for the separation of groups, S-Plot was used. The S-plot combines the covariances and correlations between the **X** matrix and O-PLS scores for a given model component. The covariance values give the magnitude of contribution of a variable while the correlation values reflect the effect and reliability of the variable for the model component scores. Variables with both very high

correlation and covariance are important for the explanation power of the model. Selection of discriminant variables was achieved using the VIP (Variable Importance in Projection) scores procedure for each validated OPLS-DA model [21]. Putative annotation of detected features was performed using both accurate mass comparison using freely available metabolite databases HMDB, LipidBlast, KEGG, and Metlin. Furthermore, CCS values were also compared to the MetCCS database [22].

2.6. Pathway and network analysis

In order to provide a broader understanding of metabolic changes in MPSI, we also explored the biochemical pathways with a network analysis approach using Mummichog which allows pathway enrichment analyses. This Python package highlights pathways that are significantly impacted in the studied groups. Significantly impacted biochemical pathways are those exhibiting an adjusted p-value <0.05. For this comparison, we focused on features that significantly changed (q-values = 0.05 and FDR = 5%). Mummichog annotates metabolites based on accurate mass m/z and tests significant pathway enrichment within a reference metabolic network using a Fisher's exact test [23]. To protect against incorrect pathway selection, redundant pathways or those enriched by less than two metabolites were excluded. MetaboAnalyst [24] has been used for Metabolite Set Enrichment Analysis on the amino acids concentration matrix. **The Figure 1.** presents an overview of the adopted metabolomics workflow.

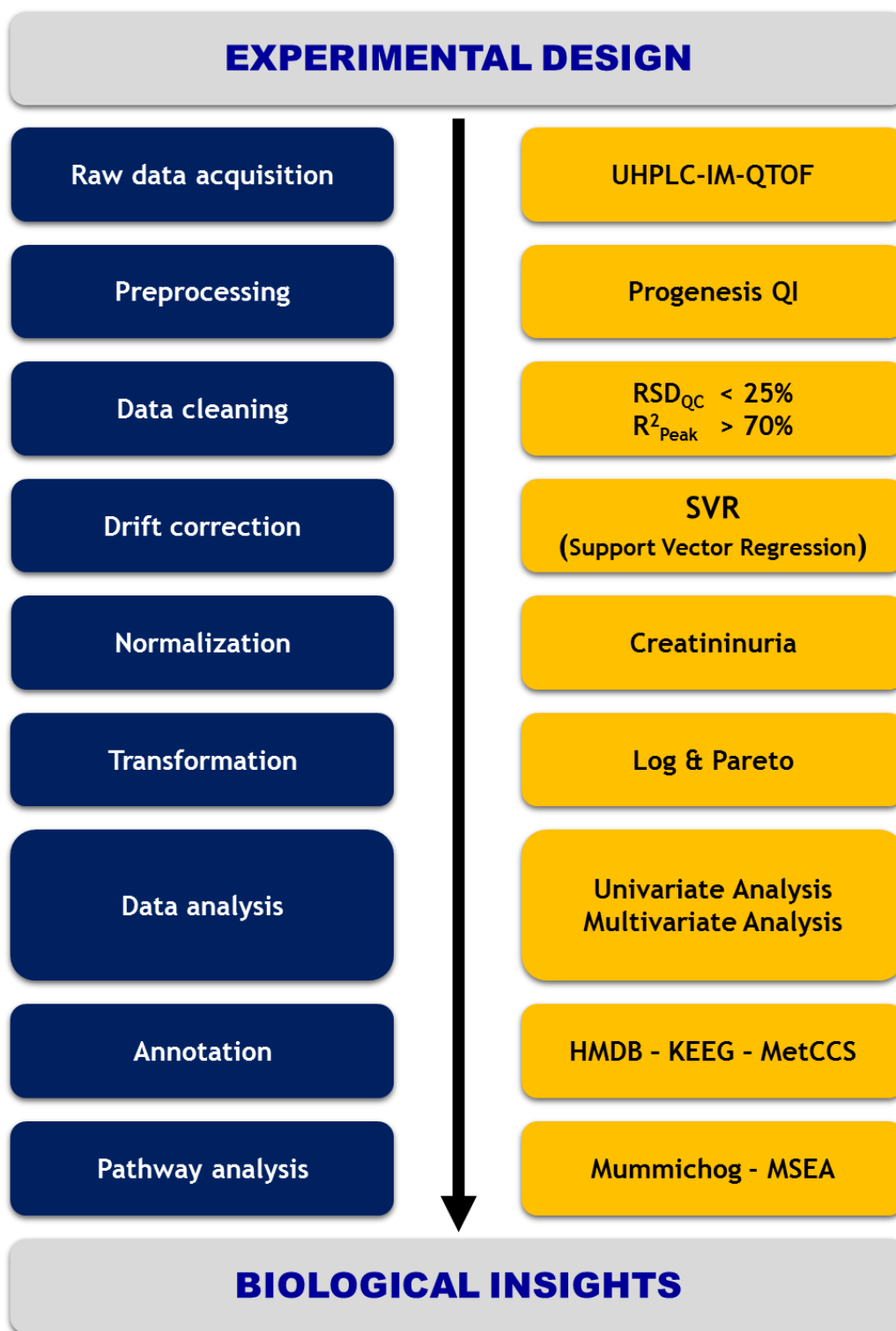


Figure 1. Illustration of the untargeted metabolomics workflow spanning from experimental design to pathway analysis and biological interpretation. HMDB: Human Metabolome Database. KEEG: Kyoto Encyclopedia of Genes and Genomes. MetCCS: Metabolite CCS database. MSEA: Metabolite Set Enrichment Analysis. RSD: Relative Standard Deviation.

3. Results

Untargeted analysis

The untargeted analysis of urine samples of control individuals, MPSI patients and MPSIT patients yielded 854 features. The analysis by independent ANOVA test resulted in 511 metabolites above the $p < 0.05$ cut-off (FDR 5%). A hierarchical clustering analysis was first applied to group samples with similar profiles of variable intensity. The heatmap in **Figure 2A** represents the top 100 features ranked by analysis of variance (ANOVA). The results show that all samples belonging to the same group were correctly clustered together. The dendrogram structure highlights two main clusters of variable intensities represented by its two longest branches (maximum dissimilarity according to the Euclidean distance). According to the color gradient, the intensity differences between groups are substantial. This first analysis allowed us to easily detect natural clusters in the data, although it did not facilitate extraction of discriminant variables among the all the dataset. To further explore natural separation between metabolic profiles and reduce data dimensionality, the dataset underwent a principal component analysis (PCA). The number of significant components was estimated using internal cross-validation with seven exclusion groups giving a three-component PCA model accounting for 21% of the total variance. The resulting scores plot was used to identify trends, groups and potential outliers within the data. Figure 2B shows PCA scores plot. There is a clear separation between MPSI and treated MPSI samples. However, there is an overlap with the control samples. Thus, to address our classification purposes, supervised methods are more suitable since they allow to accurately model the relationship between controls, MPS I and treated MPSI (MPSIT) samples. OPLS-DA classification was first applied to the data set. Samples were labelled according to the corresponding groups, MPSI, MPSIT and Control. A model was considered predictive if the Q₂ (cross-validation measure of the predictive power) regression line intercept resulting from the permutation test was negative. This means that the random labeled models exhibit lower predictive performance than the true one. The final model had an R² = 0.96 and Q₂ = 0.54. The OPLSDA scores plot (**Figure 3C**) revealed that each class was well separated, suggesting that the OPLS-DA model successfully discriminated samples according to their underlying metabolic profile. This model was internally validated both by CV-ANOVA (p -value = 4×10^{-20}) and by the permutation test (999 permutations gave a negative Q₂ intercept). Model validation details are shown in supplementary information (**Figure S5**). To go further in data modelling, binary OPLSDA classification models have been built. The first OPLSDA model was built using a dataset including Control and MPSI samples. The model had one predictive and two orthogonal components, and its validation parameters were as follows: R² = 0.94, Q₂ = 0.63 and CV-ANOVA p -value = 1.75×10^{-15} (**Figure S6**). The corresponding scores plot is shown in **Figure 3D**. It exhibited a clear separation between the two classes on the predictive component. A second OPLS-DA model was built using a dataset including MPSI and MPSIT samples following the same procedure. The OPLSDA model had 1 + 3 component model with R² = 0.97, Q₂ = 0.63 and CV-ANOVA p -value = 5×10^{-4} (**Figure S7**). Selection of discriminant variables was achieved using the VIP scores procedure

for each validated OPLS-DA model. Based on 1 as a cutoff value, 216 features out of the 854 were selected for the MPSI *vs.* Control model and 169 for the MPSI *vs.* MPSIT model. We then refined the two lists of variables by retaining only the most discriminant variables along with their putative annotation. The list included Carnitine, Arginine, Tetrahydrocorticosteron, Prolyl-Lysine, S-(5'-Adenosyl)-L-Methionine, oleic acid and Phenylalanylalanine.

These discriminant variables are presented in **Table 1** for both models along with their respective statistical metrics and annotation accuracy. Boxplot of the main discriminant features are presented in **Figure S-8**. The discriminant performance of these features is also assessed using area under the receiver operating characteristic (ROC) curves. Carnitine has the highest AUC (AUC = 0.93). The overall ROC results are shown in **Figure S9**. Furthermore, to explore the underlying pathways dysregulated in MPS I we used the Mummichog software to look for significant pathways related to variation in the significantly disturbed features. Different metabolism pathways were affected such as glycerophospholipid metabolism, vitamins and amino acids are shown in **Table 2**. Interestingly, a series of amino acid metabolic pathways were markedly dysregulated.

MPS I vs. Control

Pathway	Overlap size	p-value (FDR = 5%)
Aspartate and asparagine metabolism	13	0.0075
Lysine metabolism	9	0.0128
Glutathione Metabolism	3	0.0145
Vitamin H (biotin) metabolism	3	0.0145
De novo fatty acid biosynthesis	6	0.0280
Tyrosine metabolism	19	0.0357
Ascorbate (Vitamin C) and Aldarate Metabolism	4	0.0359
Omega-3 fatty acid metabolism	3	0.0449
Vitamin B5 - CoA biosynthesis from pantothenate	3	0.0449

MPS I vs. MPS I treated

Pathway	Overlap size	p-value (FDR = 5%)
Lysine metabolism	13	0.0014
Glycerophospholipid metabolism	11	0.0028
Methionine and cysteine metabolism	8	0.0123
Tyrosine metabolism	26	0.0216
Biopterin metabolism	7	0.0222
Urea cycle/amino group metabolism	14	0.0260
Ascorbate (Vitamin C) and Aldarate Metabolism	5	0.0264
Arginine and Proline Metabolism	8	0.0341
Glutathione Metabolism	3	0.0388
Vitamin H (biotin) metabolism	3	0.0388

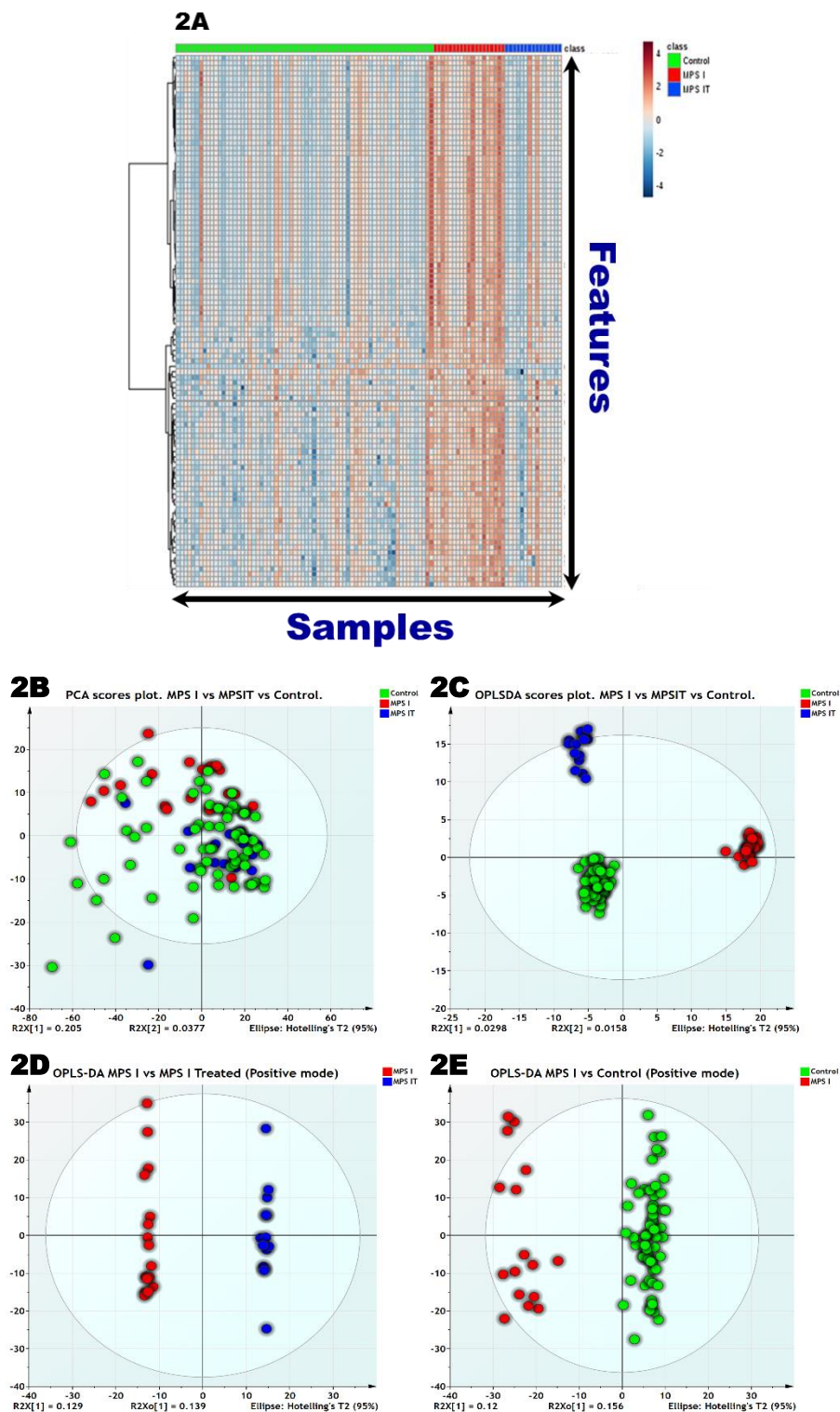


Figure 2. 2A) Hierarchical cluster analysis and heat map visualization of top 100 variables (y-axis) ranked by ANOVA. The urine sample class within are represented along the x-axis. The color code was used to represent log-scaled intensities of features between -4 (blue) and +4 (brown), showing the features relative abundance according to the groups. 2B) PCA scores plot of the normalized dataset. The three groups are represented by different colors. MPS I and MPSI treated samples are well separated on PC1 according to their class membership however. Control samples show an overlap. 2C) OPLSDA scores plot ($R^2 = 0.96$, $Q^2 = 0.54$) shows a clear separation between the different groups. PC 1 separates the MPS I samples from the controls. However, PC2 separates treated MPS I from control samples. 2D) Clear separation between treated MPS I from treated samples is observed ($R^2 = 0.97$, $Q^2 = 0.63$). 2E) Clear separation between MPS I samples from the controls is observed ($R^2 = 0.94$, $Q^2 = 0.63$). Detailed model characteristics and validation are given in supporting information.

Table 1. Some discriminant features extracted by OPLSDA models allowing the discrimination of control subjects, MPS I and MPS IT.

MPS I vs. Control												
HMDB	Putative annotation	Formula	Mr	m/z	Adduct	Δ m/z (ppm)	t_r	D_i	CCS	FDR	%RSD	VIP
HMDB00062	Carnitine	C ₇ H ₁₅ NO ₃	161.1053	203.1518	M+CH ₃ OH+H	0.48	1.41	2.43	145.1	4.83E-10	10.0	1.90
HMDB00207	Oleic acid	C ₁₈ H ₃₄ O ₂	282.2546	283.2616	M+H	-4.36	10.75	4.00	189.2	1.02E-06	24.53	1.85
HMDB29022	Prolyl-Lysine	C ₁₁ H ₂₁ N ₃ O ₃	243.1595	282.1226	M+K	4.88	1.49	3.08	163.1	6.60E-09	5.97	1.73
HMDB00268	Tetrahydrocortison	C ₂₁ H ₃₄ O ₄	350.2465	351.2538	M+H	2.33	10.75	4.48	200.0	1.13E-04	17.26	1.51
HMDB01185	S-(5'-Adenosyl)-L-Methionine	C ₁₅ H ₂₃ N ₆ O ₅ S	398.1375	399.1447	M+H	0.59	5.26	4.27	193.7	2.08E-04	9.8	1.34
HMDB28988	Phenylalanylanine	C ₁₂ H ₁₆ N ₂ O ₃	236.1152	237.1225	M+H	-3.62	7.67	2.70	152.7	1.49E-06	10.51	1.18
MPS I vs. MPS IT												
HMDB	Putative annotation	Formula	Mr	m/z	Adduct	Δ m/z (ppm)	t_r	D_i	CCS	FDR	%RSD	VIP
HMDB00517	Arginine	C ₆ H ₁₄ N ₄ O ₂	174.1112	175.1212	M+H	-2.70	1.23	2.11	135.0	7.31E-06	22.62	2.38
HMDB00207	Oleic acid	C ₁₈ H ₃₄ O ₂	282.2546	283.2616	M+H	-4.36	10.75	4.00	189.2	1.02E-06	24.53	1.93
HMDB00062	Carnitine	C ₇ H ₁₅ NO ₃	161.1053	203.1518	M+CH ₃ OH+H	0.48	1.41	2.43	145.1	4.83E-10	9.97	1.93
HMDB28988	Phenylalanylanine	C ₁₂ H ₁₆ N ₂ O ₃	236.1152	237.1225	M+H	-3.62	7.67	2.70	152.7	1.49E-06	10.51	1.91
HMDB29022	Prolyl-Lysine	C ₁₁ H ₂₁ N ₃ O ₃	243.1595	282.1226	M+K	4.88	1.49	3.08	163.1	6.60E-09	5.97	1.77
HMDB00268	Tetrahydrocortison	C ₂₁ H ₃₄ O ₄	350.2465	351.2538	M+H	2.33	10.75	4.48	200.0	1.13E-04	17.26	1.61
HMDB01185	S-(5'-Adenosyl)-L-Methionine	C ₁₅ H ₂₃ N ₆ O ₅ S	398.1375	399.1447	M+H	0.59	5.26	4.27	193.7	2.08E-04	9.8	1.34

Targeted analysis

The first targeted analysis addressed urinary glycosaminoglycans concentrations. As expected, total GAGs, dermatan sulfate (DS) and Heparan sulfate (HS) were significantly elevated in MPSI patients (**Table 3** and **Figure S11**). Of note, Keratan sulfate (KS) is slightly elevated in MPS I patients compared to control samples. Given these results of the above untargeted approach, we also performed a targeted amino acid profiling on all the samples. Free amino acid profiles in urine samples of MPSI, MPSIT patients and the control group were obtained. Twenty-four amino acids were quantified in all samples and their concentrations were subjected to subsequent statistical and pathway analysis. **Table 3S** presents absolute urine concentrations of amino acids. Boxplot of normalized amino acid concentrations are presented in **Figure S10**. The statistical analysis of amino acids is listed in **Table 3**. Regarding Control *vs.* MPS I comparison, nine amino acids have showed significant difference between the two groups; Glutamic acid, Arginine, Aspartic acid, Proline, Valine, Alanine, Glycine and Threonine. Besides, six amino acids showed statistically different concentrations between MPSI and MPSIT samples; Glutamic acid, Aspartic acid, Valine, Alanine and Isoleucine. To determine the amino acids profile differences between controls and MPSI patients and MPSI patients, the 24 amino acids were first analyzed by an ANOVA test. The analysis yielded eight amino acids above the $p < .05$ cut-off (FDR 5%). A hierarchical clustering analysis was first applied to group samples with similar profiles. The heatmap in **Figure 3A** represents the 24 amino acids ranked by analysis of variance (ANOVA). Even though, there is no an obvious visual pattern, the results show that all samples belonging to the same group were correctly clustered together. The dendrogram structure highlights two main clusters of variables represented by its two longest branches (maximum dissimilarity according to the Euclidean distance). Furthermore, a correlation analysis of the overall concentrations matrix has been performed. **Figures 3B** and **3C** presents the heatmap of the correlation analysis. Both figures show a clear cluster of variables that have high correlation. **Figure 3B** (MPSI *vs* Control) show a main cluster including Alanine, Leucine, Valine, Tyrosine, Threonine, Isoleucine, Histidine, Tryptophane, Serine, Asparagine, Glutamine, Phenylalanine. Regarding MPSI *vs* MPS IT, **Figure 3C** shows two main clusters. One including Isoleucine, Asparagine, Threonine, Serine, Tyrosine, Glutamine and Tryptophane. The second includes Alanine, Phenylalanine, Leucine, Valine, Glycine and Histidine. To assess the diagnostic performance of the different amino acids, a univariate ROC curve analyses for MPSI *vs.* Control groups indicated four amino acids with high AUC above 0.80 and are: Arginine (AUC = 0.91), Glutamic acid (AUC = 0.88), Aspartic acid (AUC = 0.83) and Proline (AUC = 0.83). The same procedure has been performed for MPSI *vs.* MPS IT groups and indicated tow amino acids with high AUC above 0.80 and were: Aspartic acid (AUC = 0.85) and Isoleucine (AUC = 0.82). The overall univariate and ROC analysis results are presented in **Table 3**. A comparison of different combinations of the main significant amino acids using a PLSDA model with three components each is presented in **Figure S12**. Furthermore, HS, KS, DS and total GAGs showed significant change between MPS I and control samples with DS as the most discriminant ($p < 0.0001$). Moreover, a pathway analysis has been performed on the amino acids concentrations dataset. Using these quantitative data, we performed pathway analysis that yielded the main impaired metabolisms. For MPSI *vs.* Control analysis, Arginine and Proline, Malate-Aspartate Schuttle, Cysteine, Urea cycle and alanine

metabolism were the most affected pathways. Regarding, MPSI vs. MPSIT analysis, Alanine, Malate-Aspartate Schuttle, branched amino acids metabolisms were the most affected. The overall results are shown in **Figure 4A** and **4B** for all the studied groups.

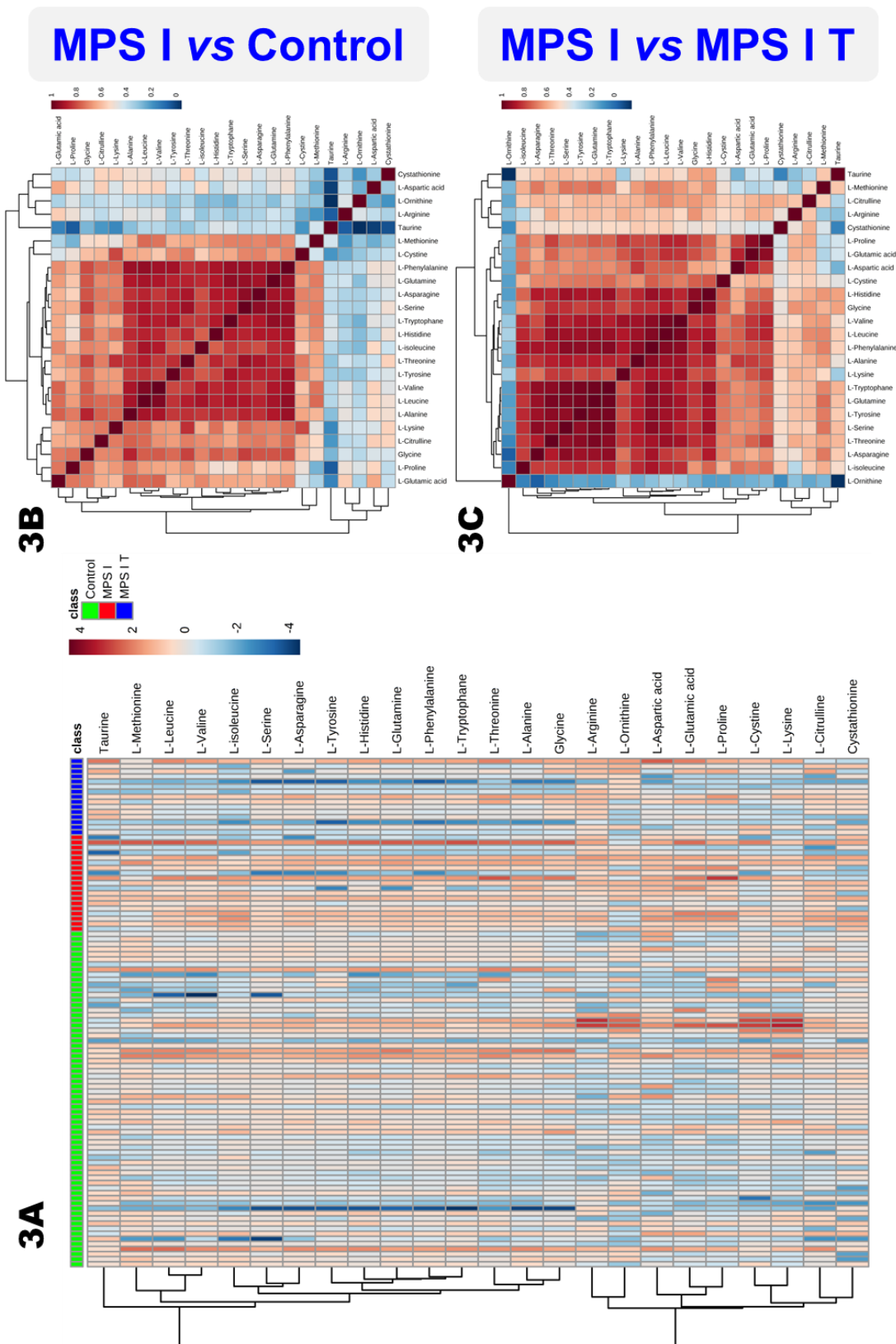


Figure 3. A) Heat map representing the clustering of 24 amino acids across the 3 groups of samples (MPS I, MPS IT and Controls). Columns represent individual samples and rows refer to amino acid. Shades of red or blue represent elevation or decrease, respectively, of an amino acid. B and C) Spearman rank-order correlation matrix 24 amino acids based on their concentrations profiles across all samples. Shades of red or blue represent low-to-high correlation coefficient between markers. R) MPS vs Control. C) MPS I vs MPS IT.

Table 3. T-test statistics, fold change and area under the curve of the receiver operating curves (ROC) for 24 amino acids and the GAGs. ($p < 0.05$). Significant features are highlighted in bold. (FDR = 5%)

	MPS I vs Control			MPS I vs Treated MPS I		
	AUC	q-value (FDR)	Fold Change	AUC	q-value (FDR)	Fold Change
<i>Amino acids</i>						
L-Arginine	0.91	1.50E-04	-2.20	0.51	4.78E-01	-0.41
L-Glutamic acid	0.88	2.37E-05	-0.38	0.75	3.58E-02	0.28
L-Aspartic acid	0.83	2.22E-04	-0.75	0.86	1.52E-02	1.30
L-Proline	0.83	2.22E-04	-0.52	0.73	7.34E-02	0.43
L-Valine	0.79	6.57E-03	-0.01	0.78	3.33E-02	0.50
L-Alanine	0.78	1.34E-02	0.15	0.76	3.58E-02	0.83
L-Leucine	0.76	2.62E-02	0.28	0.76	4.72E-02	0.35
Glycine	0.74	3.75E-02	0.40	0.58	2.01E-01	0.04
L-Glutamine	0.73	9.44E-02	0.49	0.68	2.01E-01	0.15
L-Phenylalanine	0.72	2.70E-01	0.62	0.75	1.51E-01	0.70
L-isoleucine	0.72	1.23E-01	0.32	0.83	1.85E-02	1.81
L-Threonine	0.72	3.75E-02	0.31	0.67	2.01E-01	0.12
L-Serine	0.71	3.42E-01	0.71	0.68	2.82E-01	0.40
L-Histidine	0.70	1.78E-01	0.52	0.61	2.82E-01	-0.05
L-Ornithine	0.69	3.42E-01	-0.46	0.57	7.94E-01	-0.19
L-Asparagine	0.68	4.40E-01	0.57	0.70	2.31E-01	0.74
L-Lysine	0.68	3.51E-01	0.39	0.75	7.34E-02	0.47
L-Citrulline	0.68	3.20E-01	0.41	0.66	3.96E-01	0.19
L-Tryptophane	0.67	3.20E-01	0.80	0.66	2.08E-01	0.16
L-Tyrosine	0.64	5.95E-01	1.02	0.69	2.08E-01	0.55
Cystathionine	0.63	3.42E-01	0.61	0.65	3.65E-01	0.15

Taurine	0.56	3.20E-01	2.02	0.59	4.00E-01	-1.95
L-Cystine	0.54	6.48E-01	1.30	0.58	5.65E-01	-0.31
L-Methionine	0.52	5.30E-01	0.93	0.65	1.93E-01	0.09

GAGs

Total GAGs	0.92	4.76E-04	-3.74	0.79	6.39E-01	1.56
Dermatan sulfate	0.92	1.37E-06	-5.14	0.78	3.65E-01	2.48
Keratan sulfate	0.91	2.75E-02	-1.84	0.77	5.65E-01	1.24
Heparan sulfate	0.89	1.14E-03	-3.13	0.75	6.02E-01	1.46

To go further in the pathway analysis, we performed a comparative analysis between the pathway analysis results from both untargeted and targeted results along with the recently *in-silico* systems analysis data reported by Salazar *et al.* [25]. Salazar *et al.* performed a system biology approach using a genome-scale human metabolic reconstruction to understand the effect of metabolism alterations in MPS. The *In-silico* MPSI model was generated by silencing *IDUA* gene of MPSI related enzyme, and were analyzed through a flux balance and variability analysis. Thus, to depict the interrelationships between our untargeted and targeted results along with these *in-silico* metabolic impairment data, we used a Venn diagram approach. Two main metabolisms were identified using the three approaches: Arginine and Proline metabolism and Glutathione metabolism (Figure 6C). Detailed data are presented in Table S4.

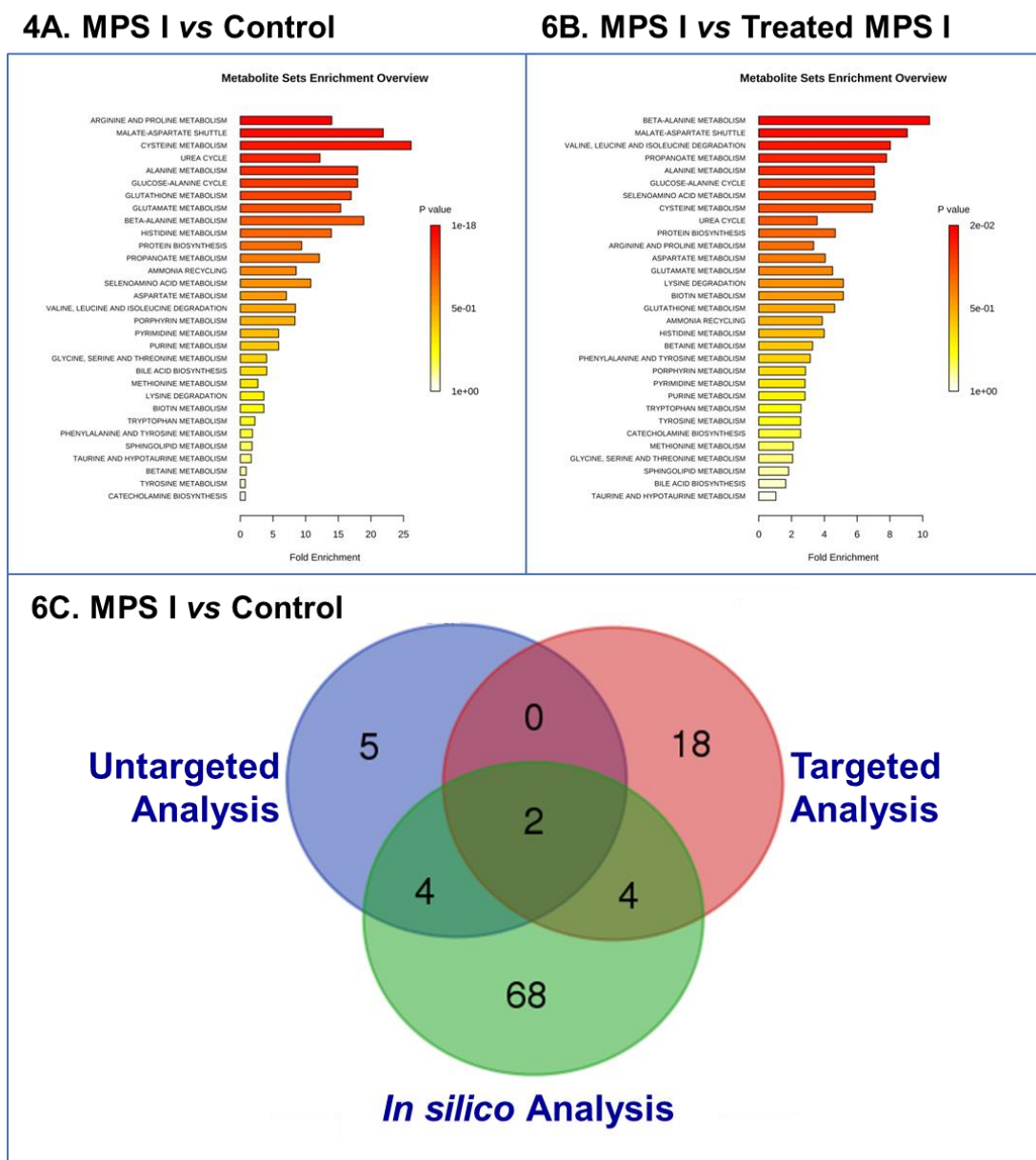


Figure 4. Metabolite Set Enrichment Analysis using amino acid concentrations. 4A) MPS I vs. Control. 4B) MPS I vs. MPS II. 4C) Venn diagram of the significant pathways retrieved from untargeted, targeted approaches and *in silico* systems biology approach from Salazar *et al.* (2013). The diagram shows two common metabolisms: Glutathione and Arginine and proline metabolisms. Detailed pathway information is given in supporting information (Table S4).

4. Discussion

In this study, the potential of metabolomics to identify biomarkers related to MPSI in urine was investigated. The data demonstrates that lysosomal accumulation of GAGs triggers deep metabolic turnover in MPSI patients with and urinary global metabolomics profiling may provide better understanding MPSI disease mechanisms and may pave the way for potential biomarkers. The metabolic alterations seen were mainly relevant to amino acid pathways contributing significantly to the clear discrimination of the different studied groups, MPSI, MPSIT and Control samples using their metabolic differences. Indeed, based on the untargeted urinary metabolic profiles retrieved from the different studied groups, we were able to build a predictive model that clearly separates the different studied groups, MPSI, MPSIT and Control samples using their metabolic differences. This study showed metabolic impairments mainly amino acid metabolism and related metabolisms such as vitamin and Glutathione metabolisms. In the light of these results, we performed a targeted analysis focusing on amino acid profiles which confirmed the amino acid profile alterations. This pathway analysis yielded different dysregulated metabolic pathways. Most important metabolisms are arginine-proline metabolism (**Figure S13**), cysteine-glutathione metabolism (**Figure S14**). The latter metabolism is tightly linked to oxidative stress, recent studies have shown oxidative damage involvement in the pathophysiology of several genetic diseases, including in LSD [26]. Indeed, the GAGs biosynthesis requires recycled substrates, however in MPSs these recycled substrates are lacking which may lead to an increase in cellular energy needs [27]. This energy requirement may trigger an active mitochondrion turnover with finally an excess production of reactive oxidative species. Oxidative stress has been recognized as a mechanism of cell damage in MPSs and has been reported in MPS patients undergoing ERT [28-32]. Moreover, oxidative stress has been observed in MPSI mouse model [33,34]. Pereira *et al.* assessed oxidative stress in MPSI patients, compared with control subjects [28]. The authors detected a decrease in superoxide dismutase activity in erythrocytes from MP I patients after ERT, while Catalase activity increased after ERT compared to baseline levels. This could be explained through antioxidant pathway impairments due to cellular damages following GAG accumulation. These findings could suggest that potential antioxidants that might be included as adjuvant for current MPS therapies. Regarding arginine, its classification performance is interestingly comparable with that of the quantified GAGs (**Table 3**). Arginine is an amino acid which is involved in several key metabolisms, urea cycle, nitric oxide, polyamines, glutamate, proline and homoarginine. Furthermore, changes in arginine levels act as nutritional sensors and regulate cellular metabolism through its interaction with mammalian target of rapamycin complex 1 (mTORC1) [35]. Recently, Chantranupong *et al.* demonstrated that arginine sensing by mTORC1 is specific and depends on CASTOR1. CASTOR1 interacts with GATOR2 to inhibit mTORC1 in low arginine condition. In the presence of arginine, CASTOR1 is bound to arginine, and thus free up GATOR2 and activate mTORC1 [36]. mTORC1 is already defined as a key regulator of protein synthesis, cell growth and autophagy [37]. Autophagy has been described as impaired in several LSDs [38] which may be attributed to

lysosome dysfunction but may be also linked to arginine metabolism impairment. Woloszynek *et al.* observed profound metabolic alterations in energy expenditure in MPSI mice, similar to those observed in the current study with an increase in most amino acids concentrations including dipeptides, amino acid derivatives, and urea [39]. The authors attributed these changes to an increase protein catabolism and an autophagy disruption as a consequence of lysosome dysfunction. Interestingly, autophagic vacuoles number are increased in several LSDs, and reduced in MPSI mice fed a high fat diet [39].

5. Conclusion

Metabolic phenotyping might help generate new hypotheses about MPS etiology and develop further targeted studies of a smaller number of potentially important metabolites. These findings enabled us to unveil profound metabolic impairments beyond the primary deficiency in MPSI. The pathophysiological bases understanding this disease may open new therapeutic strategies such as antioxidants adjuvants and diet intervention as complementary treatments for MPS and may be other LSDs.

Acknowledgements

The authors are grateful to Carine Pilon, Thomas Plichet and Tony Pereira (Rouen University Hospital, France) for their technical assistance. The authors gratefully acknowledge the Region Haute-Normandie, the Labex SynOrg (ANR-11-LABX-0029), the European Regional Development Fund (ERDF 31708), Sanofi-Genzyme, Shire and Vaincre les Maladies Lysosmales for financial support.

RÉFÉRENCES

1. Tebani, A.; Afonso, C.; Marret, S.; Bekri, S. Omics-based strategies in precision medicine: Toward a paradigm shift in inborn errors of metabolism investigations. *Int J Mol Sci* **2016**, *17*.
2. Nicholson, J.K.; Lindon, J.C.; Holmes, E. 'Metabonomics': Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological nmr spectroscopic data. *Xenobiotica; the fate of foreign compounds in biological systems* **1999**, *29*, 1181-1189.
3. Dunn, W.B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J.D.; Halsall, A.; Haselden, J.N., *et al.* Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature protocols* **2011**, *6*, 1060-1083.
4. Benton, H.P.; Want, E.; Keun, H.C.; Amberg, A.; Plumb, R.S.; Goldfain-Blanc, F.; Walther, B.; Reily, M.D.; Lindon, J.C.; Holmes, E., *et al.* Intra- and interlaboratory reproducibility of ultra performance liquid chromatography-time-of-flight mass spectrometry for urinary metabolic profiling. *Analytical chemistry* **2012**, *84*, 2424-2432.
5. Tebani, A.; Abily-Donval, L.; Afonso, C.; Marret, S.; Bekri, S. Clinical metabolomics: The new metabolic window for inborn errors of metabolism investigations in the post-genomic era. *Int J Mol Sci* **2016**, *17*.
6. Pitt, J.J. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *The Clinical biochemist. Reviews / Australian Association of Clinical Biochemists* **2009**, *30*, 19-34.

7. Pitt, J.J. Newborn screening. *The Clinical biochemist. Reviews / Australian Association of Clinical Biochemists* **2010**, *31*, 57-68.
8. Spacil, Z.; Tatipaka, H.; Barcenas, M.; Scott, C.R.; Turecek, F.; Gelb, M.H. High-throughput assay of 9 lysosomal enzymes for newborn screening. *Clin Chem* **2013**, *59*, 502-511.
9. Ballabio, A.; Gieselmann, V. Lysosomal disorders: From storage to cellular damage. *Biochimica et biophysica acta* **2009**, *1793*, 684-696.
10. Wraith, J.E. The mucopolysaccharidoses: A clinical review and guide to management. *Archives of disease in childhood* **1995**, *72*, 263-267.
11. Neufeld EF, M.J. The mucopolysaccharidoses. In *The metabolic and molecular basis of inherited disease*, Scriver C, B.A., Sly W, Vaele D, Ed. Mc Graw-Hill: New York, NY, 2001; pp 3421-3452.
12. Parenti, G.; Andria, G.; Ballabio, A. Lysosomal storage diseases: From pathophysiology to therapy. *Annu Rev Med* **2015**, *66*, 471-486.
13. Tebani, A.; Zanoutene-Cheriet, L.; Adjoutah, Z.; Abily-Donval, L.; Brasse-Lagnel, C.; Laquerriere, A.; Marret, S.; Chalabi Benabdellah, A.; Bekri, S. Clinical and molecular characterization of patients with mucopolysaccharidosis type i in an algerian series. *Int J Mol Sci* **2016**, *17*.
14. OA, B. Clinical characteristics of mps i patients in the mps i registry. 2007 san diego. *The American Society of Human Genetics* **2007**.
15. Tebani, A.; Schmitz-Afonso, I.; Rutledge, D.N.; Gonzalez, B.J.; Bekri, S.; Afonso, C. Optimization of a liquid chromatography ion mobility-mass spectrometry method for untargeted metabolomics using experimental design and multivariate data analysis. *Analytica Chimica Acta* **2016**, *913*, 55-62.
16. Paglia, G.; Angel, P.; Williams, J.P.; Richardson, K.; Olivos, H.J.; Thompson, J.W.; Menikarachchi, L.; Lai, S.; Walsh, C.; Moseley, A., *et al.* Ion mobility-derived collision cross section as an additional measure for lipid fingerprinting and identification. *Analytical chemistry* **2015**, *87*, 1137-1144.
17. de Jong, J.G.; Wevers, R.A.; Liebrand-van Sambeek, R. Measuring urinary glycosaminoglycans in the presence of protein: An improved screening procedure for mucopolysaccharidoses based on dimethylmethylene blue. *Clin Chem* **1992**, *38*, 803-807.
18. Langereis, E.J.; Wagemans, T.; Kulik, W.; Lefeber, D.J.; van Lenthe, H.; Oussoren, E.; van der Ploeg, A.T.; Ruijter, G.J.; Wevers, R.A.; Wijburg, F.A., *et al.* A multiplex assay for the diagnosis of mucopolysaccharidoses and mucopolipidoses. *PLoS one* **2015**, *10*, e0138622.
19. Shen, X.; Gong, X.; Cai, Y.; Guo, Y.; Tu, J.; Li, H.; Zhang, T.; Wang, J.; Xue, F.; Zhu, Z.-J. Normalization and integration of large-scale metabolomics data using support vector regression. *Metabolomics* **2016**, *12*, 89.
20. Eriksson, L.; Trygg, J.; Wold, S. A chemometrics toolbox based on projections and latent variables. *Journal of Chemometrics* **2014**, *28*, 332-346.
21. Galindo-Prieto, B.; Eriksson, L.; Trygg, J. Variable influence on projection (vip) for orthogonal projections to latent structures (opls). *Journal of Chemometrics* **2014**.
22. Zhou, Z.; Xiong, X.; Zhu, Z.J. Metccs predictor: A web server for predicting collision cross-section values of metabolites in ion mobility-mass spectrometry based metabolomics. *Bioinformatics (Oxford, England)* **2017**.
23. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F.H.; Khan, N.; Soltow, Q.A.; Jones, D.P.; Pulendran, B. Predicting network activity from high throughput metabolomics. *PLoS Comput. Biol.* **2013**, *9*, e1003123.
24. Xia, J.; Sinelnikov, I.V.; Han, B.; Wishart, D.S. Metaboanalyst 3.0-making metabolomics more meaningful. *Nucleic acids research* **2015**.
25. Salazar, D.A.; Rodriguez-Lopez, A.; Herreno, A.; Barbosa, H.; Herrera, J.; Ardila, A.; Barreto, G.E.; Gonzalez, J.; Almeciga-Diaz, C.J. Systems biology study of mucopolysaccharidosis using a human metabolic reconstruction network. *Molecular genetics and metabolism* **2016**, *117*, 129-139.

26. Donida, B.; Jacques, C.E.D.; Mescka, C.P.; Rodrigues, D.G.B.; Marchetti, D.P.; Ribas, G.; Giugliani, R.; Vargas, C.R. Oxidative damage and redox in lysosomal storage disorders: Biochemical markers. *Clinica Chimica Acta* **2017**, *466*, 46-53.
27. Woloszynek, J.C.; Kovacs, A.; Ohlemiller, K.K.; Roberts, M.; Sands, M.S. Metabolic adaptations to interrupted glycosaminoglycan recycling. *The Journal of biological chemistry* **2009**, *284*, 29684-29691.
28. Pereira, V.G.; Martins, A.M.; Micheletti, C.; D'Almeida, V. Mutational and oxidative stress analysis in patients with mucopolysaccharidosis type i undergoing enzyme replacement therapy. *Clinica chimica acta; international journal of clinical chemistry* **2008**, *387*, 75-79.
29. Jacques, C.E.; Donida, B.; Mescka, C.P.; Rodrigues, D.G.; Marchetti, D.P.; Bitencourt, F.H.; Burin, M.G.; de Souza, C.F.; Giugliani, R.; Vargas, C.R. Oxidative and nitrate stress and pro-inflammatory cytokines in mucopolysaccharidosis type ii patients: Effect of long-term enzyme replacement therapy and relation with glycosaminoglycan accumulation. *Biochimica et biophysica acta* **2016**, *1862*, 1608-1616.
30. Negretto, G.W.; Deon, M.; Biancini, G.B.; Burin, M.G.; Giugliani, R.; Vargas, C.R. Glycosaminoglycans can be associated with oxidative damage in mucopolysaccharidosis ii patients submitted to enzyme replacement therapy. *Cell biology and toxicology* **2014**, *30*, 189-193.
31. Filippin, L.; Wayhs, C.A.; Atik, D.M.; Manfredini, V.; Herber, S.; Carvalho, C.G.; Schwartz, I.V.; Giugliani, R.; Vargas, C.R. DNA damage in leukocytes from pretreatment mucopolysaccharidosis type ii patients; protective effect of enzyme replacement therapy. *Mutation research* **2011**, *721*, 206-210.
32. Donida, B.; Marchetti, D.P.; Biancini, G.B.; Deon, M.; Manini, P.R.; da Rosa, H.T.; Moura, D.J.; Saffi, J.; Bender, F.; Burin, M.G., *et al.* Oxidative stress and inflammation in mucopolysaccharidosis type iva patients treated with enzyme replacement therapy. *Biochimica et biophysica acta* **2015**, *1852*, 1012-1019.
33. Simonaro, C.M.; D'Angelo, M.; He, X.; Eliyahu, E.; Shtraizent, N.; Haskins, M.E.; Schuchman, E.H. Mechanism of glycosaminoglycan-mediated bone and joint disease: Implications for the mucopolysaccharidoses and other connective tissue diseases. *The American journal of pathology* **2008**, *172*, 112-122.
34. Reolon, G.K.; Reinke, A.; de Oliveira, M.R.; Braga, L.M.; Camassola, M.; Andrades, M.É.; Moreira, J.C.F.; Nardi, N.B.; Roesler, R.; Dal-Pizzol, F. Alterations in oxidative markers in the cerebellum and peripheral organs in mps i mice. *Cellular and molecular neurobiology* **2009**, *29*, 443-448.
35. Morris, S.M., Jr. Arginine metabolism revisited. *The Journal of nutrition* **2016**, *146*, 2579s-2586s.
36. Chantranupong, L.; Scaria, S.M.; Saxton, R.A.; Gygi, M.P.; Shen, K.; Wyant, G.A.; Wang, T.; Harper, J.W.; Gygi, S.P.; Sabatini, D.M. The castor proteins are arginine sensors for the mtorc1 pathway. *Cell* **2016**, *165*, 153-164.
37. Goberdhan, D.C.; Wilson, C.; Harris, A.L. Amino acid sensing by mtorc1: Intracellular transporters mark the spot. *Cell metabolism* **2016**, *23*, 580-589.
38. Chévrier, M.; Brakch, N.; Céline, L.; Genty, D.; Ramdani, Y.; Moll, S.; Djavaheri-Mergny, M.; Brasse-Lagnel, C.; Annie Laquerrière, A.L.; Barbey, F. Autophagosome maturation is impaired in fabry disease. *Autophagy* **2010**, *6*, 589-599.
39. Woloszynek, J.C.; Kovacs, A.; Ohlemiller, K.K.; Roberts, M.; Sands, M.S. Metabolic adaptations to interrupted glycosaminoglycan recycling. *The Journal of biological chemistry* **2009**, *284*, 29684-29691.

Supplementary material for:

Urinary metabolic phenotyping of mucopolysaccharidosis type I combining untargeted and targeted strategies with data modeling

Abdellah TEBANI^{1,2,3}, Isabelle SCHMITZ-AFONSO³, Lenaig ABILY-DONVAL^{2,4}, Bénédicte HERON⁵, Monique PIRAUD⁶, Jérôme AUSSEIL⁷, Anais BRASSIER⁸, Pascale De LONLAY⁸, Farid ZERMICHE⁹, Frédéric M. Vaz¹⁰, Bruno J. GONZALEZ², Stephane MARRET^{2,4}, Carlos AFONSO³, Soumeya BEKRI^{1,2,*}

¹ Department of Metabolic Biochemistry, Rouen University Hospital, Rouen, 76000, France

² Normandie Univ, UNIROUEN, CHU Rouen, INSERM U1245, 76000 Rouen, France

³ Normandie Univ, UNIROUEN, INSA Rouen, CNRS, COBRA, 76000 Rouen, France

⁴ Department of Neonatal Pediatrics and Intensive Care, Rouen University Hospital, Rouen, 76031, France

⁵ Centre de Référence des Maladies Lysosomales (CRML), Neuropédiatrie, CHU Trousseau, APHP, Paris, France.

⁶ Service Maladies Héritaires du Métabolisme et Dépistage Néonatal, Centre de Biologie et de Pathologie Est CHU de Lyon, Lyon, France.

⁷ INSERM U1088, Laboratoire de Biochimie Spécialisée, Centre de Biologie Humaine, CHU Sud, 80054, Amiens Cedex, France.

⁸ Reference Center of Inherited Metabolic Diseases, Imagine Institute, Hospital Necker Enfants Malades, APHP, University Paris Descartes, Paris, France.

⁹ Laboratoire de Biochimie et Biologie Moléculaire, Université de Lille et Pôle de Biologie Pathologie Génétique du CHRU de Lille, 59000, Lille, France.

¹⁰ Department of Clinical Chemistry and Pediatrics, Laboratory Genetic Metabolic Diseases, Academic Medical Center, Amsterdam, The Netherlands

Technical details

1. Reagents
2. GAG analysis
3. LC-ESI-MS/MS methodology for determination of amino acids
4. UHPLC-IM-MS analysis
 - 4.1. Data acquisition
 - 4.2. Quality Control
 - 4.3. Data analysis and modelling
 - 4.4. Feature selection and annotation
 - 4.5. Pathway analysis

Tables

Table S1. MRM transitions for each amino acid and its corresponding internal standard.

Table S2. Instrumental settings for UHPLC-IM-MS analysis.

Table S3. Data for Venn diagram of the significantly pathways retrieved from untargeted, targeted approaches and *in-silico* systems biology.

Figures

Figure S1. Illustration of the analysis sequence.

Figure S2. PCA score plot showing the tight clustering of quality control (QC) samples before and after normalization.

Figure S3. PCA scores showing the effect of the support vector normalization step on the QC samples scores and the drift correction.

Figure S4. The RSD distribution bar plots before and after support vector normalization step.

Figure S5. OPLSDA model validation including the three groups: MPS I, MPSI treated and controls.

Below the figure, model parameters and CV-ANOVA results are presented.

Figure S6. OPLSDA model validation for MPS I vs. Control.

Figure S7. OPLSDA model validation for MPS I vs. Treated MPSI.

Figure S8. Boxplots of selected discriminant features in the three groups: MPS I, treated MPS I and Control samples.

Figure S9. Area under the receiver operating characteristic (ROC) curves, comparing diagnostic performance of different discriminant features retrieved from untargeted metabolomics to differentiate MPS I from Control samples.

Figure S10. Boxplots of amino acids concentration in the 3 groups: MPS I, treated MPS I and Control samples.

Figure S11. Area under the receiver operating characteristic (ROC) curves, comparing diagnostic performance of the most significant quantified amino acids to differentiate MPS I from Control samples.

Figure S12. Pathways of arginine metabolism.

Figure S13. Schematic representation of aminothiols synthesis and metabolism of glutathione.

1. Reagents and chemicals

Acetonitrile were purchased from VWR Chemicals (France), ultrapure water (18 MX) from Millipore (Molsheim, France) and formic acid from Fluka (Saint Quentin Fallavier, France). The chemicals used were of analytical grade. Leucine Enkephalin (Sigma–Aldrich) at a concentration of 2 ng/L (in acetonitrile/water, 50/50) was used as reference for mass measurements. Poly-DL-alanine and Dextran (Sigma–Aldrich) were prepared in 50:50 (v/v) water/acetonitrile 10 mg/L and used for ion mobility cell calibration. The aTRAQ Kit for Amino Acid Analysis of Physiological Fluids (Ref. 4442674) was purchased from Sciex (Life Science Holdings, France). HPLC gradient grade methanol was purchased from VWR Chemicals (Fontenay-sous-Bois, France).

2. GAG-derived disaccharides analysis (Adapted from [1])

Heparan sulfate, dermatan sulfate and keratin sulfate were enzymatically digested into disaccharides in a mixture containing 100 mM NH₄Ac (pH7.0), 10 mM Ca(Ac)₂, 2 mM DTT, 5 mIU each of heparinase I, II, III, 50 mIU chondroitinase B, 10 µl of KerII and 50 µL urine diluted to 2 mM creatinine in a final volume of 150 µL. After 2 h of incubation at 30°C, 15 µL of 150 mM EDTA (pH7.0) was added along with 125 ng of the internal standard, 4UA-2S-GlcNCOEt-6S (HD009, Iduron, Manchester, UK), and the reaction was stopped by boiling for 5 min to denature the proteins. The reaction mixture was centrifuged at 20,000×g for 5 min at room temperature. The supernatant was subsequently applied to an Amicon Ultra 30 K centrifugal filter (Millipore) and centrifuged at 14,000×g for 15 min at 25°C. The filtrate was stored at -20°C until analysis. The disaccharides were quantified on a Waters Quattro Premier XE (tandem) mass spectrometer (Waters Corporation, Milford, MA, USA) coupled to an Acquity UPLC system (UPLC-MS/MS) with a ESI source and a capillary voltage of 3.5 kV, a source temperature of 130°C, desolvation temperature of 350°C, cone gas flow of 50 L/hr, desolvation gas flow of 900 L/hr and a collision gas pressure of 2.5*10⁻³ mbar. The disaccharides were separated on a Thermo Hypercarb HPLC column (100×2.1 mm, 5 µm). The mobile phase consisted of 10 mM NH₄HCO₃ (pH10), and the disaccharides were eluted with an acetonitrile gradient of 0% to 20% for 2.5 min, held at 20% for the next 2.5 min, with 2 min of equilibration at 0% before the next injection; the flow rate was 0.2 mL/min, and the total run time was 7.1min. All disaccharides were detected and quantified in the MRM acquisition mode, using the transition m/z 378.1>175.1, cone voltage of 25 V and collision energy of 14 V for D0A0, m/z 416.1>138.0, cone voltage of 40 V and collision energy of 22 V for D0S0, m/z 458.1>97.0, cone voltage of 40 V and collision energy of 34 V for D0A6 and D2A0, m/z 496.0>416.0, cone voltage of 25 V and collision energy of 16 V for D2S0 and D0S6, m/z 458.0>299.9, cone voltage of 35 V and collision energy of 22 V for D0a4, m/z 538.0>458.0, cone voltage of 20 V and collision energy of 15 V for D0a10, m/z 462.1>361, cone voltage of 55 V and collision energy of 25 V for g0A6 and g6A6 (which could be separated by their retention time of 4.7 and 5.2 min, respectively) and m/z 472.0>97.0, cone voltage of 45 V and collision energy of 25 V for the 4UA-2S-GlcNCOEt-6S internal standard.

The identity of the peaks for the KS disaccharides was confirmed by analysis of samples spiked with ¹³C-labelled g0A6 (Glycosyn, New Zealand) and g6A6 and for HS and DS disaccharides with samples spiked with unlabeled disaccharide solutions. The concentration of the disaccharides was calculated using a calibration curve of each disaccharide with 4UA-2S-GlcNCOEt-6S (HD009,

Iduron) as an internal standard. The following disaccharides were analyzed and used to calculate GAG concentrations: D0A0, D0S0, D0A6, D2A0, D0S6 and D2S0 (HS), D0a4 and D0a10 (DS), g0A6 and G6A6 (KS).

3. LC-ESI-MS/MS methodology for determination of amino acids

The aTRAQ Kit for Amino Acid Analysis of Physiological Fluids was purchased from Sciex (Framingham, MA, USA). It consisted of amine-modifying labeling aTRAQ reagent $\Delta 8$, aTRAQ internal standard set of amino acids labeled with the aTRAQ reagent $\Delta 0$, 10 % sulfosalicylic acid, borate buffer of pH 8.5, 1.2 % hydroxylamine and mobile phase modifiers – formic acid and heptafluorobutyric acid. HPLC gradient grade methanol was purchased from J.T. Baker (Center Valley, PA, USA). Deionized water obtained from Millipore Simplicity UV water purification system (Waters Corporation, Milford, MA, USA) was used. Aminoacids standards were purchased from SIGMA-ALDRICH (acidic and neutral aminoacids Ref: A6407-5ML. Basic aminoacids Ref: A6282-5ML).

The following protocol was used for preparation of urine samples. The urine were first diluted using 50/50 v/v using deionized water. An aliquot of 40 μ l of the sample was added to 10 μ l of 10% sulfosalicylic acid in order to precipitate proteins. After mixing and centrifugation (10 000 \times g for 2 min) the supernatant was mixed with 40 μ l of borate buffer. Next an aliquot of 10 μ l the obtained solution was labeled with aTRAQ reagent solution (aTRAQ reagent $\Delta 8$), mixed and centrifuged. After 30 min of incubation at room temperature the labeling reaction was stopped by addition of 5 μ l 1.2% hydroxylamine solution and the sample was incubated at room temperature for 15 min. In the next step 32 μ l of the internal standard solution was added to the sample. After mixing and centrifugation the sample was evaporated in a vacuum concentrator for 15 min in order to reduce volume to about 20 μ l. Then the residue was diluted with 20 μ l of water. Each determined amino acid had its corresponding internal standard (the same amino acid labeled with the aTRAQ reagent $\Delta 0$). Two non-proteinogenic amino acids (norleucine and norvaline) were used to evaluate the labeling efficiency and recovery. A calibration curve was constructed using five or concentrations derived from the peak area ratio of each aminoacid and the internal standard. A calibrations curve points has been used are as follows:

Aminoacids (μM)	5	25	100	250	500	1200
Glutamine (μM)	10	50	200	500	1000	2500
Cystine (μM)	5	25	100	250	500	

Concentrations were calculated from these area ratios using the calibration curve established by simple regression. The determination of free amino acid levels was conducted using the liquid chromatography instrument Prominence Shimadzu UFLC system consisting of a DGU-20A3

degasser, a LC-20AB pump, a SIL-20AHT autosampler, and a CTO-20AC oven (Shimadzu, Prominence, Kyoto, Japan) coupled to the 4000 QTRAP mass spectrometer (Sciex, Framingham, MA, USA) with an electrospray ion source. The chromatographic separation was achieved on Sciex C18 column (5 μ m, 4.6 mm x 150 mm) maintained at 50 °C with a flow rate of 800 μ L/min. A mobile phase gradient of eluent A (0.1% formic acid and 0.01% heptafluorobutyric acid in water) and eluent B (0.1% formic acid and 0.01% heptafluorobutyric acid in methanol) was applied. A gradient profile was the following: from 2% to 40% of B from 0 till 6 min, maintained at 40% of B for 4 min, then increased to 90% of B till 11 min and held at 90% of B for 1 min. After 12 min the gradient decreased to 2% of B. From 13 to 18 min the mobile phase composition was unaltered. The injection volume was set at 2 μ l. The ion source settings were: curtain gas, 20 psig; ion spray voltage, 4500 V; source temperature, 600 °C; ion source gas 1, 60 psig and ion source gas 2, 50 psig. The mass spectrometer operated in positive ionization mode with the following parameters: entrance potential, 10 V; declustering potential, 30 V and collision cell exit potential, 5 V. Collision energy of 30 eV was applied. The list of measured MRM transitions are presented in **Table S1**. Scheduled multiple reaction monitoring mode was used with nitrogen as a collision gas. A system suitability test was conducted before each batch of the samples (analysis of a standard mixture) to warm up the LC-MS/MS system and check the inter-day performance of the system. Data acquisition and processing were performed using the Analyst 1.5 software (Sciex, Framingham, MA, USA).

Table S1. MRM transitions for each amino acid and its corresponding internal standard.

Amino Acid	Analyte		Internal standard	
	Q1	Q2	Q1	Q2
Alanine	238.2	121.1	230.2	113.1
Arginine	323.2	121.1	315.2	113.1
Asparagine	281.2	121.1	273.2	113.1
Aspartic acid	282.1	121.1	274.1	113.1
Citruline	324.2	121.1	316.2	113.1
Cysteine	537.2	121.1	521.2	113.1
Glutamate	296.2	121.1	288.2	113.1
Glutamine	295.2	121.1	287.2	113.1
Glycine	224.1	121.1	216.1	113.1
Histidine	304.2	121.1	296.2	113.1
Isoleucine	280.2	121.1	272.2	113.1
Leucine	280.2	121.1	272.2	113.1
Lysine	443.3	121.1	427.3	113.1
Methionine	298.2	121.1	290.2	113.1
Ornithine	429.3	121.1	413.3	113.1
Phenylalanine	314.2	121.1	306.2	113.1
Proline	264.2	121.1	256.2	113.1
Serine	254.2	121.1	246.2	113.1
Taurine	274.1	121.1	266.1	113.1
Threonine	268.2	121.1	260.2	113.1
Tryptophane	353.2	121.1	345.2	113.1
Tyrosine	330.2	121.1	322.2	113.1
Valine	266.2	121.1	258.2	113.1

4. UHPLC-IM-MS analysis

2.1.1.1.1. Sample handling

The sample handling component was a Waters 2777C sample manager (Waters Corp., Milford, MA, USA) equipped with a 25 µl Hamilton syringe, a 2 µl loop used for full loop injections of prepared

sample, and a 2-drawer sample chamber thermo-stated at 4°C with a constant flow of dry nitrogen gas to prevent the buildup of condensation.

2.1.1.1.2. Chromatographic conditions

The chromatography was performed on a Waters NanoAcquity UPLC module (Saint Quentin en Yvelines, France) upgraded to work with 1 mm columns and composed with a binary solvent manager and column heater/cooler module. Separation was carried out at 40 °C using a 1.0 x 100 mm, 1.8 µm Acquity UPLC HSS T3 column (Waters), with a particle size of 1.8 µm, equipped with a 0.2 µm prefilter. Urine was eluted from the LC column using the following linear gradient (curve number 6): 0–1 min: 99% A; 1–3 min, 99–85% A; 3–6 min, 85-50% A; 6–9 min, 50-0% A; 9–12 min, 100% 12–16 min, 99% A for re-equilibration. Solvent A was water and solvent B was acetonitrile, both solvents contained 0.1% formic acid. The duration of column equilibration was adjusted to provide sufficient retention and chromatographic precision of early eluting species in subsequent analyses at the minimal expense of time. Sample analysis order has been randomized to avoid potential for confounding critical variables with analytical run order effects. Peak splitting and column overload was avoided by using small injection volumes (2µL) for LC-IM-MS analysis. The instrumental parameters have been optimized following the strategy previously described [2].

2.1.1.1.3. Mass spectrometry

The U-HPLC system was coupled to a hybrid quadrupole orthogonal time-of-flight (TOF) mass spectrometer (SYNAPT G2 HDMS, Waters MS Technologies, Manchester, UK). The mass spectrometer was operated in both negative and positive electrospray ionization mode. A mass range of m/z 50–1200 was used in both modes. The sample cone voltage, extraction cone voltage, source temperature, desolvation temperature, desolvation gas flow and cone gas flow were optimized as previously described [2]. Leucine enkephalin was used as the lock mass $[M-H]^+$ at m/z 556.2771. Sodium formate solution was used for external instrument calibration.

2.1.1.1.4. Ion mobility

Synapt G2 HDMS (Waters MS Technologies, Manchester, UK) was used in our study for Ion Mobility. It is equipped with a travelling wave “Triwave™” geometry in which the ion mobility cell (IMS T-wave) is placed between two traveling wave ion guides (trap T-wave and transfer T-wave). After ionization in the source and transfer through the quadrupole, the ions arrive at the first

travelling-wave ion guide that acts as an ion trap, namely “trap TWIG”. In this region, the ions are accumulated before being released in packets and accelerated using the trap-bias voltage to the second cell “IMS-TWIG” for mobility separation. In the IMS-TWIG a travelling wave is continuously applied at a given wave height and velocity to enhance separation through the mobility cell, which is filled with a gas. In this study, the IMS drift gas flow (nitrogen) and the wave velocity settings were assessed and optimized. The helium cell gas flow, wave height, Trap Bias and IMS wave delay were set at 180 mL/min, 40 V, 45 V and 450 μ s respectively. The TOF analyzer was operated in the V resolution mode with an average mass resolution of $m/\Delta m$ 20,000 (full-width at half-maximum definition). Data acquisition of an ion mobility experiment consisted of 200 bins. CCS values, obtained in nitrogen, were experimentally determined using singly charged Poly-DL-alanine oligomers as the TWIM calibrant species for ESI+. CCS values were derived according to previously reported procedures [3]. The ion mobility resolution was $\sim 40 \Omega/\Delta\Omega$ (fwhm). The he CCS values reported were determined at the apex of the ion-mobility peak. Detailed instrument settings are presented in Table S-1.

2.1.1.1.5. Raw data preprocessing

All LC-IM-MS raw data files data processing, peak detection and peak matching across samples using retention time (tR) correction and chromatographic alignment along with drift time and CCS calculation were performed using Progenesis QI (Waters MS Technologies, Manchester, UK) to yield a data matrix containing retention times, accurate masses, CCS and peak intensities. The preprocessing step resulted in an X -matrix where tR, CCS and m/z values were concatenated into “tR_m/z_CCS” features (in columns) present in each sample (in rows) with corresponding peak areas.

2.1.1.1.6. Quality Control

Aliquoted 10 μ L of each urine sample are mixed together to generate a pooled quality control sample (QCs). QCs and solvent blank samples (mobile phase) were injected sequentially in-between the urine samples. In addition, a dilution series of QC samples (6%, 12.5%, 25%, 50% and 100% original concentration) are used to assess the quality of the extracted features. An analysis sequence is presented in Figure. S-1. Indeed, feature extraction algorithms including automatic peak detection, grouping, and integration often yield a data matrix containing analytical system noise such as mobile phase chemical contaminants signals. Depending on the software and the used parameters for feature extraction, such noise can represent a significant portion of the total number of detected features. Therefore, this may mislead further data analysis such as transformation, normalization

and data modeling. Simple noise filtering strategies such as the minimum fraction filter may remove infrequently observed signals within the considered distinct sample classes. In this study, we used a filter strategy in which the features intensity must be correlated to the matrix concentration in a series of diluted QC samples in order to be included in further analysis. Beyond its role as a system noise marker, the dilution series filter is very useful to assess the informative quality of the extracted features. It ensures that the observed signal of a given feature and its relative concentration in the sample are positively correlated. This approach is used to identify feature groups that are not correlated to the gradient of concentration generated by the dilutions series and therefore should not be considered as reliable features. Thus, feature groups with correlation coefficient of less than 0.7 were removed from the dataset. Furthermore, datasets are refined by removal of feature groups that do not meet threshold of peak area measurement precision prior to data analysis. This approach uses RSD values derived from repeated measurements of a pooled QC sample. The threshold was set to $RSD < 25\%$. Thus enhancing the biological interpretation of metabolomics data. The system stability assessment has been done using the Principal Component Analysis and assessing the clustering of the QC samples. Figures are shown supporting information (Figures S-2) presenting the PCA score plot derived from the metabolomics analysis of all the urine samples and QCs replicates. In the PCA score plot, each point corresponds to a different sample or QC. The tightness of the clustering reveals the similarity of the samples (QCs), and thus the systems stability. The QCs are tightly clustered indicating a good instrumental stability over the metabolomics analysis.

Table S2. Instrumental settings for UHPLC-IM-MS analysis.

Chromatography		Column Temperature	40 (°C)
		Flow rate	80 (mL/min)
		Injection volume	2 µL
ESI-MS		Capillary voltage	2 V
		Sampling cone voltage	30 V
		Extraction cone voltage	5
		Source temperature	120 (°C)
		Desolvation temperature	500 (°C)
		Desolvation gas flow	400 (liters/h)
		Cone gas flow	50
		Optic mode	Resolution
		MS scan rate	0.2 scan/s
		Lock mass solution	Leu-enk (2 µg/ml)
		Lock mass flow rate	6 µL/min
Triwave DC	Trap DC	Entrance	3
		Bias	45
		Trap DC	0
		Exit	3
	IMS DC	Entrance	25
		Helium cell DC	35
		Helium exit	-5
		Bias	3
		Exit	0
	Transfer DC	Entrance	4
		Exit	15
Gas controls	IMS gas	Nitrogen	90 (ml/min)
	Helium cell	Nitrogen	180 (ml/min)
Triwave	Trap	Wave velocity	300 (m/s)
		Wave height	0.5 (V)
	IMS	Wave velocity	587 (m/s)
		Wave height	40 (V)

Transfer	Wave velocity	300 (m/s)
	Wave height	5 (V)

5. Data analysis and modeling

Support vector regression normalization method was applied using the MetNormalizer R package [4] before any data analysis, to remove the unwanted intra- and inter-batch measurements analytical variations. The effect of this normalization step on the raw data is shown in Figures S-2, S-3 and S-4 (Supporting information). Then normalized data matrix has been log-transformed and pareto-scaled. All data analyzes and modeling were done using SIMCA 14.0 (MKS DAS, Umeå, Sweden). First, hierarchical cluster analysis has been applied to the data set to get an overview the clustering trends of samples with similar profiles of variable intensity. Furthermore, multivariate data analysis and modeling is performed using Principal Component Analysis (PCA) as unsupervised method PCA was first applied to get an overview of the data and identify potential severe outliers which are defined as observations whose scores mapped outside the Hotelling's T2 ellipse (confidence interval = 0.95) in a cross-validated seven-component model. The DmodX was used to detect moderate outliers [5]. Orthogonal Partial Least-Squares-Discriminant Analysis (OPLSDA) as a supervised method. To select the most relevant features, the training group has been repeatedly split into a training set and a test set. A permutation test (999 iterations) was performed to prevent the OPLSDA over fitting of the model by comparing diagnostic statistic metrics of the generated model with those of randomly generated models. R2X is the cumulative modeled variation in X (X = features), R2Y is the cumulative modeled variation in Y (Y = sample groups), and Q2Y is the cumulative predicted variation in Y, based on the cross-validation. The range of these parameters is between 0 and 1, where 1 indicates a perfect fit. Furthermore, cross-validated analysis of variance (CV-ANOVA) was systematically performed based on the cross-validated model [6]. The Figure S-2 shows a PCA score plot of the raw data compared to normalized, log-transformed and pareto-scaled data which emphasis the importance of these data pretreatment steps before data modeling. The X matrix was 100 x 854 variables. The Y matrix was 100 analyses x 3 groups. Characteristics and validation results from all OPLS-DA models are provided in supplemental material (Figures S-5, S-6 and S-7).

6. Feature selection and annotation

To select the most discriminant variables for the separation of groups, S-Plot was used. The S-plot combines the covariances and correlations between the **X** matrix and O-PLS scores for a given model component. The covariance values give the magnitude of contribution of a variable while the correlation values reflect the effect and reliability of the variable for the model component scores. Variables with both very high correlation and covariance are important for the explanation power of the model. Furthermore, selection of discriminant variables was achieved using the VIP scores procedure for each validated OPLS-DA model [7]. Putative annotation of detected features was performed using both accurate mass comparison using freely available metabolite databases HMDB, LipidBlast, KEGG, and Metlin. Furthermore, CCS values were also compared to the MetCCS database [8].

7. Pathway and network analysis

In order to provide a broader understanding of metabolic changes in MPS I, we also explored the biochemical pathways using a network analysis approach using Mummichog (v.1.0.5) which allows pathway enrichment analyses. The idea behind this metabolic network prediction strategy assumes that metabolite concentration alterations are more likely to occur within a metabolic connected network rather than in a random fashion. This Mummichog python package highlight pathways that are significantly impacted in the studied groups. Significantly impacted biochemical pathways are those exhibiting an adjusted p-value <0.05. For this comparison, we focused on features that significantly changed (511 features with q-values = 0.05 and FDR = 5%). Mummichog annotates metabolites based on accurate mass m/z (5 ppm mass error was used) and tests significant pathway enrichment within a reference metabolic network using a Fisher's exact test [9]. The matched candidates were then mapped to reference human metabolic networks from the KEGG, MetaCyc, Recon and Edinburgh Human Metabolic Network. The null distribution in pathway analysis was obtained from 1000 set of randomly permuted m/z lists draw from all features detected in the whole metabolomic dataset and modeled by Gamma distribution. To protect against incorrect pathway selection, redundant pathways or those enriched by less than two metabolites were excluded. MetaboAnalyst [10] has been used for Metabolite Set Enrichment Analysis using the amino acids concentrations.

RÉFÉRENCES

1. Langereis, E.J.; Wagemans, T.; Kulik, W.; Lefeber, D.J.; van Lenthe, H.; Oussoren, E.; van der Ploeg, A.T.; Ruijter, G.J.; Wevers, R.A.; Wijburg, F.A., *et al.* A multiplex assay for the diagnosis of mucopolysaccharidoses and mucopolipidoses. *PLoS one* **2015**, *10*, e0138622.
2. Tebani, A.; Schmitz-Afonso, I.; Rutledge, D.N.; Gonzalez, B.J.; Bekri, S.; Afonso, C. Optimization of a liquid chromatography ion mobility-mass spectrometry method for untargeted metabolomics using experimental design and multivariate data analysis. *Analytica Chimica Acta* **2016**, *913*, 55-62.
3. Paglia, G.; Angel, P.; Williams, J.P.; Richardson, K.; Olivos, H.J.; Thompson, J.W.; Menikarachchi, L.; Lai, S.; Walsh, C.; Moseley, A., *et al.* Ion mobility-derived collision cross section as an additional measure for lipid fingerprinting and identification. *Analytical chemistry* **2015**, *87*, 1137-1144.
4. Shen, X.; Gong, X.; Cai, Y.; Guo, Y.; Tu, J.; Li, H.; Zhang, T.; Wang, J.; Xue, F.; Zhu, Z.-J. Normalization and integration of large-scale metabolomics data using support vector regression. *Metabolomics* **2016**, *12*, 89.
5. Eriksson, L.; Trygg, J.; Wold, S. A chemometrics toolbox based on projections and latent variables. *Journal of Chemometrics* **2014**, *28*, 332-346.
6. Eriksson, L.; Trygg, J.; Wold, S. Cv-anova for significance testing of pls and oppls® models. *Journal of Chemometrics* **2008**, *22*, 594-600.
7. Galindo-Prieto, B.; Eriksson, L.; Trygg, J. Variable influence on projection (vip) for orthogonal projections to latent structures (oppls). *Journal of Chemometrics* **2014**.
8. Zhou, Z.; Xiong, X.; Zhu, Z.J. Metccs predictor: A web server for predicting collision cross-section values of metabolites in ion mobility-mass spectrometry based metabolomics. *Bioinformatics (Oxford, England)* **2017**.
9. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F.H.; Khan, N.; Soltow, Q.A.; Jones, D.P.; Pulendran, B. Predicting network activity from high throughput metabolomics. *PLoS Comput. Biol.* **2013**, *9*, e1003123.
10. Xia, J.; Sinelnikov, I.V.; Han, B.; Wishart, D.S. Metaboanalyst 3.0-making metabolomics more meaningful. *Nucleic acids research* **2015**.
1. Salazar, D.A.; Rodriguez-Lopez, A.; Herreno, A.; Barbosa, H.; Herrera, J.; Ardila, A.; Barreto, G.E.; Gonzalez, J.; Almeciga-Diaz, C.J. Systems biology study of mucopolysaccharidosis using a human metabolic reconstruction network. *Molecular genetics and metabolism* **2016**, *117*, 129-139

Table S3. The normalized concentrations of free amino acids in urine samples of the three studied groups.

(μ M amino acid / mM creatinine)

	Control group (n=66)			MPS I group (n = 19)			MPS IT group (n = 15)		
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
L-Alanine	30.4	36.9	0.3 - 205.5	63.0	58.3	5.9 - 244.5	26.3	34.9	0.8 - 122.5
L-Arginine	27.1	147.1	0 - 1012.5	12.3	12.5	1.6 - 50.7	9.4	7.4	0 - 28.3
L-Citrulline	2.0	2.7	0 - 17.1	4.2	4.0	0 - 12.9	2.0	2.6	0 - 10
L-Cystine	12.5	34.2	0 - 213.1	9.2	15.2	0.1 - 65.6	4.2	4.8	0.3 - 15.8
L-Glutamine	51.4	56.0	1 - 327.7	103.6	114.6	2.7 - 528.1	48.2	51.5	2.4 - 184.2
L-Glutamic acid	3.9	4.8	0.7 - 30.2	10.0	7.1	1.5 - 25.8	5.0	5.1	1 - 20.8
Glycine	145.3	169.2	1.6 - 967.3	265.1	315.1	37 - 1322.7	157.3	134.4	5.5 - 436.7
L-Histidine	78.2	89.0	0.7 - 426.2	133.5	176.1	7.6 - 810.9	72.7	68.3	2.7 - 267.5
L-isoIeucine	1.9	2.1	0 - 11.8	4.3	4.0	0.1 - 12.5	0.9	1.3	0 - 5
L-Leucine	3.8	4.6	0.1 - 25.7	7.1	7.9	1.3 - 35.9	3.3	4.2	0.4 - 17.5
L-Lysine	55.1	170.4	0.8 - 1129.8	24.8	18.6	4.2 - 66.8	15.9	23.0	1.6 - 82.5
L-Methionine	0.8	1.1	0 - 6.1	1.1	1.9	0.1 - 8.6	0.3	0.3	0 - 1
L-Ornithine	15.5	64.0	0 - 451.7	6.1	7.1	0.1 - 27.1	7.6	12.1	0 - 42.5
L-Phenylalanine	6.0	5.9	0 - 29.1	13.0	17.9	0.1 - 82.8	4.4	5.5	0 - 22.5

L-Proline	2.4	4.2	0.3 - 29.3	7.0	9.7	0.4 - 44.6	3.0	3.2	0.5 - 10.6
L-Serine	44.6	50.4	0 - 252.3	84.4	99.6	0.1 - 379.7	41.5	48.2	0 - 174.2
L-Tyrosine	9.8	11.1	0 - 64.9	19.9	26.7	0.1 - 117.2	6.3	7.8	0 - 31.7
L-Valine	5.0	5.5	0 - 26.4	10.8	10.1	1.7 - 42.2	4.4	5.0	0.8 - 20
L-Aspartic acid	1.3	2.1	0.1 - 8.8	2.8	1.7	0.4 - 6.8	1.9	4.9	0 - 20
Taurine	47.4	54.0	0.6 - 334.5	148.4	379.6	0.1 - 1704.7	111.4	200.5	1.6 - 821.7
L-Threonine	13.4	15.9	0.8 - 71	27.4	34.7	3 - 136.2	15.4	20.4	1 - 70.8
L-Asparagine	12.3	15.3	0 - 89.6	22.3	20.9	0.1 - 67.7	9.1	11.9	0 - 48.7
Cystathionine	3.7	3.7	0 - 16.8	11.4	12.5	0.1 - 48.4	3.2	3.0	0 - 9.2
L-Tryptophane	6.4	5.9	0 - 32.8	13.2	20.5	0.6 - 94.5	5.2	5.7	0.2 - 23.3

Table S4. Data for venn diagram of the significantly pathways retrieved from untargeted, targeted approaches and *in silico* systems biology approach from Salazar DA et al [11].

Data	N°	Pathways
<i>In silico</i> Targeted Untargeted	2	Glutathione Metabolism
		Arginine and Proline Metabolism
<i>In silico</i> Targeted	4	Glycerophospholipid Metabolism
		Tyrosine Metabolism
		Methionine and Cysteine Metabolism
		Lysine Metabolism
<i>In silico</i> Untargeted	4	Histidine Metabolism
		Taurine and Hypotaurine Metabolism
		Propanoate Metabolism
		Butanoate Metabolism
Untargeted	5	Ascorbate (Vitamin C) And Aldarate Metabolism
		Vitamin H (Biotin) Metabolism
		Biopterin Metabolism
		Urea Cycle/Amino Group Metabolism
Targeted	18	Lysine Biosynthesis Glycine, Serine And Threonine Metabolism Aminoacyl-trna Biosynthesis Valine, Leucine And Isoleucine Degradation Nicotinate And Nicotinamide Metabolism Nitrogen Metabolism Cyanoamino Acid Metabolism Selenoamino Acid Metabolism Beta-alanine Metabolism Porphyrin And Chlorophyll Metabolism Valine, Leucine And Isoleucine Biosynthesis Cysteine And Methionine Metabolism D-arginine And D-ornithine Metabolism Pyrimidine Metabolism Purine Metabolism Pantothenate And Coa Biosynthesis D-glutamine And D-glutamate Metabolism Alanine, Aspartate And Glutamate Metabolism

In silico

68 Squalene And Cholesterol Synthesis Sphingolipid Metabolism Fatty Acid Oxidation Purine Synthesis Phosphatidylinositol Phosphate Metabolism Glycosphingolipid Metabolism Vitamin B6 Metabolism Vitamin A Metabolism Tryptophan Metabolism Glycolysis/Gluconeogenesis Transport, Lysosomal N-glycan Degradation Exchange/Demand Reaction Phenylalanine Metabolism Thiamine Metabolism Nucleotide Interconversion Dietary Fiber Binding Coa Synthesis Transport, Endoplasmic Reticular Vitamin C Metabolism Cholesterol Metabolism Nad Metabolism Pentose Phosphate Pathway R Group Synthesis Pyruvate Metabolism Keratan Sulfate Degradation Pyrimidine Catabolism Transport, Nuclear Fructose And Mannose Metabolism Tetrahydrobiopterin Metabolism Aminosugar Metabolism Pyrimidine Synthesis Galactose Metabolism Folate Metabolism Transport, Extracellular Oxidative Phosphorylation Glyoxylate And Dicarboxylate Metabolism Glutamate Metabolism Transport, Golgi Apparatus Ros Detoxification Transport, Peroxisomal Heme Synthesis Citric Acid Cycle Inositol Phosphate Metabolism Unassigned Keratan Sulfate Synthesis Purine Catabolism Transport, Mitochondrial Bile Acid Synthesis Biotin Metabolism N-glycan Synthesis Valine, Leucine, And Isoleucine Metabolism Vitamin B2 Metabolism Urea Cycle Triacylglycerol Synthesis Nucleotide Salvage Pathway Fatty Acid Synthesis Steroid Metabolism Miscellaneous Androgen And Estrogen Synthesis And Metabolism Alanine And Aspartate Metabolism Eicosanoid Metabolism Cysteine Metabolism C5-branched Dibasic Acid Metabolism Coa Catabolism Starch And Sucrose Metabolism Glycine, Serine, Alanine And Threonine Metabolism D-alanine Metabolism

Figure S1. Illustration of the analysis sequence. Aliquoted 10 μL of each urine sample are mixed together to generate a pooled quality control sample (QCs). QCs and solvent blank samples (mobile phase) were injected sequentially in-between the urine samples. In addition, a dilution series of QC samples (6%, 12.5%, 25%, 50% and 100% of the original concentration) are used to assess the quality of the extracted features. A conditioning step is used to condition the column using ten QC injections. Sample injection order has been orthogonalized.

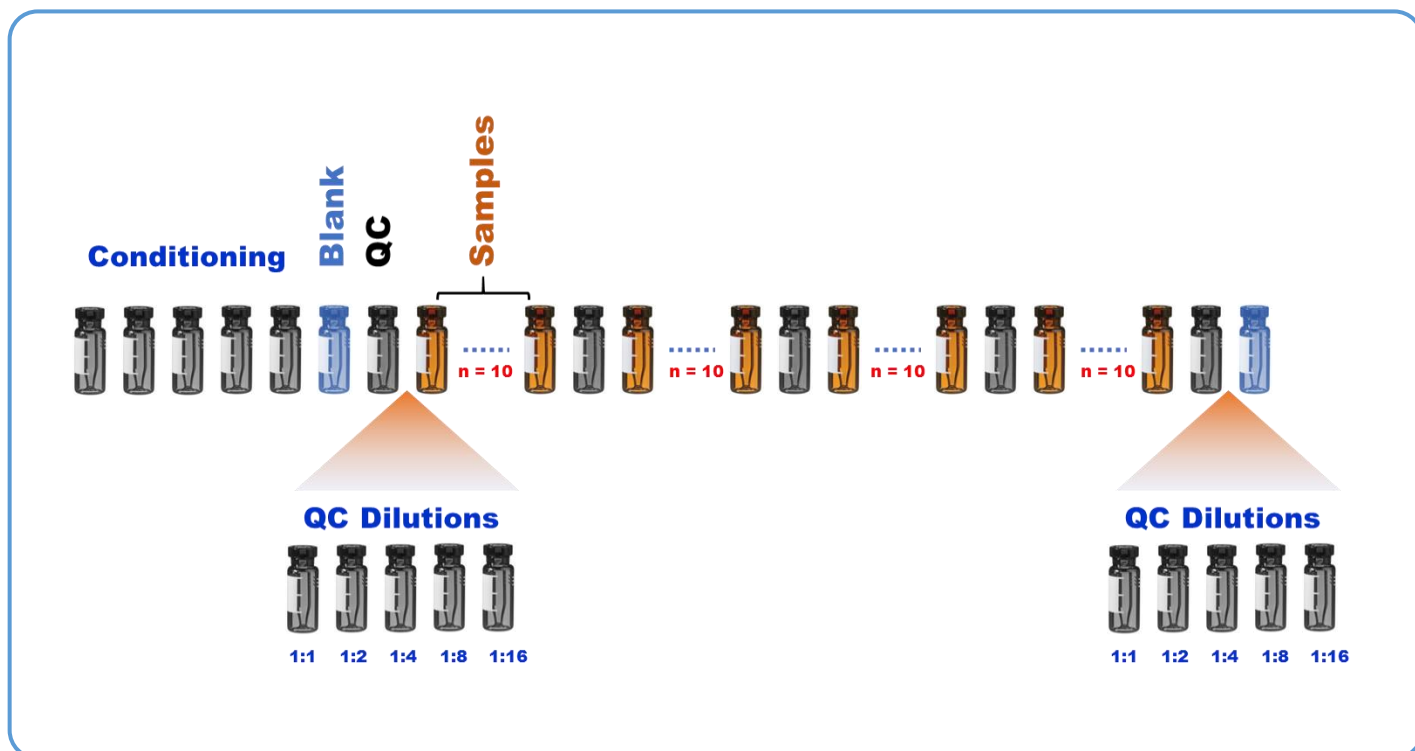


Figure S2. PCA score plot showing the tight clustering of quality control (QC) samples before and after normalization. Up: positive mode. Below: negative mode. Grey circle: samples. Green circle: QC samples. Left: raw data. Right: normalized data.

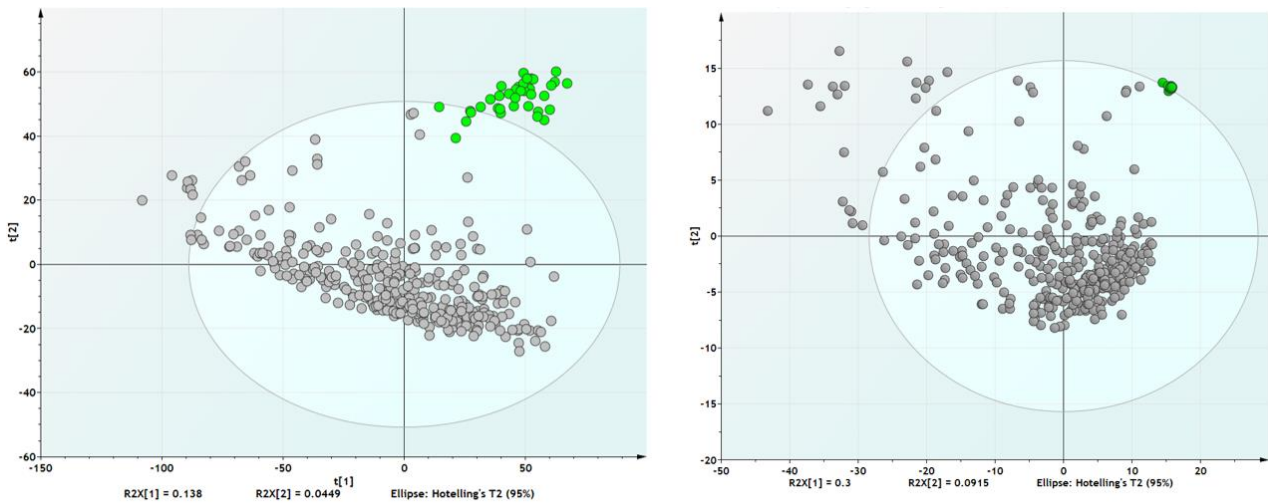


Figure S3. PCA scores showing the effect of the support vector normalization step on the QC samples scores and the drift correction. Blue: first component. Red: second component. Green: third component

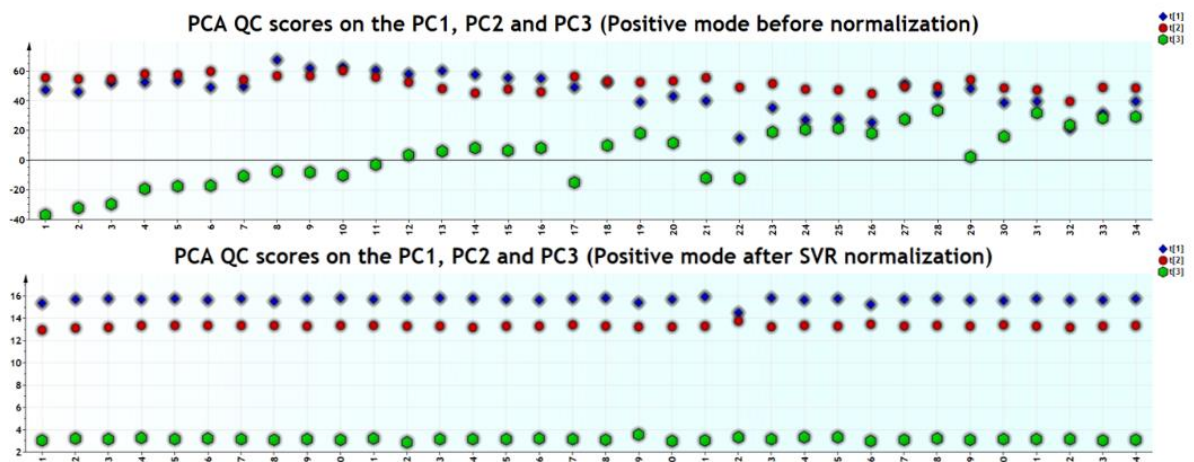


Figure S4. The RSD distribution bar plots before and after support vector normalization step.

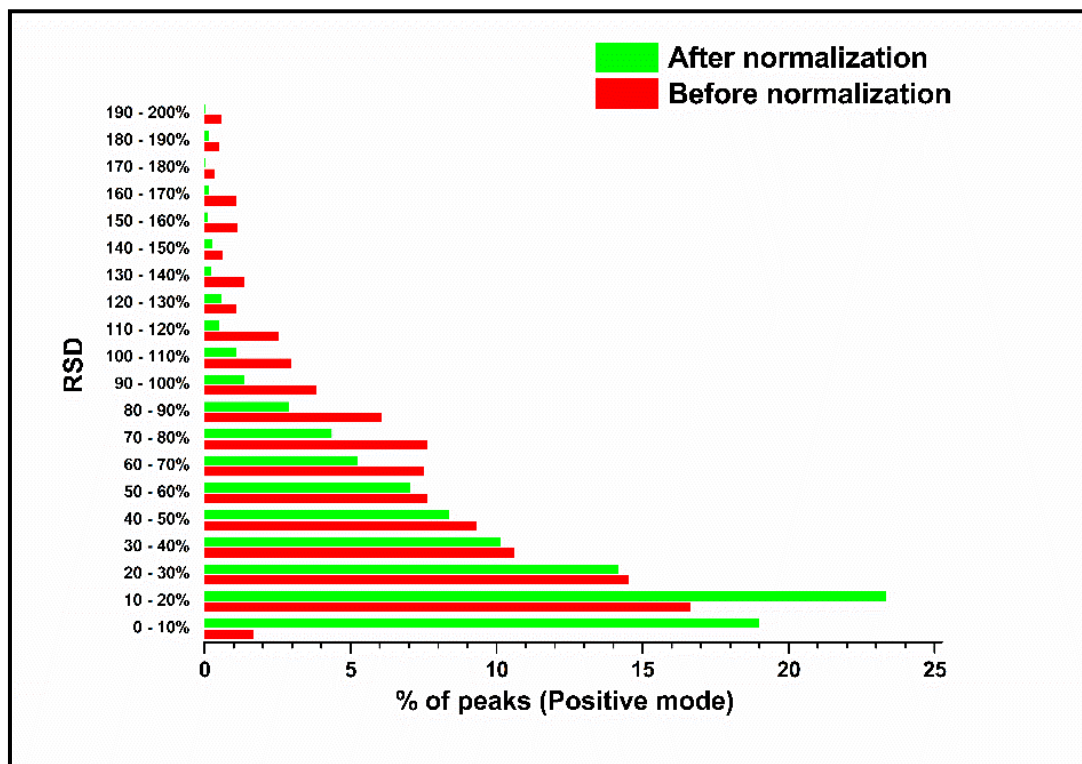
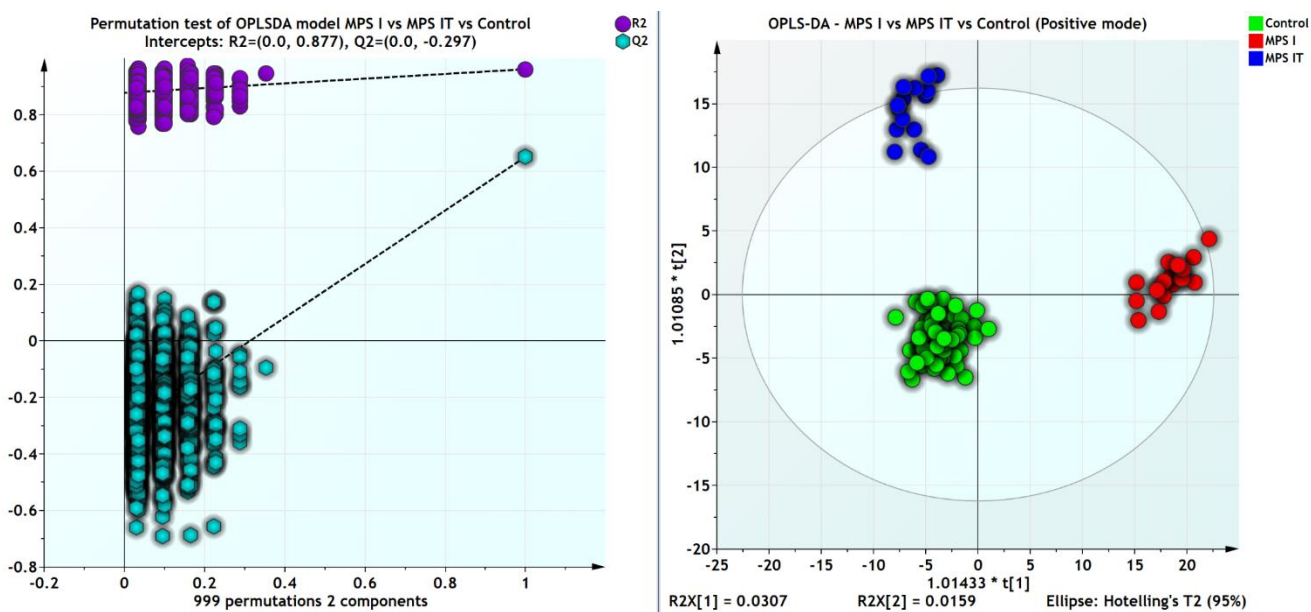


Figure S5. OPLSDA model validation including the three groups: MPS I, MPSIT and controls.

Below the figure, model parameters and CV-ANOVA results are presented.



MPS I vs. MPSIT vs. Control - OPLSDA model parameters.

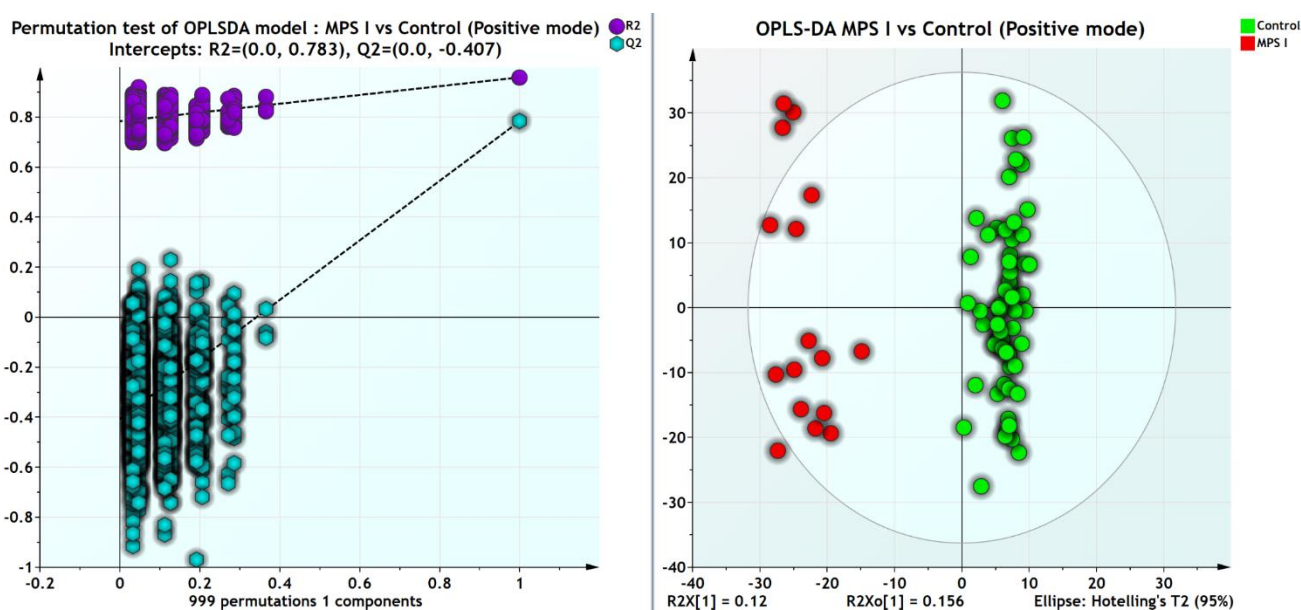
Component	R2X	R2X(cum)	Eigenvalue	R2	R2(cum)	Q2	Limit	Q2(cum)	R2Y	R2Y(cum)
Model		0.328			0.961			0.542		1
Predictive		0.0456			0.961			0.542		1
P1	0.0298	0.0298	3.04	0.51	0.51	0.296	0.01	0.296	0.52	0.52
P2	0.0158	0.0456	1.61	0.451	0.961	0.247	0.01	0.542	0.48	1
Orthogonal in X(OPLS)		0.283			0					
O1	0.198	0.198	20.2	0	0					
O2	0.0258	0.224	2.64	0	0					
O3	0.0209	0.245	2.13	0	0					
O4	0.0221	0.267	2.25	0	0					
O5	0.0158	0.283	1.61	0	0					

CV-ANOVA test results.

	SS	DF	MS	F	p	SD
Total corr.	202	202	1			1
Regression	110.774	24	4.615	9.005	3.99e-020	2.148
Residual	91.226	178	0.512			0.715

Figure S6. OPLSDA model validation for MPS I vs. Control.

Below the figure, model parameters and CV-ANOVA results are presented.



MPS I vs. Control - OPLSDA model parameters.

Component	R2X	R2X(cum)	Eigenvalue	R2	R2(cum)	Q2	Limit	Q2(cum)	R2Y	R2Y(cum)
Model		0.271			0.945			0.631		1
Predictive		0.0329			0.945			0.631		1
P1	0.0329	0.0329	2.86	0.945	0.945	0.631	0.01	0.631	1	1

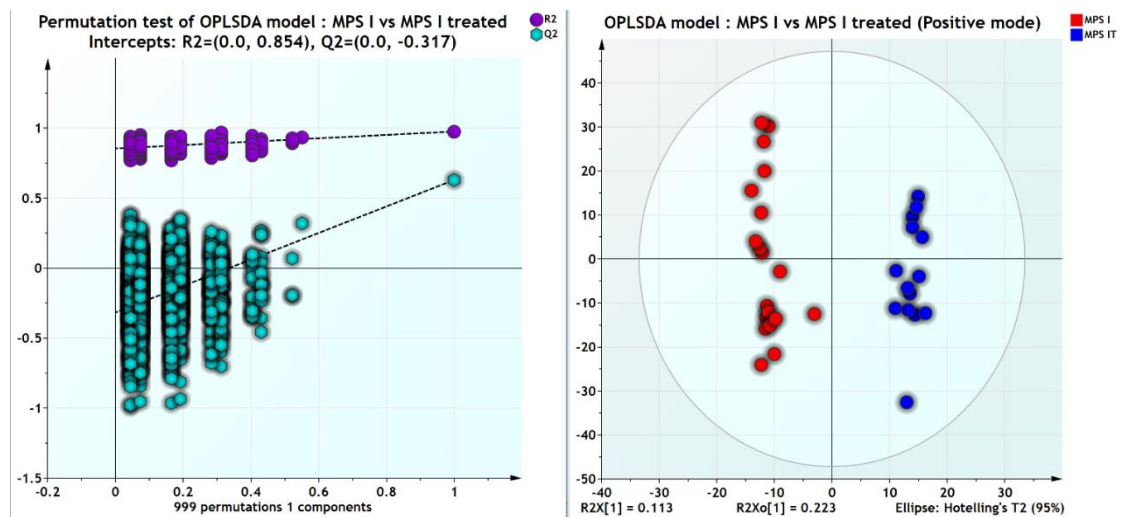
Orthogonal in X(OPLS)		0.239			0	
O1	0.21	0.21	18.3	0	0	
O2	0.0282	0.239	2.45	0	0	

MPS I vs. Control - CV-ANOVA test results.

	SS	DF	MS	F	p	SD
Total corr.	86	86	1			1
Regression	54.235	6	9.039	22.766	1.75e-015	3.006
Residual	31.764	80	0.397			0.630

Figure S7. OPLSDA model validation for MPS I vs. MPSIT.

Below the figure, model parameters and CV-ANOVA results are presented.



MPS I vs. MPSIT - CV-ANOVA test results.

	SS	DF	MS	F	p	SD
Total corr.	33	33	1			1
Regression	20.785	8	2.598	5.3179	0.000580	1.611
Residual	12.214	25	0.488			0.698

MPS I vs. MPSIT - OPLSDA model parameters.

Component	R2X	R2X(cum)	Eigenvalue	R2	R2(cum)	Q2	Limit	Q2(cum)	R2Y	R2Y(cum)
Model		0.488			0.975			0.63		1
Predictive		0.113			0.975			0.63		1
P1	0.113	0.113	3.83	0.975	0.975	0.63	0.01	0.63	1	1
Orthogonal in X(OPLS)		0.376			0					
O1	0.223	0.223	7.59	0	0					
O2	0.115	0.338	3.9	0	0					
O3	0.0375	0.376	1.28	0	0					

Figure S8. Boxplots of selected discriminant features in the three groups: MPS I, MPS IT and Control samples.

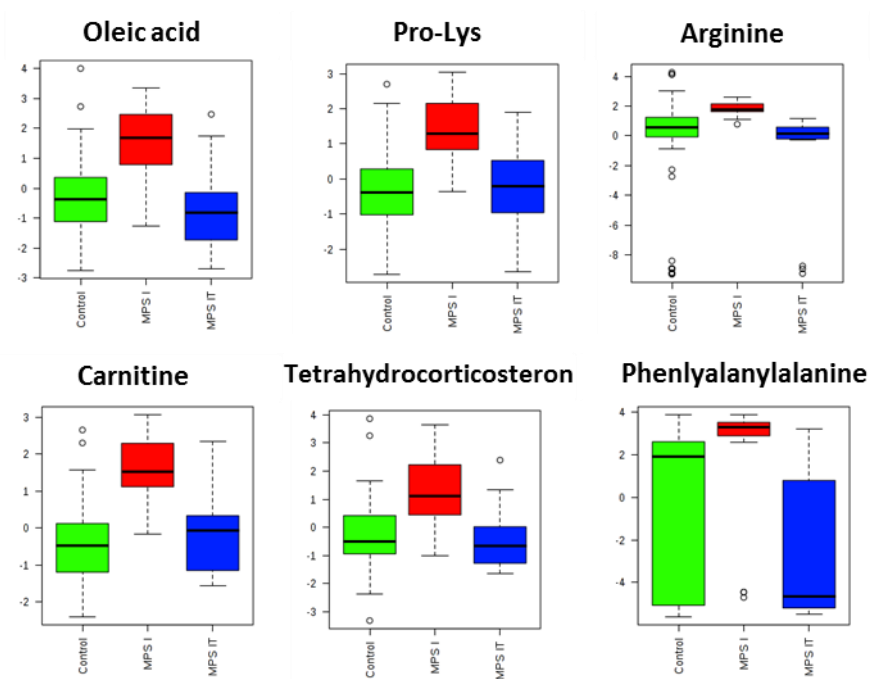


Figure S9. Area under the receiver operating characteristic (ROC) curves, comparing diagnostic performance of different discriminant features retrieved from untargeted metabolomics to differentiate MPS I from Control samples. Up) Arginine, Carnitine, Tetrahydrocorticosteron and S-(5'-Adenosyl)-L-Methionine. Down) Comparison of different combinations of features using PLSDA model with three components each. Combining features does not show improvement in AUC.

AUC: Area under the curve. False positive rate = 100 – Specificity. True positive rate = Sensitivity.

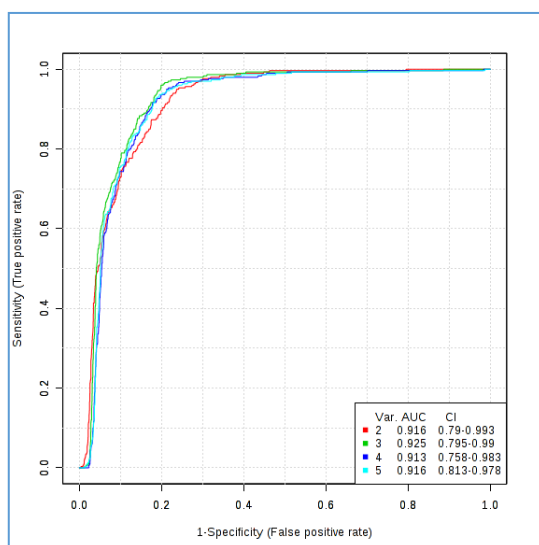
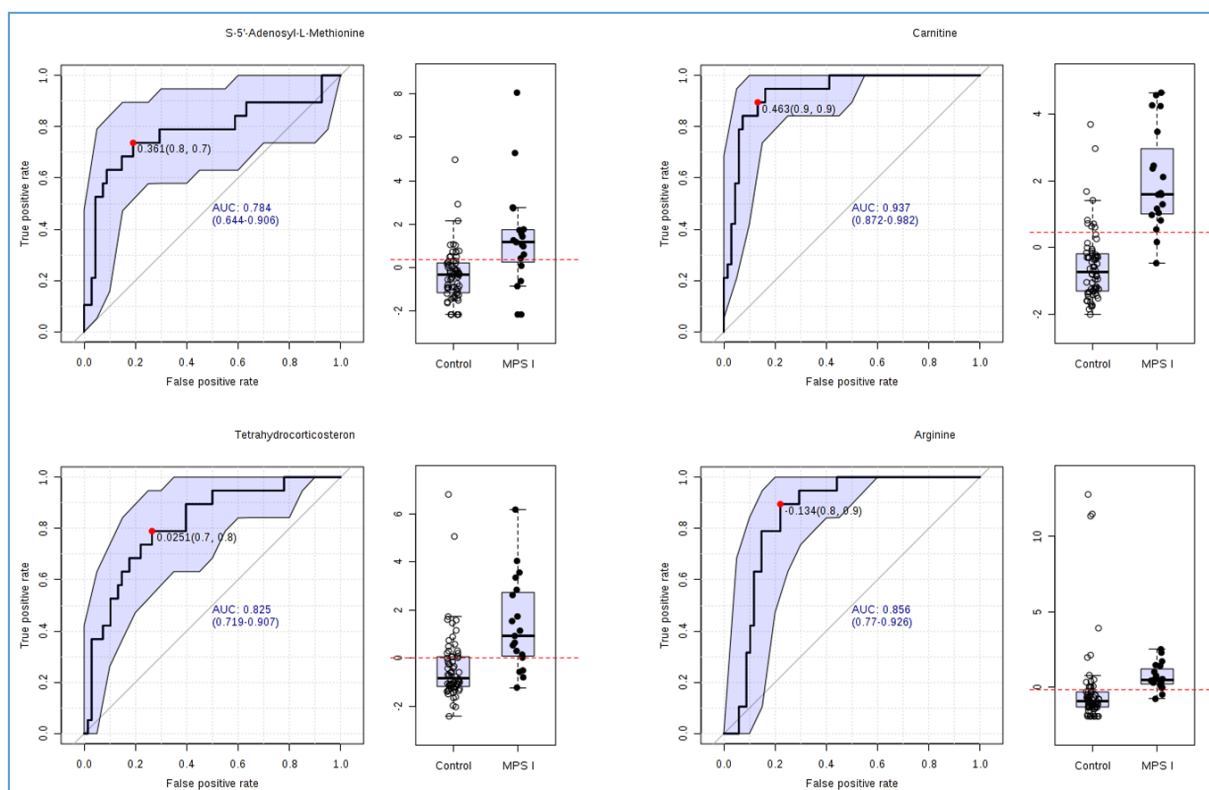


Figure S10. Boxplots of amino acids concentrations across the three studied groups: MPS I, MPS IT and Control samples.

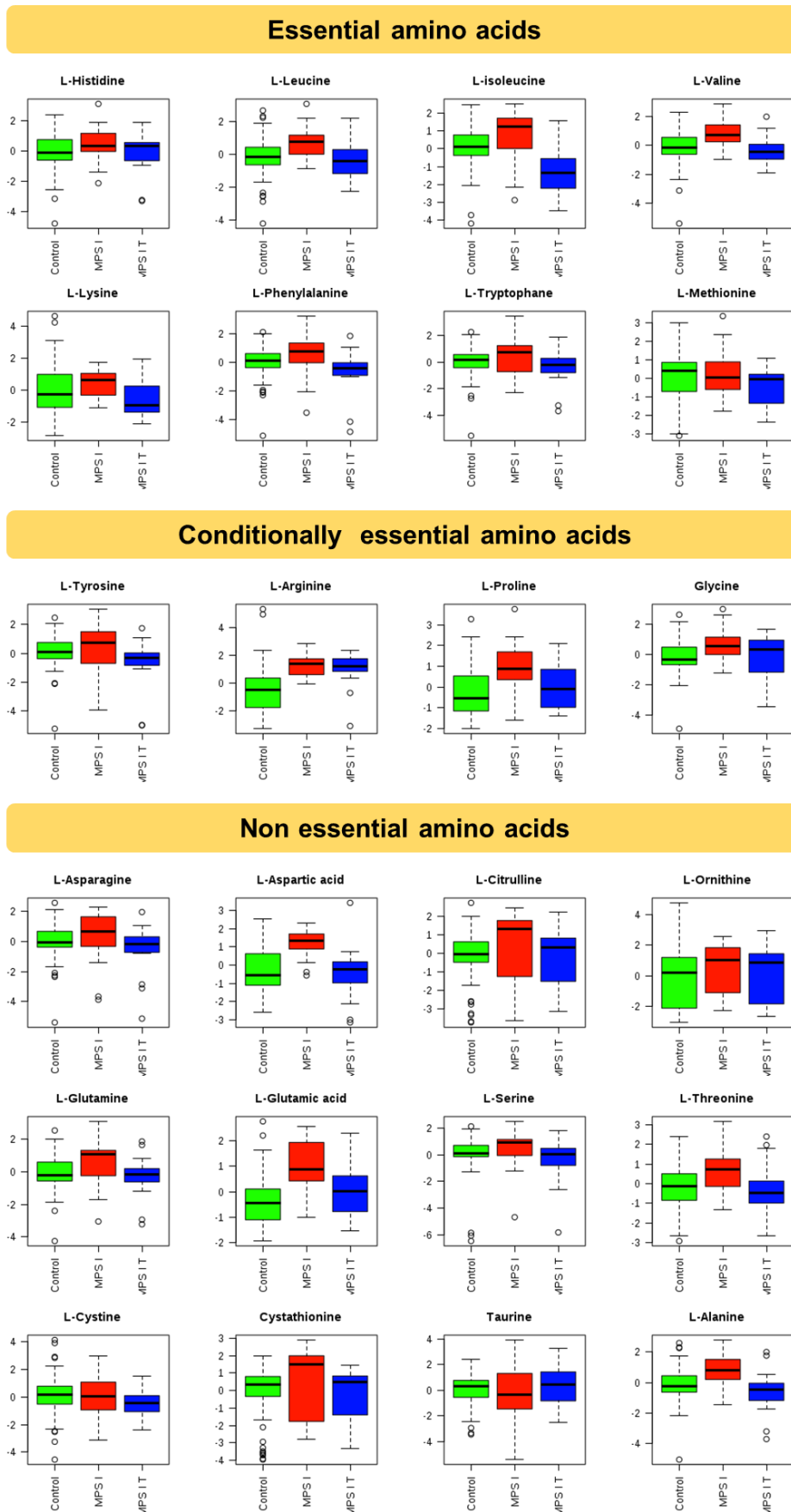


Figure S11. Boxplots of Heparan sulfate, Keratan sulfate, Dermatan sulfate and Total Glycoaminoglycans (Total GAGs) in the three groups: MPS I, MPS IT and Control samples.

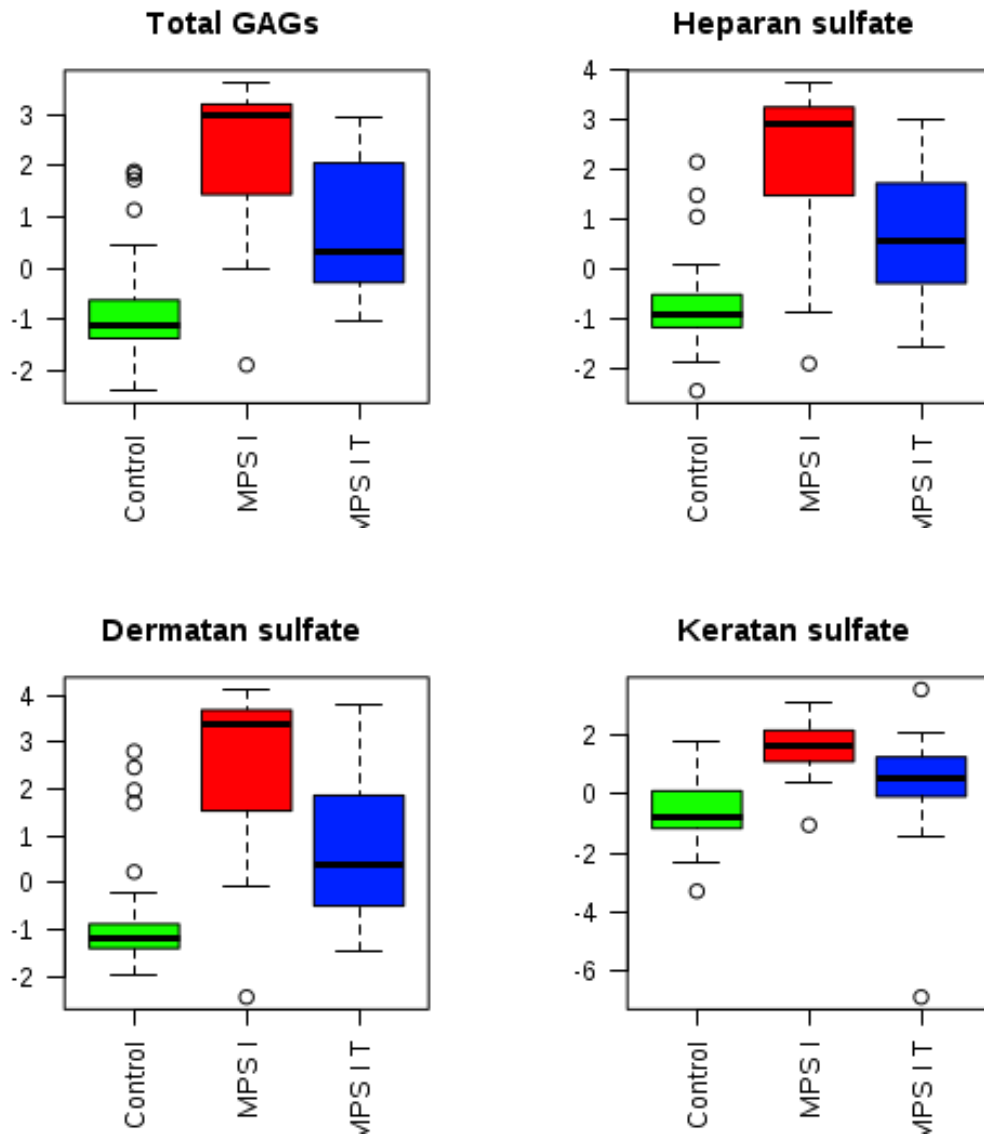
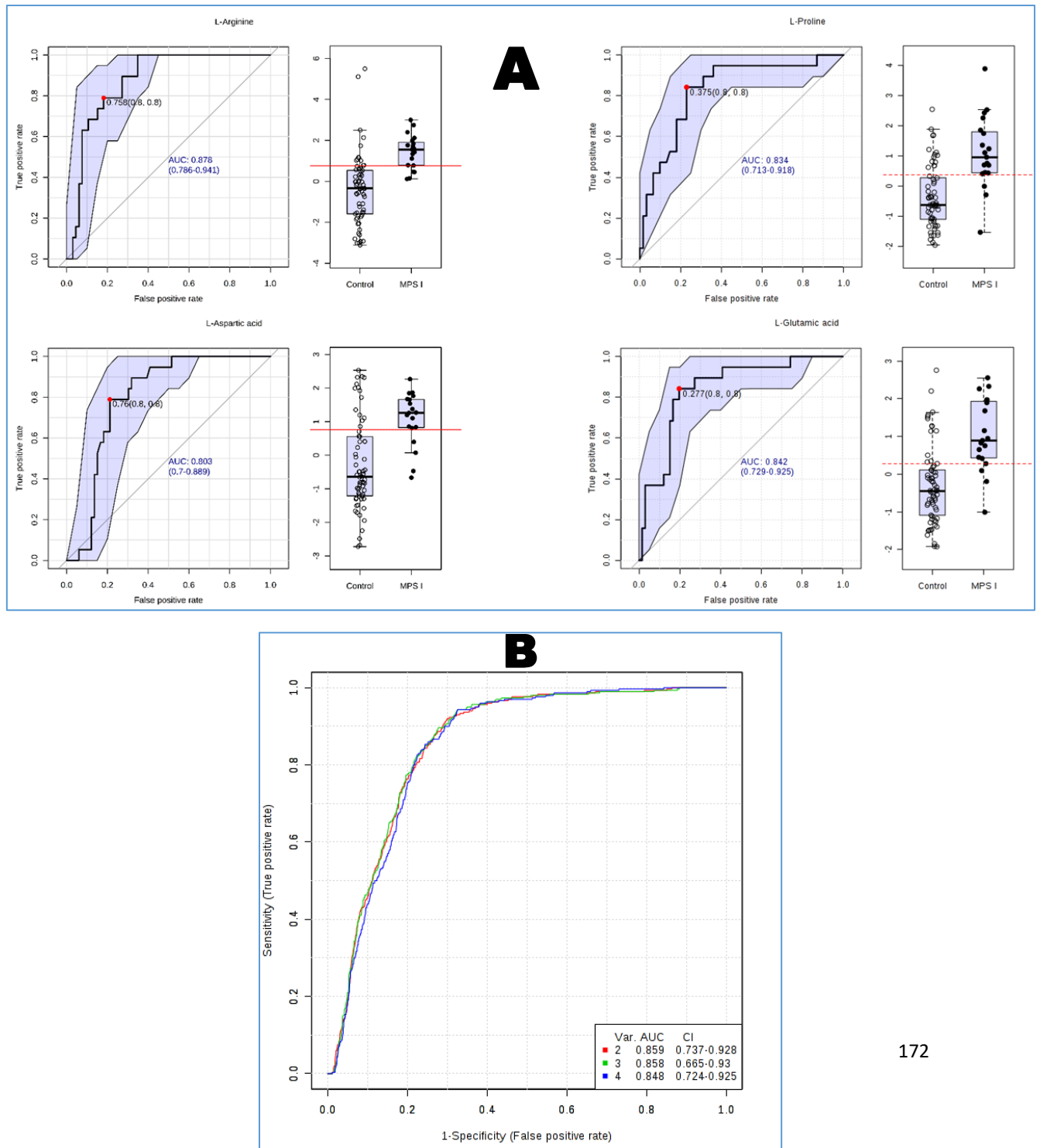


Figure S12. Area under the receiver operating characteristic (ROC) curves, comparing diagnostic performance of the most significant quantified aminoacids to differentiate MPS I from Control samples. A) Arginine, Proline, Aspartic acid, and Glutamic acid. B) A comparison of different combinations of the four amino acids using a PLSDA model with three components each is presented. Combining amino acids does not show improvement in AUC.

AUC, Area under the curve. False positive rate = 100 – Specificity. True positive rate = Sensitivity.



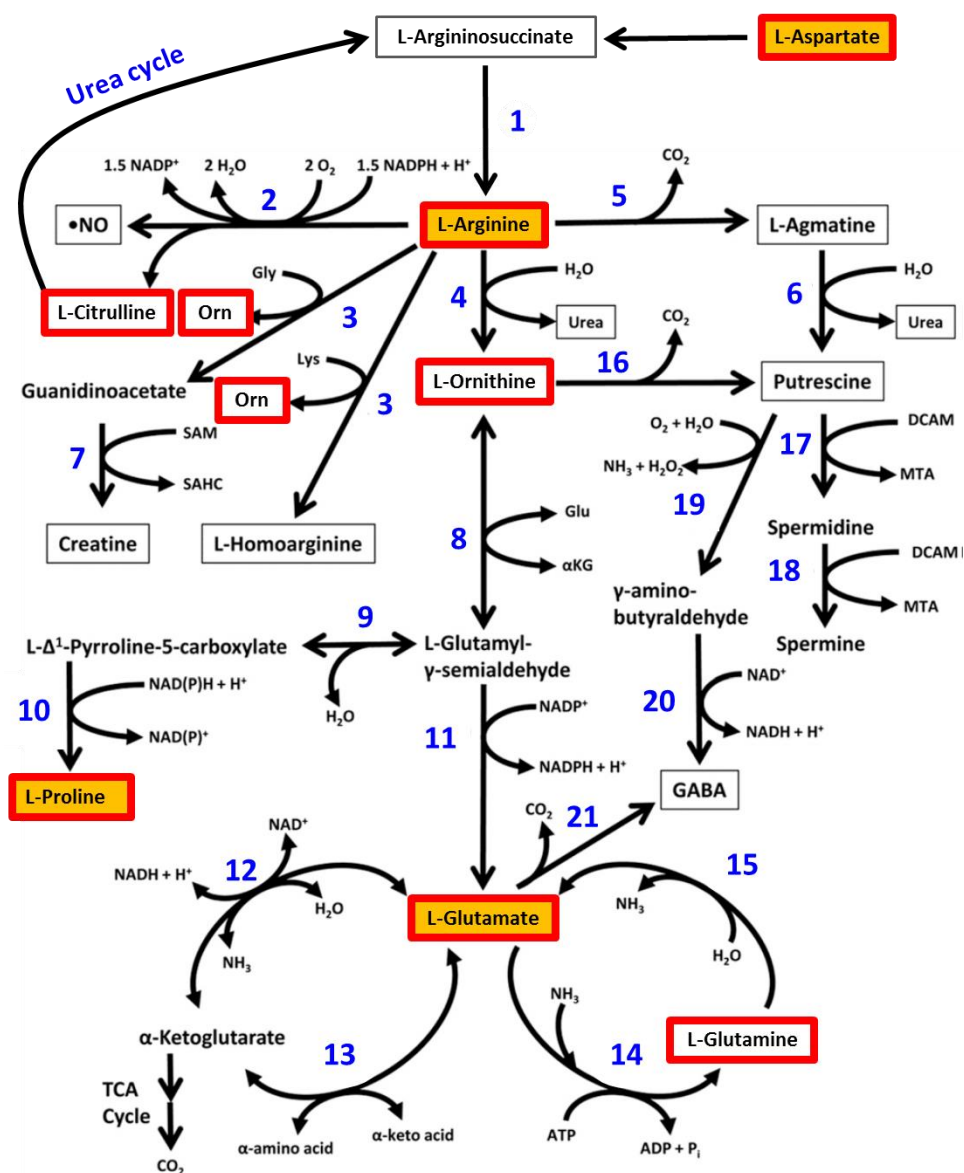
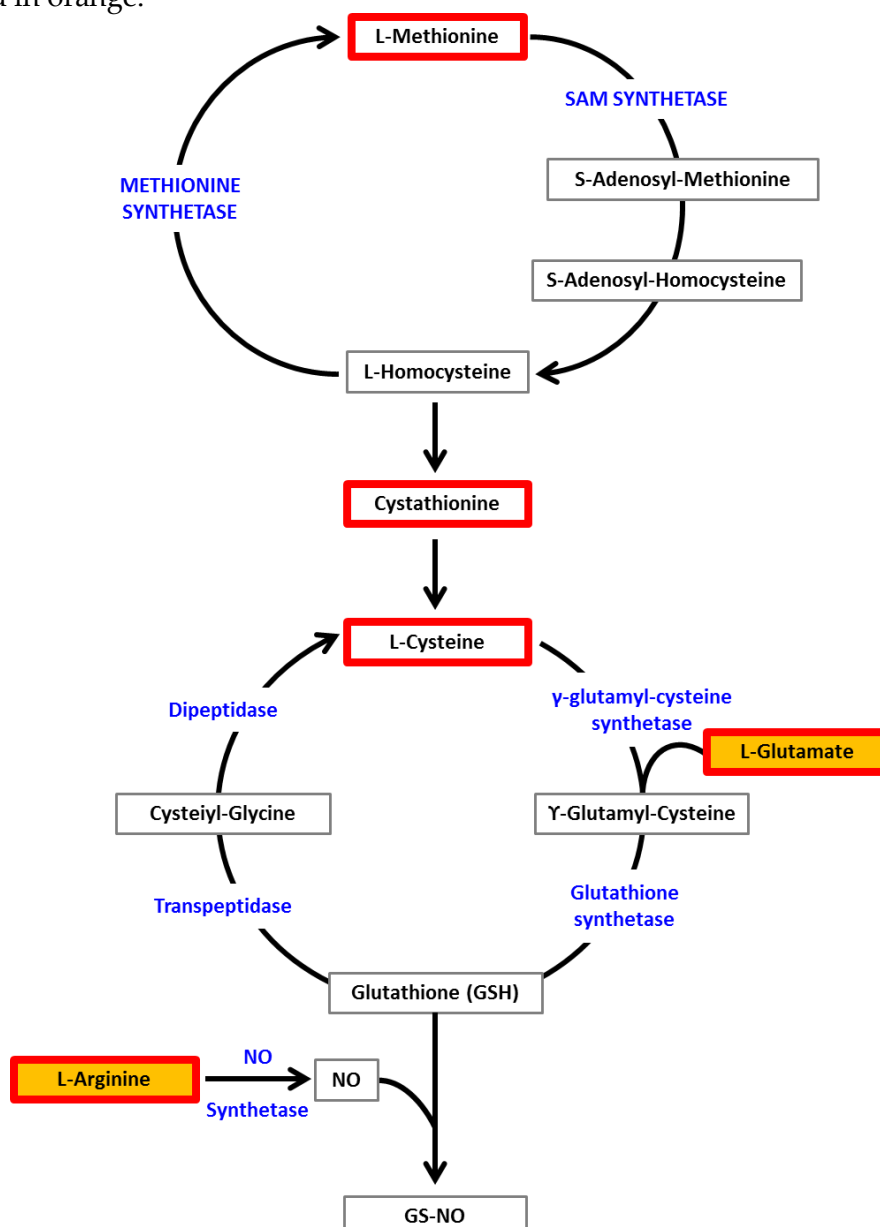


Figure S13. Pathways of arginine metabolism. Enzymes that catalyze the indicated reactions are as follows: 1: argininosuccinic lyase, 2: NO synthases. 3: arginine:glycine amidinotransferase. 4: arginase. 5, arginine decarboxylase. 6: agmatinase (agmatine ureohydrolase). 7: guanidinoacetate N-methyltransferase. 9: ornithine aminotransferase. 10: pyrroline- 5-carboxylate reductase. 11: pyrroline- 5-carboxylate dehydrogenase. 12: glutamate dehydrogenase. 13: alanine aminotransferase, aspartate aminotransferase, or branched-chain amino acid aminotransferase. 14: glutamine synthetase. 15: glutaminase. 16: ornithine decarboxylase. 17: spermidine synthase. 18: spermine synthase. 19: diamine oxidase. 20, aldehyde dehydrogenase. and 21: glutamate decarboxylase. Complete oxidation of arginine-derived α-ketoglutarate occurs via the citric acid cycle. Step 8 is a spontaneous, nonenzymatic reaction. DCAM: decarboxylated S-adenosylmethionine. Glu: L-glutamate. MTA: methylthioadenosine. SAHC: S-adenosylhomocysteine. SAM: S-adenosylmethionine. αKG: α-ketoglutarate. The metabolites that were measured in the present study are showed in red boxes. Elevated metabolites are highlighted in orange.

Figure S14. Schematic representation of aminothiols synthesis and glutathione metabolism. GSH is synthesized from cysteine, glutamate and glycine by the sequential action of GCL (glutamate–cysteine ligase) and GS (glutathione synthase). Intercellular transport of GSH involves its breakdown by the action of γ -glutamyl transpeptidase and dipeptidase, forming Cys-Gly and cysteine respectively. Alternatively, GSH may be oxidized to its disulphide, GSSG, and either GSH or GSSG may interact with protein cysteinyl thiols. Analogously, cysteine may be oxidized to cystine. Cysteine and Cys-Gly may also interact with proteins. Methionine is the precursor for cysteine synthesis, in which homocysteine and cystathionine are intermediates. Homocysteine is the substrate for the regeneration of methionine by methionine synthase. Methionine and homocysteine may also be oxidized to methionine sulphoxide and homocystine respectively as a result of ROS attacks. SAM (S-adenosylmethionine) and SAH (S-adenosylhomocysteine), which are intermediates in the conversion of methionine into homocysteine, also serve as methyl group donors. Arginine also interacts with GSH metabolism through nitric oxide (NO) metabolism. The metabolites that were measured in the present study are showed in red boxes. Elevated metabolites are highlighted in orange.




RESEARCH

Open



Unveiling metabolic remodeling in mucopolysaccharidosis type III through integrative metabolomics and pathway analysis

Abdellah Tebani^{1,2,3}, Lenaig Abily-Donval^{2,4}, Isabelle Schmitz-Afonso³, Bénédicte Héron⁵, Monique Piraud⁶, Jérôme Ausseil⁷, Farid Zerimech⁸, Bruno Gonzalez², Stéphane Marret^{2,4}, Carlos Afonso³ and Soumeya Bekri^{1,2*} 

Abstract

Background : Metabolomics represent a valuable tool to recover biological information using body fluids and may help to characterize pathophysiological mechanisms of the studied disease. This approach has not been widely used to explore inherited metabolic diseases. This study investigates mucopolysaccharidosis type III (MPS III). A thorough and holistic understanding of metabolic remodeling in MPS III may allow the development, improvement and personalization of patient care.

Methods : We applied both targeted and untargeted metabolomics to urine samples obtained from a French cohort of 49 patients, consisting of 13 MPS IIIA, 16 MPS IIIB, 13 MPS IIIC, and 7 MPS IIID, along with 66 controls. The analytical strategy is based on ultra-high-performance liquid chromatography combined with ion mobility and high-resolution mass spectrometry. Twenty-four amino acids have been assessed using tandem mass spectrometry combined with liquid chromatography. Multivariate data modeling has been used for discriminant metabolite selection. Pathway analysis has been performed to retrieve metabolic pathways impairments.

Results : Data analysis revealed distinct biochemical profiles. These metabolic patterns, particularly those related to the amino acid metabolisms, allowed the different studied groups to be distinguished. Pathway analysis unveiled major amino acid pathways impairments in MPS III mainly arginine–proline metabolism and urea cycle metabolism.

Conclusion : This represents one of the first metabolomics-based investigations of MPS III. These results may shed light on MPS III pathophysiology and could help to set more targeted studies to infer the biomarkers of the affected pathways, which is crucial for rare conditions such as MPS III.

Keywords : Metabolomics, Inborn errors of metabolism, Mucopolysaccharidosis type III, Lysosomal storage diseases, Mass spectrometry, Ion mobility

*Correspondence : soumeya.bekri@chu-rouen.fr

¹ Department of Metabolic Biochemistry, Rouen University Hospital, 76000 Rouen Cedex, France

Full list of author information is available at the end of the article



Background

Metabolism is a complex, interconnected and finely regulated network. It is composed of reactions biochemical processes that transform endogenous or exogenous substrates into vital products for cell, tissue and organism function. As a result, deregulation of this homeostasis underlies the pathophysiological mechanisms of different diseases [1]. An alteration of a metabolic pathway may be related to nutritional, environmental, or genetic factors. Inborn errors of metabolism (IEM) are rare diseases mainly due to a genetic defect enzymes or cofactors involved in a metabolic pathway or in intra- or intercellular metabolites. For better management of IEM patients, rapid and accurate biochemical and molecular tests are needed. Omics are very appealing to speed up both their molecular understanding and may lead to more efficient biomarkers. Omics are very appealing to achieve holistic and systemic aspects of diseases [1]. The metabolome refers to all metabolites present in a given biological system [2]. Metabolomics is an “omics” technology that allows metabolome characterization [3, 4]. Metabolomics is particularly interesting in exploring IEM given their intrinsic with metabolism [5]. Lysosomal storage diseases (LSD) represent a group of about 50 inherited disorders related to deficient lysosomal proteins. This impairment leads to a progressive accumulation of metabolites or macromolecules within the lysosomes. This storage causes, at least partly, various organ failures [6]. Mucopolysaccharidoses (MPS) are a subgroup of LSD. They are related to impaired catabolism of glycosaminoglycans (GAGs), chondroitin sulfate (CS), dermatan sulfate (DS), heparan sulfate (HS), keratan sulfate (KS), and hyaluronan, leading to GAG accumulation in the lysosomes and extracellular matrix [7]. This accumulation leads to multiple progressive tissue and organ failures [8]. Seven distinct forms of MPS are described and related to 11 known enzyme deficiencies [6]. Overall incidence is more than 1 in 30,000 live births [9]. Most MPS patients are asymptomatic after birth, but prenatal symptoms may be observed in MPS I, MPS IVA, MPS VI and more frequently in MPS VII. Depending on the patient and the MPS subtype, symptoms and severity may vary. Different MPS treatments are either in clinical use or under clinical trials [10]. Mucopolysaccharidosis type III (MPS III), or Sanfilippo syndrome, is caused by a congenital deficiency of one of the four enzymes involved in the degradation of HS [11]. Four subtypes, MPS IIIA, MPS IIIB, MPS IIIC and MPS IIID, have an autosomal recessive inheritance [12]. Typically, patients with Sanfilippo disease present with no obvious clinical features

prior to age 1–3 age years. Growth parameters may be higher compared to the reference range in the first years of life, while a growth delay may be observed in older patients. In all MPS III subtypes, central nervous system (CNS) involvement predominates (neurodegeneration, hyperactivity and behavioral disturbances) with less pronounced skeletal abnormalities and organomegaly. In addition, the clinical picture includes hirsutism, coarse facial features, cardiomegaly, thick hair, cloudy cornea, recurrent diarrhea, otitis and dysarthria [12–15]. MPS IIIA is the most severe type with an earlier onset and a rapid neurological deterioration. The first signs occur at around 1–3 years of age and the clinical symptoms worsen gradually and inevitably, resulting in the onset of severe dementia and a complete loss of motor functions. As other inherited metabolic diseases, the symptoms show high variability among patients even within the same family. Patients usually die before the third decade of life, although patients with a mild phenotype and allelic heterogeneity have been reported [12–15]. MPS IIIA (OMIM #252900) is caused by Heparane-*N*-sulfatase (SGSH, EC 3.10.1.1) deficiency with an incidence of 1 in 100,000 [16, 17]. MPS IIIB (OMIM #252920) is due to *N*-acetyl- α -glucosaminidase (NAGLU, EC 3.2.1.50) deficiency with an incidence of 1 in 200,000 [18]. MPS IIIC (OMIM #252930) is caused by heparan acetylCoA : α -glucosaminide-*N*-acetyltransferase (HGSNAT, EC 2.3.1.3) deficiency with an incidence of 1 in 1,500,000 [19]. MPS IIID (OMIM #252940) is due to *N*-acetylglucosamine-6-sulfatase (GNS, EC 3.1.6.14) deficiency with an incidence of 1 in 1,000,000 [20]. So far, no specific approved treatment is available. Gene therapy [21], bone marrow transplant [22], chaperon molecules [23], substrate deprivation therapy [24] and intrathecal enzyme therapy [25] are among the most active therapeutic research areas. The goal of this work is to apply both targeted and untargeted metabolomics on MPS IIIA, MPS IIIB, MPS IIIC and MPS IIID patients, compared to controls, to investigate metabolic changes in these conditions.

Methods

Urine samples

Random urine samples were collected from patients with a confirmed MPS diagnosis. Urine samples were collected within five expert centers for inherited metabolic diseases in France. The 49 untreated MPS III patients were evaluated as follows : 13 MPS IIIA patients : 6 males (age range from 5.1 to 12.0 years, mean age 6.2 years) and 7 females (age range from 1.9 to 18.4 years, mean age : 6.8 years); 16 MPS IIIB patients : 7 males (age range from 3.8

to 9.8 years, mean age 7.2 years) and 9 females (age range from 2.9 to 11.7 years, mean age 6.3 years); 13 MPS IIIC patients : 7 males (age range from 6.4 to 20.6 years, mean age :12.1 years) and 6 females (age range from 2.8 to 31.1 years, mean age 10.0 years); 7 MPS IIID patients : 3 males (age range from 3.8 to 17.5 years, mean age 8.9 years) and 4 females (age range from 3.4 to 18.7 years, mean age 7.8 years). Moreover, control urine samples were also collected from 66 healthy subjects, 27 males and 39 females (age range from 5.5 to 70 years, mean age 40.8 years). This project was approved by the Research Ethics Board of Rouen University Hospital (CERNI E2016-21).

Metabolic phenotyping

The protocol used in this study has been previously described [26]. Briefly, urine samples were processed by transferring 200 μ L of urine to 1.5 mL tubes and centrifuging at 4 °C for 10 min at 13,000g; then 100 μ L of ultrapure water was added to 100 μ L of supernatant and mixed. For untargeted metabolomics data acquisition, Ultraperformance liquid chromatography–ion mobility mass spectrometry on a Synapt G2 HDMS (Waters, Saint-Quentin-en-Yvelines, France) mass spectrometer as previously described [27]. Regarding targeted analysis, free amino acid profiles in urine was based on liquid chromatography coupled to a tandem mass spectrometry. Detailed protocols are presented in the Additional file 1.

Data analysis

A one-way analysis of variance (ANOVA) test was applied for multiple groups testing while a *t* test is used for binary comparisons. The Benjamini and Hochberg false discovery rate (FDR) method was used for multiple testing corrections with an FDR cut-off level of 5%. A Receiver operating characteristic curve (ROC) has been used to assess the diagnostic performance of the chosen classifiers. Support vector regression normalization for untargeted metabolomics data [28]. The normalized data has been log-transformed and pareto-scaled. All data modeling and analysis is done using SIMCA 15.0 (MKS DAS, Umeå, Sweden) and R software. The Mummichog algorithm has been used for pathway analysis [29] while MetaboAnalyst has been used for Metabolite Set Enrichment Analysis on the amino acid data [30]. Details regarding data modeling and validation results from all models are provided in Additional file 1. Figure 1 presents an overview of the implemented metabolomics workflow.

Results Untargeted analysis

The heatmap in Fig. 2a depicts the top 100 features ranked by ANOVA ($p < 0.05$ cut-off and FDR 5%). The results highlight correct clustering of the different sample groups and the dendrogram structure, using Euclidean distance, shows two main clusters of variables.

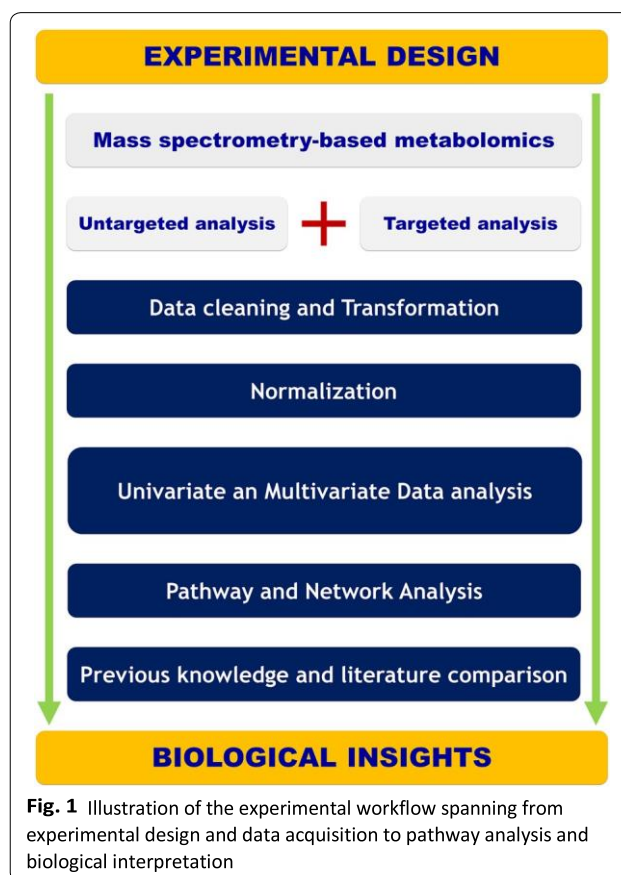
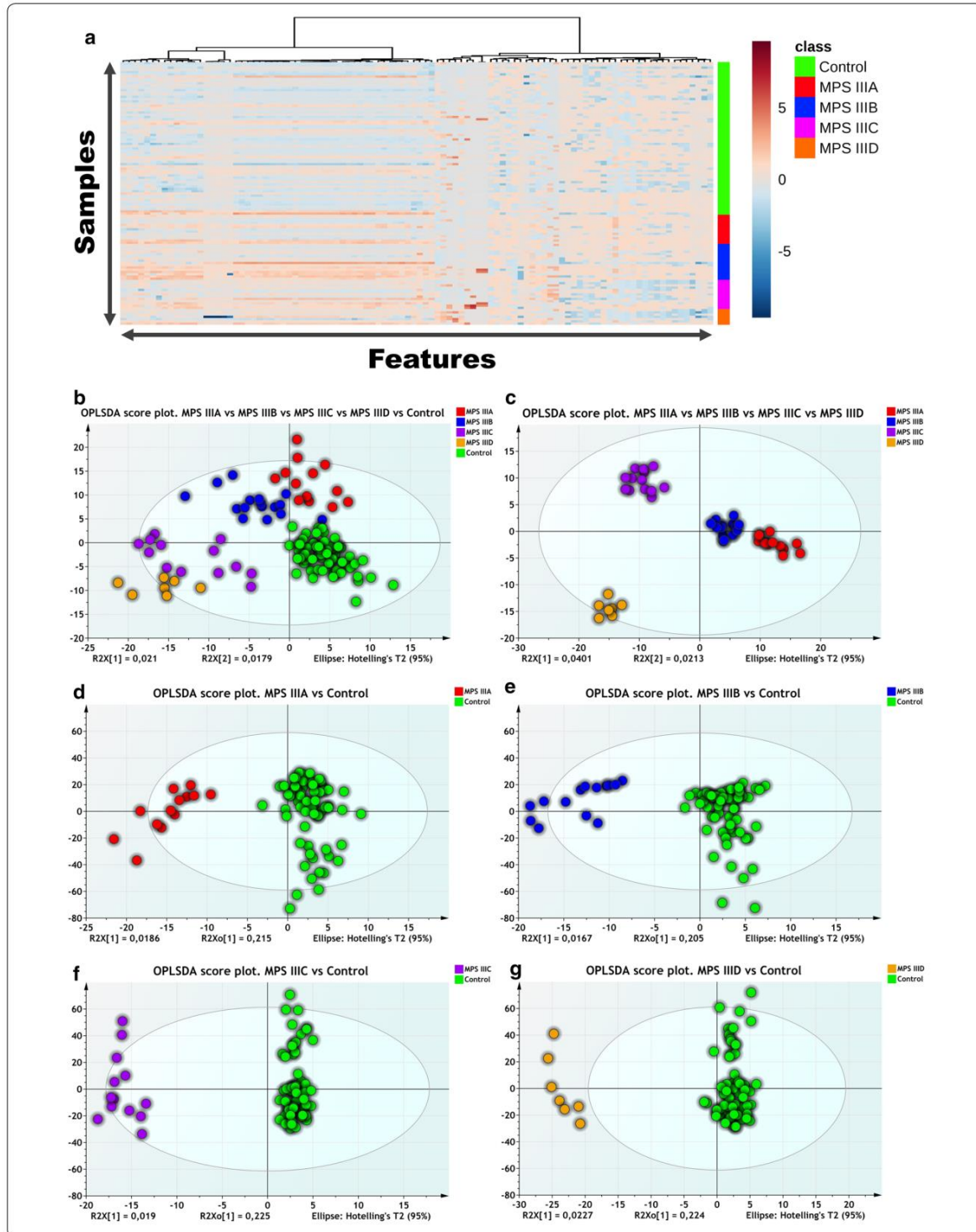


Fig. 1 Illustration of the experimental workflow spanning from experimental design and data acquisition to pathway analysis and biological interpretation

We applied principal component analysis (PCA) to further analyze the underlying differential metabolic profiles. A three-component PCA model accounting for 18% of the total variance has been built. Trends, groups and potential outliers within the data are investigated using score plots. For predictive classification purposes, supervised methods are used since they allow the accurate modeling of the relationship between controls, MPS IIIA, IIIB, IIIC and IIID samples. First, an OPLS-DA classification was applied to the whole dataset. Samples were labeled according to the corresponding groups, MPS IIIA, IIIB, IIIC, IIID and control (Fig. 2b, c). A negative Q2 regression line intercept resulting from the permutation test

allows the cross-validation of OPLSDA models. The final model had an $R^2 = 0.77$ and $Q^2 = 0.13$. The OPLS-DA scores plots (Fig. 2b) shows a clear separation, suggesting that the OPLS-DA model successfully classified samples according to their respective metabolic profiles. Model

validation is assessed both by CV-ANOVA (p -value = 3×10^{-2}) and by the permutation test (999 permutations gave a negative Q^2 intercept). Details regarding model validation are shown in Additional file 1 :



(See figure on previous page.)

Fig. 2 **a** Hierarchical cluster analysis and heat map visualization of top 100 variables (x-axis) ranked by ANOVA. The urine sample classes are represented along the y-axis. The color code was used to represent log-scaled intensities of features between -5 (blue) and $+5$ (brown), showing the relative abundance of the features according to the groups. **b** OPLSDA scores plot ($R^2 = 0.77$, $Q^2 = 0.13$) shows a clear separation between the different diseased and control groups (MPSIIIA, MPSIIIB, MPSIIIC and MPSIIID and control). **c** OPLSDA scores plot ($R^2 = 0.93$, $Q^2 = 0.05$) shows a clear separation between the different diseased groups (MPSIIIA, MPSIIIB, MPSIIIC and MPSIIIC). **d** Clear separation between MPSIIIA and control samples is observed ($R^2 = 0.89$, $Q^2 = 0.23$). **e** Clear separation of MPSIIIB samples from the controls is observed ($R^2 = 0.89$, $Q^2 = 0.21$). **f** Clear separation of MPSIIIC samples from the controls is observed ($R^2 = 0.98$, $Q^2 = 0.39$). **g** Clear separation of MPSIIID samples from the controls is observed ($R^2 = 0.95$, $Q^2 = 0.36$). Detailed model characteristics and validation are given in Additional file 1

Fig. S3). Furthermore, separate binary OPLS-DA classification models have been built for each disease sample vs control. For control and MPS IIIA samples the model had one predictive and two orthogonal components, and its validation parameters were as follows : $R^2 = 0.89$, $Q^2 = 0.23$ and CV-ANOVA p-value = 2.84×10^{-3} (Additional file 1 : Fig. S4). The corresponding score plot is shown in Fig. 2d. It shows a clear separation between the two classes on the predictive component. For MPSIIIB and control samples the model had one predictive and two orthogonal components model with $R^2=0.89$, $Q^2=0.21$ and CV-ANOVA p-value= 5.28×10^{-3} (Fig. 2e). For MPSIIIC and control samples, the model has one predictive and three orthogonal components with $R^2=0.98$, $Q^2=0.39$ and CV-ANOVA p-value= 1.35×10^{-5} (Fig. 2f). Another OPLS-DA model was built for MPSIIID and control samples with one predictive and two orthogonal components model with $R^2=0.95$, $Q^2=0.36$ and CV-ANOVA p-value= 1.83×10^{-5} (Fig. 2g). To select discriminant variables, their respective VIP scores for each validated OPLS-DA model have been used. Based on 1 as a cutoff value, 25 features out of 854 were selected for the MPSIIIA VS Control model, 243 for MPSIIIB vs control, 247 for MPS IIIC vs control and 262 for the MPSIIID vs control model. The variables lists have been refined by retaining only the most discriminant variables and their putative annotation. The list included N-acetylserotonin, N-succinyl-l,l-2,6-diaminopimelate, octanoylglucuronide and 3-2-hydroxyphenyl-propanoic acid. These discriminant variables are depicted in Tables 1 and 2 with their respective statistical metrics and annotation. Boxplots of the main discriminant features are presented in Additional file 1 : Fig. S7. Using the area under the ROC curves (AUC), the discriminant performances of these features are also investigated. N-Acetylserotonin has the highest AUC for MPS IIIA (AUC = 0.83) and MPS IIIB (AUC = 0.83). N-Succinyl-l,l-2,6-diaminopimelate has the highest AUC (0.73) for MPS IIIC and octanoylglucuronide performed best for MPS IIID with an AUC = 0.79. The results are shown in Tables 1 and 2. Furthermore, the underlying impaired pathways in each disease are explored using Mummichog. The results are shown in Table 3. Interestingly, amino acid metabolisms and fatty acid pathways were markedly dysregulated.

Targeted analysis

We also quantified twenty-four amino acids and Additional file 1 : Table S4 presents their absolute urine concentrations. Boxplots of normalized amino acid concentrations are shown in Additional file 1 : Fig. S8 and the statistical metrics are presented in Table 4. MPS IIIA yielded 11 significantly changed amino acids compared to controls : arginine, aspartic acid, alanine, threonine, histidine, phenylalanine, glycine, proline, asparagine and tyrosine. For MPS IIIB vs control, arginine, aspartic acid, alanine, threonine, histidine, phenylalanine, glycine, proline, glutamine, asparagine, tyrosine and leucine showed significant differences. regarding MPS IIIC vs control, 6 amino acids showed differences : arginine, aspartic acid, serine, isoleucine, methionine and citrulline. For MPS IIID vs control, 6 amino acids showed differences : arginine, alanine, threonine, glycine, glutamine and citrulline. To holistically determine the amino acid profile differences between controls and each of the MPS III subtype patients, the amino acids concentrations were assessed using an ANOVA test. The analysis yielded 17 amino acids above the $p < 0.05$ cut-off (FDR 5%). A hierarchical clustering analysis was applied to group samples according to their profile similarities. The heatmap in Fig. 3a represents the 24 amino acids ranked by ANOVA. The results show that all samples belonging to the same group were correctly clustered. The dendrogram structure, using Euclidean distance, highlights two main clusters of variables. Furthermore, a correlation analysis has been performed. Figure 3b–e presents the heatmap of the correlation analysis for MPS IIIA, MPS IIIB, MPS IIIC and MPS IIID, respectively. Both figures show a clear cluster of variables that have high correlation. Figure 3b (MPS IIIA vs control) shows a main cluster including alanine, leucine, valine, glycine, tyrosine, threonine, isoleucine, histidine, lysine, tryptophan, serine, asparagine, glutamine, phenylalanine, cystine and methionine. Regarding MPS IIIB vs control, Fig. 3c shows two main clusters : the main one includes methionine, isoleucine, serine,

Table 1 Some discriminant features, putatively annotated, extracted by the different OPLS-DA models for MPSIIIA, MPSIIIB, MPSIIIC and MPSIIID

HMDB	Putative annotation	Formula	M	m/z	Adduct	Δ m/z (ppm)	t_R (min)	tD (ms)	CCS (Å ²)	%RSD
HMDB01238	N-Acetylsertonin	C ₁₂ H ₁₄ N ₂ O ₂	218.1055	241.0985	M + Na	13	5.44	2.65	146.0	16.76
HMDB12267	N-Succinyl-L,L-2,6-diaminopimelate	C ₁₁ H ₁₈ N ₂ O ₇	290.1114	291.1223	M + H	10	7.48	3.24	162.5	9.90
HMDB33752	3-2-Hydroxyphenyl-propanoic acid	C ₉ H ₁₀ O ₃	166.0629	199.0970	M + CH ₃ OH + H	3	6.77	2.38	138.8	21.89
HMDB10347	Octanoylglucuronide	C ₁₄ H ₂₄ O ₈	320.1471	303.1445	M - H ₂ O + H	0	7.05	3.40	166.8	16.52

M monoisotopic mass, ppm parts per million, t_R retention time, tD drift time, CCS cross collision section, VIP variable importance in projection

Table 2 Statistical and discriminant metrics of the selected annotated features

HMDB	Putative annotation	MPSIIIA			MPSIIIB			MPSIIIC			MPSIIID		
		FDR	AUC	VIP	FDR	AUC	VIP	FDR	AUC	VIP	FDR	AUC	VIP
HMDB01238	<i>N</i> -Acetylserotonin	<i>2.28E-03</i>	<i>0.83</i>	<i>1.21</i>	<i>6.05E-01</i>	<i>0.83</i>	<i>0.93</i>	8.61E-01	0.61	0.99	9.70E-01	0.54	0.35
HMDB12267	<i>N</i> -Succinyl-L,L-2,6-diaminopimelate	3.88E-01	0.60	0.56	7.44E-01	0.63	0.64	<i>9.06E-03</i>	<i>0.73</i>	<i>1.29</i>	6.56E-01	0.69	0.05
HMDB33752	3-2-Hydroxyphenyl-propanoic acid	<i>1.76E-02</i>	<i>0.76</i>	<i>1.53</i>	5.30E-01	0.77	1.76	3.31E-01	0.66	0.93	7.56E-01	0.62	0.22
HMDB10347	Octanoylglucuronide	6.75E-01	0.58	0.69	3.86E-01	0.72	1.95	8.54E-01	0.64	0.51	<i>2.57E-02</i>	<i>0.79</i>	<i>1.90</i>

FDR false discovery rate, AUC area under the curve, VIP variable importance in projection

Significant features are highlighted in italics for control and each disease comparison (false discovery rate FDR = 5%)

Table 3 Significantly dysregulated pathways

	Pathway	Overlap size	p-value (FDR = 5%)
MPS IIIA	Vitamin B1 (thiamin) metabolism	2	2.77E-03
	Pyrimidine metabolism	2	9.86E-03
MPS IIIB	TCA cycle	2	1.41E-03
	Aspartate and asparagine metabolism	4	1.81E-03
	Vitamin E metabolism	3	3.00E-03
	Methionine and cysteine metabolism	3	4.78E-03
	Fatty acid activation	2	4.89E-03
	Lysine metabolism	2	9.66E-03
	De novo fatty acid biosynthesis	2	1.31E-02
	Tryptophan metabolism	3	3.31E-02
MPS IIIC	Vitamin B1 (thiamin) metabolism	2	5.54E-04
	Omega-3 fatty acid metabolism	2	8.27E-04
	Butanoate metabolism	2	1.71E-03
	Tryptophan metabolism	3	8.07E-03
	Linoleate metabolism	2	1.18E-02
	Tyrosine metabolism	3	2.28E-02
MPS IIID	Methionine and cysteine metabolism	2	2.60E-02
	TCA cycle	2	1.41E-03
	Vitamin B1 (thiamin) metabolism	2	1.41E-03
	Aspartate and asparagine metabolism	4	3.08E-03
	Butanoate metabolism	2	5.14E-03
	Carnitine shuttle	2	2.32E-02
	Arginine-proline metabolism	2	2.91E-02
	Tryptophan metabolism	3	4.36E-02

FDR false discovery rate

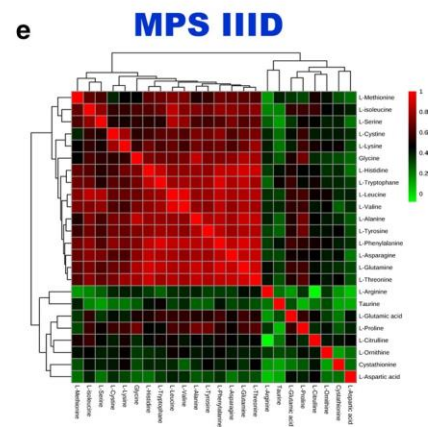
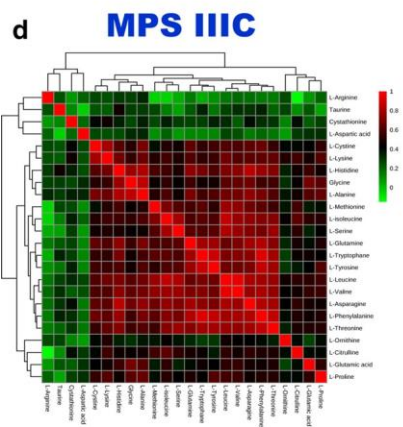
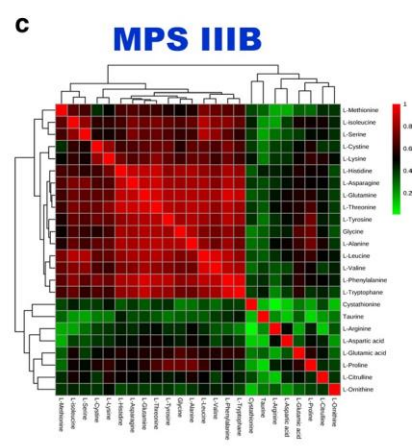
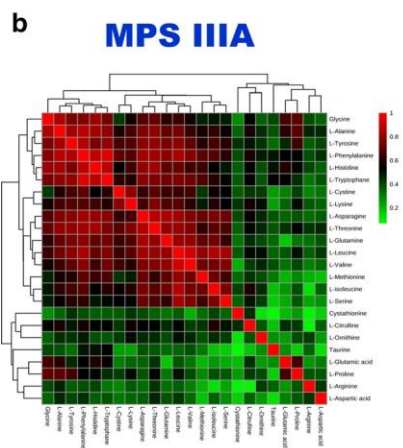
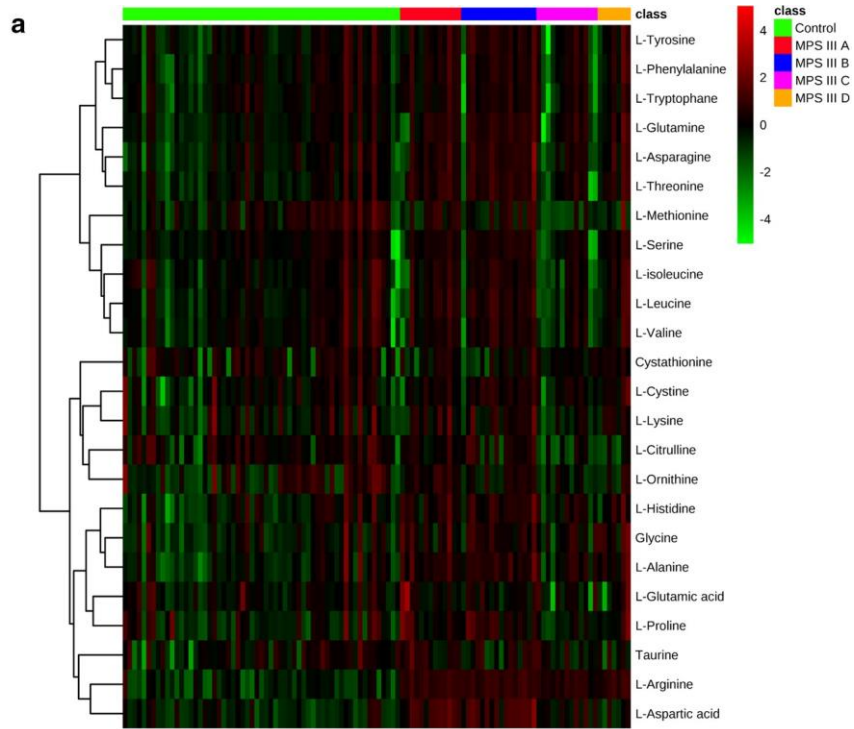
cystine, lysine, histidine, asparagine, glutamine, threonine, tyrosine, glycine, alanine, leucine, valine, phenylalanine, and tryptophan. Regarding MPS IIIC vs control, Fig. 3d shows three clusters : the main one includes cystine, lysine, histidine, glycine, alanine, methionine, isoleucine, serine, glutamine, tryptophan, tyrosine, leucine, valine, asparagine, phenylalanine and threonine. For MPS IIID vs control, Fig. 3e shows two clusters : the main one includes methionine, isoleucine, serine, cystine, lysine, glycine, histidine, tryptophan, leucine, valine, leucine, valine, tyrosine, phenylalanine, asparagine, glutamine and threonine. To assess the diagnostic performance of the

different amino acids, we performed univariate ROC curve analyses for the different MPS III subtype compared to controls. For MPS IIIA, there were four amino acids with a high AUC above 0.80, including : arginine (0.98), aspartic acid (0.95), alanine (0.85) and threonine (0.81). The same procedure was performed for MPS IIIB vs control and indicated seven amino acids with a high AUC above 0.80 and these were : arginine (0.98), aspartic acid (0.94), Alanine (0.87) and threonine (0.86), histidine (0.81), glutamine (0.87), asparagine (0.83). For MPS IIIC, the results showed only Arginine with a high AUC (0.95).

Table 4 Fold change, t-test statistics, and area under the curve (AUC) of the receiver operating curves (ROC) for 24 amino acids, free carnitine and acylcarnitines ($p < 0.05$)

	MPSIIIA vs control			MPSIIIB vs control			MPSIIIC vs control			MPSIIID vs control		
	AUC	q-value (FDR)	Fold change	AUC	q-value (FDR)	Fold change	AUC	q-value (FDR)	Fold change	AUC	q-value (FDR)	Fold change
L-Arginine	0.98	7.77E-09	-4.53	0.98	1.45E-10	-3.90	0.95	4.26E-06	-3.60	0.98	1.08E-04	-3.75
L-Aspartic acid	0.95	2.10E-09	-3.83	0.94	1.29E-10	-3.30	0.77	1.43E-02	-1.45	0.56	6.06E-01	0.30
L-Alanine	0.86	1.32E-03	-1.11	0.87	1.81E-04	-0.45	0.67	2.55E-01	-0.33	0.81	4.63E-02	-0.50
L-Threonine	0.81	8.18E-03	-1.13	0.86	2.68E-04	-0.72	0.50	2.86E-01	0.91	0.78	4.63E-02	-0.61
L-Histidine	0.77	1.76E-02	-0.98	0.81	1.60E-03	-0.43	0.59	5.02E-01	-0.11	0.69	2.89E-01	-0.09
L-Phenylalanine	0.76	2.82E-02	-0.78	0.77	6.94E-02	0.08	0.52	4.02E-01	0.67	0.63	2.27E-02	-0.03
Glycine	0.75	1.76E-02	-0.82	0.78	3.34E-03	-0.13	0.63	4.09E-01	-0.12	0.81	2.27E-02	-0.69
L-Proline	0.73	5.17E-02	-0.69	0.77	3.93E-03	-0.16	0.52	6.61E-01	0.44	0.71	2.89E-01	0.03
L-Glutamine	0.73	1.98E-01	-0.46	0.87	1.04E-03	-0.49	0.55	4.02E-01	0.77	0.72	4.63E-02	-0.52
L-Asparagine	0.72	7.76E-02	-0.69	0.83	1.48E-03	-0.51	0.55	6.61E-01	0.46	0.70	1.67E-01	-0.30
L-Tyrosine	0.71	6.88E-02	-0.63	0.80	4.01E-03	-0.21	0.50	2.86E-01	0.89	0.64	2.97E-01	0.04
L-Tryptophan	0.70	1.19E-01	-0.49	0.74	2.59E-01	0.36	0.58	1.29E-01	1.13	0.55	5.60E-01	0.44
L-Leucine	0.67	1.84E-01	-0.49	0.79	2.31E-02	-0.08	0.64	2.55E-01	0.93	0.62	3.62E-01	0.08
Taurine	0.66	1.84E-01	-0.71	0.52	8.38E-01	0.73	0.56	9.70E-01	0.27	0.55	6.53E-01	1.12
L-Serine	0.65	4.20E-01	-0.44	0.78	1.01E-01	-0.20	0.63	8.53E-02	1.89	0.68	3.73E-01	-0.22
L-Glutamic acid	0.62	1.34E-01	-0.48	0.55	8.23E-01	0.78	0.52	2.86E-01	0.76	0.52	5.37E-01	1.12
L-Lysine	0.61	3.99E-01	-0.43	0.65	2.01E-01	0.10	0.59	3.72E-01	0.88	0.66	4.09E-01	0.02
L-Isoleucine	0.61	7.85E-01	-0.02	0.67	2.51E-01	0.36	0.68	4.42E-02	1.35	0.59	4.09E-01	0.18
L-Valine	0.60	6.68E-01	-0.10	0.70	2.64E-01	0.36	0.66	2.55E-01	0.99	0.58	4.09E-01	0.15
L-Ornithine	0.60	2.65E-01	-1.01	0.53	8.53E-01	0.70	0.61	2.86E-01	1.53	0.53	6.53E-01	1.31
L-Cystine	0.59	6.68E-01	-0.19	0.71	1.01E-01	-0.08	0.51	9.70E-01	0.32	0.69	2.89E-01	-0.38
Cystathionine	0.55	9.48E-01	0.25	0.60	5.76E-01	1.36	0.63	9.70E-01	0.25	0.53	5.60E-01	0.09
L-Methionine	0.54	9.48E-01	0.11	0.51	9.71E-01	0.83	0.79	3.29E-03	2.15	0.69	3.62E-01	1.60
L-Citrulline	0.52	8.20E-01	0.03	0.60	6.61E-01	0.54	0.74	9.43E-03	2.11	0.76	7.15E-02	2.48

Significant features are highlighted in italics (false discovery rate FDR = 5%)



(See figure on previous page.)

Fig. 3 a Heat map representing the clustering of 24 amino acids across the five groups of samples (MPS IIIA, MPS IIIB, MPS IIIC, MPS IIID and Controls). Columns represent individual samples and rows refer to amino acid. Shades of green or red represent elevation or decrease, respectively, of an amino acid. **b–e** Spearman rank-order correlation matrix 24 amino acids based on their concentrations profiles across all samples in MPS IIIA, MPS IIIB, MPS IIIC and MPS IIID respectively. Shades of green to red represent low-to-high correlation coefficient between markers

Regarding MPS IIID, three amino acids showed a high AUC : arginine (0.98), alanine (0.81) and glycine (0.81). The overall univariate and ROC analysis results are shown in Table 4 and Fig. 4. The ROC curves along with a comparison of the different combinations of the main significant amino acids have been performed using PLS-DA models with three components each. The results are presented in Additional file 1 : Fig. S9. Pathway analysis yielded the main impaired metabolisms. For MPS IIIA vs control and MPS IIIB vs control analyses, beta-alanine metabolism, malate–aspartate shuttle, arginine–proline, urea cycle and aspartate metabolism were among the most affected pathways. For the MPS IIIC vs control analysis, methionine metabolism, in addition to the abovementioned metabolic pathways, was the most affected. For the MPS IIID vs control analysis, urea cycle, arginine–proline metabolism, porphyrin metabolism and pyrimidine and purine metabolism were the most affected pathways. The overall results are shown in Fig. 5a–d for all the studied groups.

Discussion

In this study, MPS III urine patterns of metabolites have been studied to unveil the biochemical indicators that may differentiate MPS III patients from control individuals. Of note, the mean-age difference between the studied groups represent a drawback which is mainly due to pediatric recruitment difficulties and ethical considerations. However, given the stringency of the statistical cut-off and the applied multiple testing correction might circumvent some of these biases. Using untargeted metabolomics, we succeeded in building a predictive model with a clear separation between the different studied groups—MPS IIIA, MPS IIIB, MPS IIIC, MPS IIID and control sample—which is underlined by the metabolic pattern similarity in each group. The retrieved data revealed a profound metabolic modeling mainly of amino acid-related metabolism. In light of these results, targeted amino acid analysis has been performed, which confirmed the deep metabolic alterations. Using these data, pathway analysis succeeded in identifying the main disrupted pathways. Salazar et al. [31] reported a genome-scale human

metabolic reconstruction based approach to understand the effect of metabolism alterations in MPS. This *in silico* approach applied to MPS III subtypes (MPS IIIA, MPS IIIB, MPS IIIC and MPS IIID) by silencing, respectively, *SGSH*, *NAGLU*, *HGSNAT* and *GNS* genes, allowed the generation of models which were analyzed through flux balance and variability analysis. We performed a comparative analysis between the *in silico* systems based analysis data and the pathway analysis results of this present study. This comparison is illustrated by a Venn diagram (Fig. 5e) and showed two main common metabolisms : arginine–proline metabolism and urea cycle. Detailed data are presented in Additional file 1 : Table S4. Arginine–proline metabolism and its connections to urea cycle is depicted in Additional file 1 : Fig. S10.

As observed in MPS I patients [26] the arginine metabolism is the most altered pathway, aspartic acid is highly elevated in MPS IIIA and IIIB, significantly elevated in IIIC and shows a rising tendency in IIID. These metabolisms (arginine–proline, urea cycle, aspartic acid) have been reported to be upregulated along with high autophagic activity upon oxygen and glucose reduction using cultured fibroblasts [32] which is consistent with their involvement in bioenergetic balance. As in other LSDs, arginine metabolism may be challenged in MPS III due to lysosome dysfunction and its subsequent autophagic block [33]. Aspartic-acid contributes to the synthesis of *N*-acetyl-l-aspartate (NAA) and its derivative *N*-acetylaspartylglutamate (NAAG). NAA plays a central role in neuronal osmosis and myelin synthesis whereas NAAG is a key neurotransmitter. NAA and NAAG are highly present in the brain; their synthesis and catabolism take place in the brain and are highly regulated and compartmentalized [34]. This high-level homeostasis is consistent with a key function of these components in the central nervous system. Thus, the impact of NAA metabolism is illustrated by the brain damages associated with the NAA catabolic enzyme called aspartoacylase in Canavan disease, an earlyonset spongiform leukodystrophy [35]. It has also been reported that the NAA signal obtained using magnetic resonance spectroscopy is reduced in metachromatic leukodystrophy,

Krabbe disease and other lysosomal storage diseases [36, 37].

A recent study reported metabolomics profiling in serum from MPS IIIA and MPS IIIB patients. Our results are in

accordance with this study, which showed notable metabolic disturbance of key amino acids indicating profound metabolic pathway remodeling. Interestingly, NAA

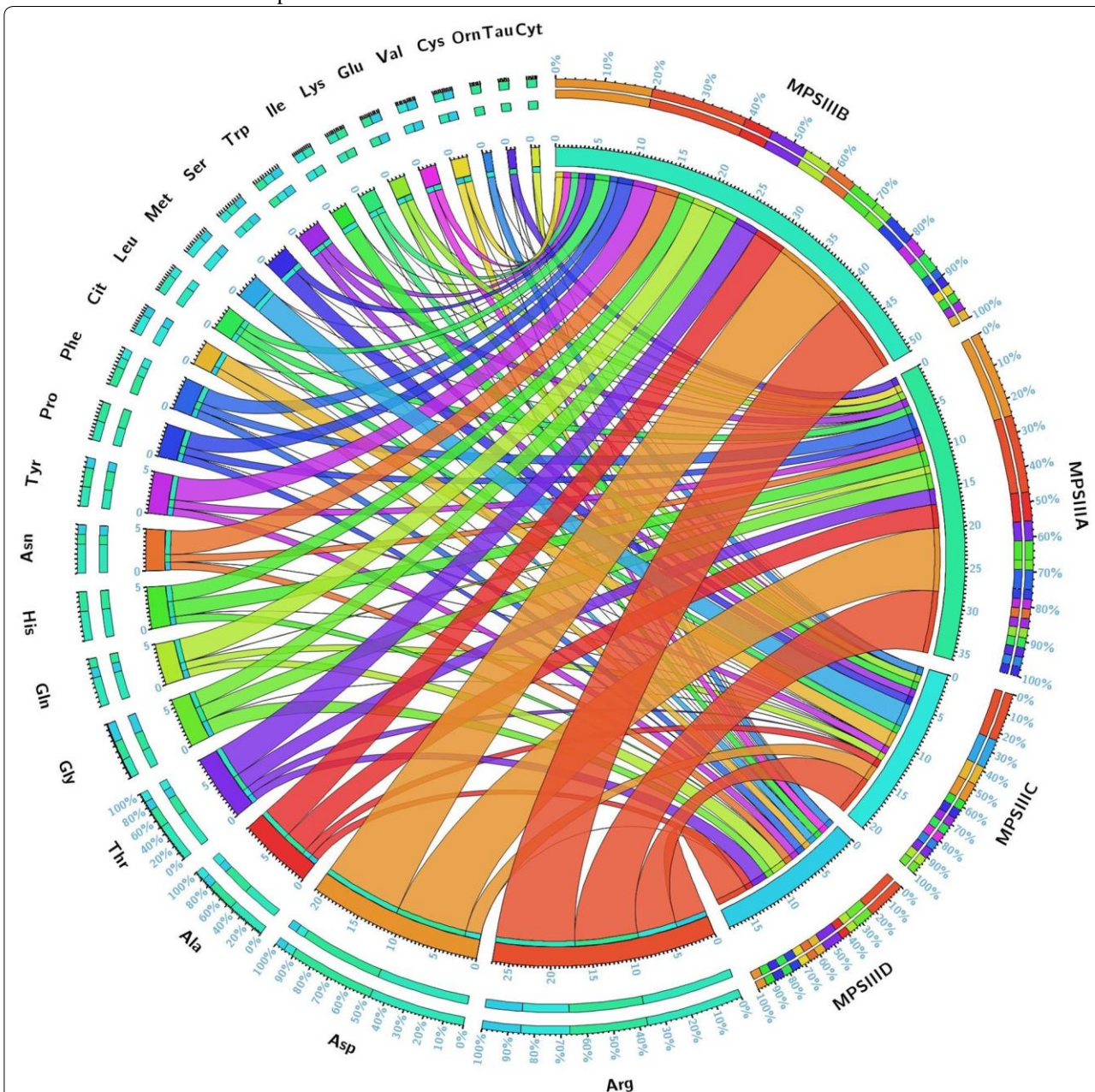
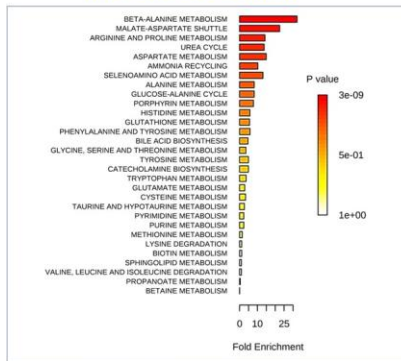


Fig. 4 Circular plot of the 24 amino acids and their related $-\log(p)$ values in the different studies MPS III groups. Segments are color-coded according to amino acids and ribbon size represents $-\log(p)$ values (large ribbons mean low p -values). Corresponding p -values are presented in Table 4

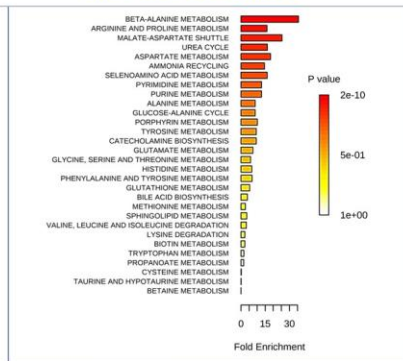
(See figure on next page.)

Fig. 5 Metabolite Set Enrichment Analysis using amino acid concentrations. **a** MPS IIIA vs Control. **b** MPS IIIB vs Control. **c** MPS IIIC vs Control. **d** MPS IIID vs Control. **e** Venn diagram of the significant pathways retrieved from experimental metabolomics data and in silico systems biology approach from Salazar et al. [37]. The diagram shows two common metabolisms : arginine–proline metabolism and urea cycle. Detailed pathway information is given in Additional file 1 : Table S6

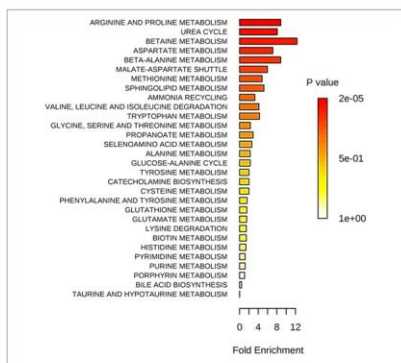
a MPS IIIA vs. Control



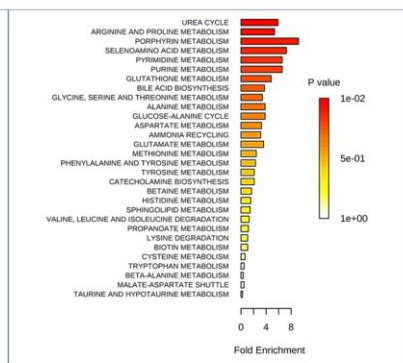
b MPS IIIB vs. Control



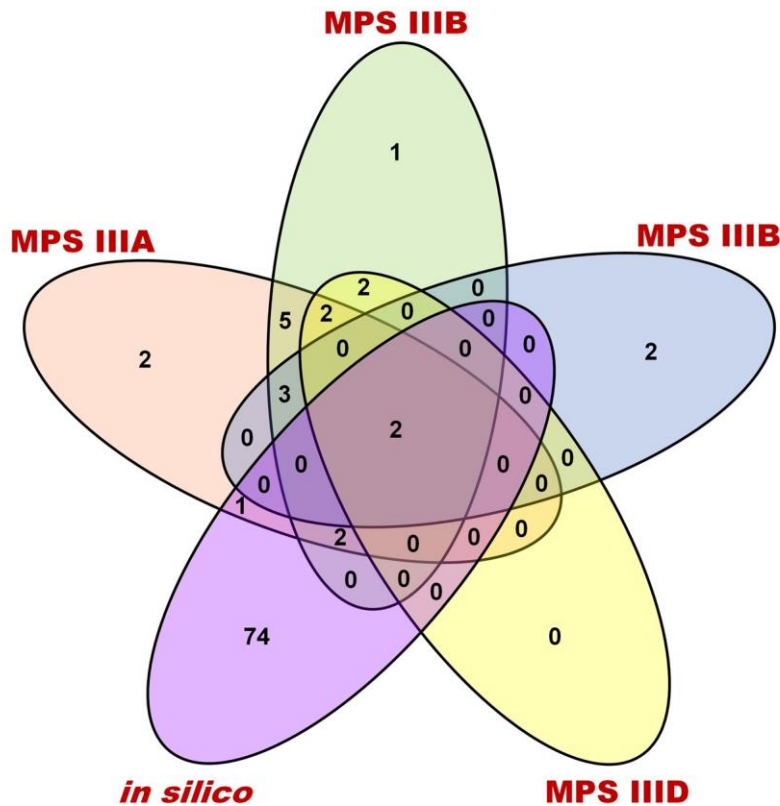
c MPS IIIC vs. Control



d MPS IIID vs. Control



e Venn diagram of experimental vs. in silico analysis



levels were decreased in these patients compared to the control patients [38].

Conclusion

In this study, urine global metabolomics profiling revealed profound metabolic impairments in patients with MPS III. The identification of pathological metabolomics signatures may provide better understanding of the pathophysiological mechanism underlying these diseases and thus allow therapeutic innovation in such rare conditions.

Additional file

[Additional file 1.](#) Detailed analytical protocols and data modeling.

Abbreviations

IEM : inborn errors of metabolism; LSD : lysosomal storage diseases; MPS : mucopolysaccharidoses; GAGs : glycosaminoglycans; MPS III : mucopolysaccharidosis type III; UPLC-IM-MS : ultraperformance liquid chromatography-ion mobility mass spectrometry; CCS : collision cross section; ERT : enzyme replacement therapy; QC : quality control; HS : heparan sulfate; DS : dermatan sulfate; KS : keratan sulfate; ROC : receiver operating characteristic; FDR : false discovery rate; PCA : principal component analysis; OPLS-DA : orthogonal partial leastsquares-discriminant analysis; VIP : variable influence in projection; AUC : area under curve.

Authors' contributions

AT, work conception and design, data acquisition, data analysis, figures drawing and interpretation, paper drafting and final approval; LAD and SM, data interpretation, patient samples and paper revising; ISA, data acquisition and paper revising; MP, BH, JA and FZ, patient samples and paper revising; BG, paper revising and final approval; CA, work conception, data acquisition and interpretation, paper revising and final approval; SB, work conception and design, data analysis and interpretation, paper drafting, revising and final approval. All authors read and approved the final manuscript.

Author details

¹ Department of Metabolic Biochemistry, Rouen University Hospital, 76000 Rouen Cedex, France. ² Normandie Univ, UNIROUEN, CHU Rouen, INSERM U1245, 76000 Rouen, France. ³ Normandie Univ, UNIROUEN, INSA Rouen, CNRS, COBRA, 76000 Rouen, France. ⁴ Department of Neonatal Pediatrics, Intensive Care and Neuropediatrics, Rouen University Hospital, 76031 Rouen, France. ⁵ Department of Pediatric Neurology, Reference Center of Lysosomal Diseases, Trousseau Hospital, APHP and Sorbonne Université, GRC No 19, Pathologies Congénitales du Cervelet-LeucoDystrophies, AP-HP, Hôpital Armand Trousseau, 75012 Paris, France. ⁶ Service de Biochimie et Biologie Moléculaire Grand Est, Unité des Maladies Hérititaires du Métabolisme et Dépistage Néonatal, Centre de Biologie et de Pathologie Est, CHU de Lyon, Lyon, France. ⁷ INSERM U1088, Laboratoire de Biochimie Métabolique, Centre de Biologie Humaine, CHU Sud, 80054 Amiens Cedex, France. ⁸ Laboratoire de Biochimie et Biologie Moléculaire, Université de Lille et Pôle de Biologie Pathologie Génétique du CHRU de Lille, 59000 Lille, France.

Acknowledgements

The authors are grateful to Carine Pilon, Thomas Plichet and Tony Pereira (Rouen University Hospital, France) for their technical assistance.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

All authors have given consent for publication.

Ethics approval and consent to participate

The study was approved by the Institutional Ethics Committee Research (Ethics Board of Rouen University Hospital-CERNI E2016-21) and was performed in accordance with the Declaration of Helsinki.

Funding

The authors gratefully acknowledge the Region Haute-Normandie, the European Regional Development Fund (ERDF 31708), Sanofi-Genzyme, Shire and Vaincre les Maladies Lysosomales for financial support. None of the funders of this research was involved in the design of the study and collection, analysis, and interpretation of data or in writing the manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received : 4 June 2018 Accepted : 30 August 2018

Published online: 04 September 2018

References

1. Tebani A, Afonso C, Marret S, Bekri S. Omics-based strategies in precision medicine : toward a paradigm shift in inborn errors of metabolism investigations. *Int J Mol Sci.* 2016;17 :1555.
2. Nicholson JK, Lindon JC, Holmes E. 'Metabonomics' : understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica.* 1999;29 :1181-9.
3. Bekri S. The role of metabolomics in precision medicine. *Exp Rev Prec Med Drug Dev.* 2016;1 :517-32.
4. Benton HP, Want E, Keun HC, Amberg A, Plumb RS, Goldfain-Blanc F, Walther B, Reily MD, Lindon JC, Holmes E, et al. Intra- and interlaboratory reproducibility of ultra performance liquid chromatography-time-offlight mass spectrometry for urinary metabolic profiling. *Anal Chem.* 2012;84 :2424-32.
5. Tebani A, Abily-Donval L, Afonso C, Marret S, Bekri S. Clinical metabolomics : the new metabolic window for inborn errors of metabolism investigations in the post-genomic era. *Int J Mol Sci.* 2016;17 :1167.
6. Ballabio A, Gieselmann V. Lysosomal disorders : from storage to cellular damage. *Biochim Biophys Acta.* 2009;1793 :684-96.
7. Wraith JE. The mucopolysaccharidoses : a clinical review and guide to management. *Arch Dis Child.* 1995;72 :263-7.
8. Neufeld EF. The mucopolysaccharidoses. In : Scriver CBA, Sly W, Vaele D, editors. *The metabolic and molecular basis of inherited disease.* New York : Mc Graw-Hill; 2001. p. 3421-52.
9. Baehner F, Schmiedeskamp C, Krummenauer F, Miebach E, Bajbouj M, Whybra C, Kohlschütter A, Kampmann C, Beck M. Cumulative incidence rates of the mucopolysaccharidoses in Germany. *J Inher Metab Dis.* 2005;28 :1011-7.

10. Parenti G, Andria G, Ballabio A. Lysosomal storage diseases : from pathophysiology to therapy. *Annu Rev Med*. 2015;66 :471–86.
11. Elizabeth N, Joseph M. *The mucopolysaccharidoses*. New York : McGrawHill; 2001.
12. Delgadillo V, O'Callaghan Mdel M, Gort L, Coll MJ, Pineda M. Natural history of Sanfilippo syndrome in Spain. *Orphanet J Rare Dis*. 2013;8 :189.
13. Buhrman D, Thakkar K, Poe M, Escolar ML. Natural history of Sanfilippo syndrome type A. *J Inherit Metab Dis*. 2014;37 :431–7.
14. Wijburg FA, Wegrzyn G, Burton BK, Tylki-Szymanska A. Mucopolysaccharidosis type III (Sanfilippo syndrome) and misdiagnosis of idiopathic developmental delay, attention deficit/hyperactivity disorder or autism spectrum disorder. *Acta Paediatr*. 2013;102 :462–70.
15. Khan SA, Peracha H, Ballhausen D, Wiesbauer A, Rohrbach M, Gautschi M, Mason RW, Giugliani R, Suzuki Y, Orii KE, et al. Epidemiology of mucopolysaccharidoses. *Mol Genet Metab*. 2017;121 :227–40.
16. Meyer A, Kossow K, Gal A, Muhlhause C, Ullrich K, Braulke T, Muschol N. Scoring evaluation of the natural course of mucopolysaccharidosis type IIIA (Sanfilippo syndrome type A). *Pediatrics*. 2007;120 :e1255–61.
17. Shapiro EG, Nestrail I, Delaney KA, Rudser K, Kovac V, Nair N, Richard CW, Haslett P, Whitley CB. A prospective natural history study of mucopolysaccharidosis type IIIA. *J Pediatrics*. 2016;170(278–287) :e274.
18. Valstar MJ, Bruggenwirth HT, Olmer R, Wevers RA, Verheijen FW, Poorthuis BJ, Halley DJ, Wijburg FA. Mucopolysaccharidosis type IIIB may predominantly present with an attenuated clinical phenotype. *J Inherit Metab Dis*. 2010;33 :759–67.
19. Ruijter GJ, Valstar MJ, van de Kamp JM, van der Helm RM, Durand S, van Diggelen OP, Wevers RA, Poorthuis BJ, Pshezhetsky AV, Wijburg FA. Clinical and genetic spectrum of Sanfilippo type C (MPS IIIC) disease in The Netherlands. *Mol Genet Metab*. 2008;93 :104–11.
20. Valstar MJ, Bertoli-Avella AM, Wessels MW, Ruijter GJ, de Graaf B, Olmer R, Elfferich P, Neijs S, Kariminejad R, Suheyl Ezgu F, et al. Mucopolysaccharidosis type IIID : 12 new patients and 15 novel mutations. *Hum Mutat*. 2010;31 :E1348–60.
21. Tardieu M, Zerah M, Gougeon ML, Ausseil J, de Bournonville S, Husson B, Zafeiriou D, Parenti G, Bourget P, Poirier B, et al. Intracerebral gene therapy in children with mucopolysaccharidosis type IIIB syndrome : an uncontrolled phase 1/2 clinical trial. *Lancet Neurol*. 2017;16 :712–20.
22. Lange MC, Teive HA, Troiano AR, Bitencourt M, Funke VA, Setubal DC, Zanis Neto J, Medeiros CR, Werneck LC, Pasquini R, Bonfim CM. Bone marrow transplantation in patients with storage diseases : a developing country experience. *Arq Neuropsiquiatr*. 2006;64 :1–4.
23. Moskot M, Jakobkiewicz-Banecka J, Kloska A, Smolinska E, Mozolewski P, Malinowska M, Rychlowski M, Banecki B, Wegrzyn G, Gabig-Ciminska M. Modulation of expression of genes involved in glycosaminoglycan metabolism and lysosome biogenesis by flavonoids. *Sci Rep*. 2015;5 :9378.
24. Roberts AL, Fletcher JM, Moore L, Byers S. Trans-generational exposure to low levels of rhodamine B does not adversely affect litter size or liver function in murine mucopolysaccharidosis type IIIA. *Mol Genet Metab*. 2010;101 :208–13.
25. Savas PS, Hemsley KM, Hopwood JJ. Intracerebral injection of sulfamidase delays neuropathology in murine MPS-IIIa. *Mol Genet Metab*. 2004;82 :273–85.
26. Tebani A, Schmitz-Afonso I, Abily-Donval L, Heron B, Piraud M, Ausseil J, Brassier A, De Lonlay P, Zerimech F, Vaz FM, et al. Urinary metabolic phenotyping of mucopolysaccharidosis type I combining untargeted and targeted strategies with data modeling. *Clinica Chimica Acta*. 2017;475 :7–14.
27. Tebani A, Schmitz-Afonso I, Rutledge DN, Gonzalez BJ, Bekri S, Afonso C. Optimization of a liquid chromatography ion mobility-mass spectrometry method for untargeted metabolomics using experimental design and multivariate data analysis. *Anal Chim Acta*. 2016;913 :55–62.
28. Shen X, Gong X, Cai Y, Guo Y, Tu J, Li H, Zhang T, Wang J, Xue F, Zhu Z-J. Normalization and integration of large-scale metabolomics data using support vector regression. *Metabolomics*. 2016;12 :89.
29. Li S, Park Y, Duraisingham S, Strobel FH, Khan N, Soltow QA, Jones DP, Pulendran B. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol*. 2013;9 :e1003123.
30. Xia J, Sinelnikov IV, Han B, Wishart DS. MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Res*. 2015;43 :W25.
31. Salazar DA, Rodriguez-Lopez A, Herrero A, Barbosa H, Herrera J, Ardila A, Barreto GE, Gonzalez J, Almeciga-Diaz CJ. Systems biology study of mucopolysaccharidosis using a human metabolic reconstruction network. *Mol Genet Metab*. 2016;117 :129–39.
32. Weckmann K, Diefenthaler P, Baeken MW, Yusifli K, Turck CW, Asara JM, Behl C, Hajieva P. Metabolomics profiling reveals differential adaptation of major energy metabolism pathways associated with autophagy upon oxygen and glucose reduction. *Sci Rep*. 2018;8 :2337.
33. Pshezhetsky AV. Lysosomal storage of heparan sulfate causes mitochondrial defects, altered autophagy, and neuronal death in the mouse model of mucopolysaccharidosis III type C. *Autophagy*. 2016;12 :1059–60.
34. Baslow MH. Functions of *N*-acetyl-*D*-aspartate and *N*-acetyl-*D*-aspartylglutamate in the vertebrate brain : role in glial cell-specific signaling. *J Neurochem*. 2000;75 :453–9.
35. Matalon R, Michals-Matalon K. Molecular basis of Canavan disease. *Eur J Paediatr Neurol*. 1998;2 :69–76.
36. Brockmann K, Dechent P, Wilken B, Rusch O, Frahm J, Hanefeld F. Proton MRS profile of cerebral metabolic abnormalities in Krabbe disease. *Neurology*. 2003;60 :819–25.
37. Dali C, Hanson LG, Barton NW, Fogh J, Nair N, Lund AM. Brain *N*-acetylaspartate levels correlate with motor function in metachromatic leukodystrophy. *Neurology*. 2010;75 :1896–903.
38. Fu H, Meadows AS, Pineda RJ, Mohny RP, Stirdivant S, McCarty DM. Serum global metabolomics profiling reveals profound metabolic impairments in patients with MPS IIIA and MPS IIIB. *Metab Brain Dis*. 2017;32 :1403–15.

Ready to submit your research ? Choose BMC and benefit from :

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research : over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions





Article

Analysis of Mucopolysaccharidosis Type VI through Integrative Functional Metabolomics

Abdellah Tebani ^{1,2,3}, Lenaig Abily-Donval ^{2,4}, Isabelle Schmitz-Afonso ³, Monique Piraud ⁵, Jérôme Ausseil ⁶, Farid Zerimech ⁷, Carine Pilon ¹, Tony Pereira ⁸, Stéphane Marret ^{2,4}, Carlos Afonso ³ and Soumeya Bekri ^{1,2,*}

¹ Department of Metabolic Biochemistry, Rouen University Hospital, 76000 Rouen, France; abdellah.tebani@yahoo.com (A.T.); Carine.Pilon@chu-rouen.fr (C.P.)

² Normandie University, UNIROUEN, CHU Rouen, INSERM U1245, 76000 Rouen, France; lenaig.abily-donval@chu-rouen.fr (L.A.-D.); stephane.marret@chu-rouen.fr (S.M.)

³ Normandie Univ, UNIROUEN, INSA Rouen, CNRS, COBRA, 76000 Rouen, France; isabelle.schmitz-afonso@univ-rouen.fr (I.S.-A.); carlos.afonso@univ-rouen.fr (C.A.)

⁴ Department of Neonatal Pediatrics, Intensive Care and Neuropediatrics, Rouen University Hospital, 76031 Rouen, France

⁵ Service de Biochimie et Biologie Moléculaire Grand Est, Unité des Maladies Héréditaires du Métabolisme et Dépistage Néonatal, Centre de Biologie et de Pathologie Est, Hospices Civils de Lyon, 69002 Lyon, France; monique.piraud@chu-lyon.fr

⁶ INSERM U1088, Laboratoire de Biochimie Métabolique, Centre de Biologie Humaine, CHU Sud, 80054 Amiens CEDEX, France; jerome.ausseil@chu-amiens.fr

⁷ Laboratoire de Biochimie et Biologie Moléculaire, Université de Lille et Pôle de Biologie Pathologie Génétique du CHRU de Lille, 59000 Lille, France; farid.zerimech@chru-lille.fr

⁸ Department of Pharmacology, Rouen University Hospital, 76000 Rouen, France; Tony.Pereira@chu-rouen.fr

* Correspondence : soumeya.bekri@chu-rouen.fr; Tel. : +33-232-888-124



Received : 17 December 2018; Accepted : 18 January 2019; Published : 21 January 2019

Abstract : Metabolic phenotyping is poised as a powerful and promising tool for biomarker discovery in inherited metabolic diseases. However, few studies applied this approach to mucopolysaccharidoses (MPS). Thus, this innovative functional approach may unveil comprehensive impairments in MPS biology. This study explores mucopolysaccharidosis VI (MPS VI) or Maroteaux–Lamy syndrome (OMIM #253200) which is an autosomal recessive lysosomal storage disease caused by the deficiency of arylsulfatase B enzyme. Urine samples were collected from 16 MPS VI patients and 66 healthy control individuals. Untargeted metabolomics analysis was applied using ultra-high-performance liquid chromatography combined with ion mobility and high-resolution mass spectrometry. Furthermore, dermatan sulfate, amino acids, carnitine, and acylcarnitine profiles were quantified using liquid chromatography coupled to tandem mass spectrometry. Univariate analysis and multivariate data modeling were used for integrative analysis and discriminant metabolites selection. Pathway analysis was done to unveil impaired metabolism. The study revealed significant differential biochemical patterns using multivariate data modeling. Pathway analysis revealed that several major amino acid pathways were dysregulated in MPS VI. Integrative analysis of targeted and untargeted metabolomics data with *in silico* results yielded arginine-proline, histidine, and glutathione metabolism being the most affected. This study is one of the first metabolic phenotyping studies of MPS VI. The findings might shed light on molecular understanding of MPS pathophysiology to develop further MPS studies to enhance diagnosis and treatments of this rare condition.

Keywords : metabolomics; inherited metabolic diseases; lysosomal storage diseases; mucopolysaccharidosis type VI; Maroteaux–Lamy syndrome; mass spectrometry

1. Introduction

The metabolome can be defined as the complement of all metabolites contained in a given biological system. It embodies the most functional readout of the information embedded in the genotype. Therefore, metabolite changes are usually more suitable to describe the biochemical state of a biological system given their closeness to the phenotype. Given this precious contextual biological information, the metabolome is very appealing to assess pathophysiological states [1,2]. To do so, metabolomics is an “omics” technology that allows the molecular and biochemical characterizations of the metabolome [3,4]. It also describes, holistically, its changes related to both genetic and environmental factors. Inherited metabolic diseases (IMDs) are the most obvious disease group that may directly take advantage of the potential of metabolomics given their underlying pathophysiology [5]. For years, mass spectrometry-based targeted metabolomics have been used in the assessment of IMDs such as aminoacidopathies, organic acidurias, and acylcarnitines for fatty acid oxidation disorders [6–10]. Furthermore, MS is now widely implemented in national inborn errors of metabolism (IEM) newborn screening programs worldwide [11]. Indeed, IMDs are a group of rare diseases mainly related to a genetic defect in enzymes, transporters or cofactors involved in metabolic pathways. Hence, a technology that could infer these impairments in a systematic and holistic fashion is very promising to better understand the metabolic pathways involved in the progression of the disease for better diagnosis and treatment [12,13]. Lysosomal storage disorders (LSDs) represent a group of about 50 inherited disorders related to lysosomal protein deficiencies which cause a progressive accumulation of undegraded metabolites within the lysosome. This metabolite storage leads to various organ failures and premature death [14]. Mucopolysaccharidoses (MPS) is an LSD subgroup due to glycosaminoglycan (GAG) catabolism impairment. Thus, depending on the enzymatic block, one or several GAGs (dermatan sulfate—DS, chondroitin sulfate—CS, heparan sulfate—HS, keratan sulfate—KS, and hyaluronan) may accumulate in lysosomes and the extracellular matrix [15]. The GAG storage causes progressive multiple tissue and organ failures [16]. Eleven known enzyme deficiencies lead to seven distinct forms of MPS [14]. Prenatal symptoms, mainly hydrops fetalis, may be observed in MPS I, MPS IVA, MPS VI, and more frequently in MPS VII [17,18]; however, most MPS patients are asymptomatic after birth. Mucopolysaccharidoses symptoms and severity vary with patients and MPS subtypes. Several MPS treatments are in clinical use or being investigated under clinical trials [19]. Mucopolysaccharidosis VI (MPS VI) or Maroteaux–Lamy syndrome (OMIM #253200) is an autosomal recessive LSD, first described in 1963 [20]. It is caused by the deficiency of the arylsulfatase B (ASB) enzyme also called *N*-acetylgalactosamine 4-sulfatase (E.C.3.1.6.12). This leads to incomplete degradation and accumulation of DS. The arylsulfatase B gene (*ARSB*) is located in chromosome 5 (5q13–5q14) [2].

Mucopolysaccharidoses VI is very rare; its estimated incidence ranges from 0.36 to 1.30 per 100,000 live births [21]. Clinical features, age of onset, and disease progression vary widely in MPS VI patients. A distinction is generally made between slowly and rapidly progressing onsets. However, it is important to note that the phenotype of the disease is rather a clinical continuum. The rapidly progressing phenotype of MPS VI is usually observed before the age of 2 years. The skeleton is generally severely affected with dysostosis multiplex as a typical feature [22]. Musculoskeletal abnormalities can also lead to short stature and low body weight. Other common findings in these patients are coarse facial features [23]. In the absence of treatment, patients with rapidly progressing phenotype generally die from cardiopulmonary disease, infection, or surgical complications with a life expectancy below 20 years. The slowly progressing phenotype of MPS VI is associated with a slower clinical course. The patients show normal or only mildly coarsened facial features, a slightly reduced to normal body height, and less prominent skeletal dysplasia [22,24]. In an advanced stage, patients often develop disability symptoms requiring surgery which lead, ultimately, to a reduction in lifespan [25]. Most MPS VI patients have a poor quality of life due to the high morbidity (impairments of vision, hearing, mobility, and functional capacity along with frequent surgeries), which requires constant and tight management of the disease. Like most rare diseases, for a long time the treatment of MPS VI was limited to palliative care. Recently, targeted therapies have been developed mainly as enzyme replacement therapy Naglazyme™ (Galsulfase, BioMarin) [26,27] and hematopoietic stem cell transplantation (HSCT) [28]. The only available biomarker for MPS VI diagnosis is urinary dermatan sulfate [22]. The aim of this study is to apply both

targeted and untargeted metabolic profiling on MPS VI patients, compared to controls, to explore metabolic remodeling in this disease.

2. Results

2.1. Untargeted Analysis

The heatmap in Figure 1A represents the top 100 features ranked by *t*-test ($p < 0.05$ cut-off and FDR 5%) suggesting differential patterns between the study groups and two main clusters are clearly individualized. To explore further, we performed a principal component analysis (PCA) as a dimension reduction tool. PCA is an unsupervised method which is very useful to track clustering trends and identify potential outliers. The PCA score plot also revealed a clustering trend between controls and patients' samples (Figure 1B). To unveil metabolic features underlying the group separation, supervised methods are more suitable given their predictive characteristics. The OPLS-DA was applied for this classification purpose. Hence, samples were labeled according to their respective groups, control and MPS VI. The final OPLS-DA model had an $R^2 = 0.99$ and $Q^2 = 0.60$. The OPLS-DA score plot (Figure 1C) shows that the groups are clearly separated according to their metabolic features. This model is validated by cross-validation using both CV-ANOVA (p -value = 2.14×10^{-6}) and the permutation test (999 permutations). Model validation details are presented in Supplementary Materials Figure S1. Variable importance in projection (VIP) yielded by the built model was used as a variable selection parameter. Based on 1 as a cut-off value, 172 features out of the 854 were selected for the MPS VI vs. the control model. The list of variables was refined by retaining only the most discriminant ones and their putative annotation. Some discriminant features are shown in Table 1 along with their respective annotation accuracy and statistical metrics whereas boxplots and correlation plot with dermatan sulfate are presented in Figure S2. To go further in biological inference, we performed pathway analysis using Mummichog software which is based on mapping significant pathways related to the significantly disturbed metabolite variation. The affected metabolism pathways are shown in Table 2. Interestingly, a series of amino acid and fatty acids pathways are impaired.

2.2. Targeted Analysis

For targeted analysis, twenty-four amino acids, free carnitine, and twelve acylcarnitines were assessed in all urine samples. Table S1 shows urine concentrations, and boxplots of normalized amino acid concentrations are presented in Figure S3. The statistical analysis results are listed in Table 3 and Figure 2. This differential analysis yielded sixteen metabolites that have shown significant differences highlighting aspartic acid and alanine as the most significant metabolites. A hierarchical clustering analysis has been applied to track sample clustering with similar metabolic profiles. All assessed metabolites, ranked by *t*-test, are presented in the heatmap in Figure 1D. The heatmap shows a clear pattern of sample grouping to respective groups which is also highlighted in the dendrogram structure by its two longest branches; one related mainly to free carnitine and acylcarnitines and the other cluster mainly to amino acids. A correlation analysis of targeted metabolites has been done and the heatmap is depicted in Figure 1E. Both figures show a clear clustering of variables that have high correlation. One includes carnitine and acylcarnitines, and the other amino acids. To assess the diagnostic performance of the different targeted metabolites, univariate ROC curve analyses were used (Table 3). The top metabolites with a high AUC were : aspartic acid (0.85), valine (0.83) and glutamic acid (0.79). A comparison of different classifiers combining the main significant metabolites was performed using both support vector machine (SVM) and PLS-DA models with three components each (Figure S4). This combination did not improve the predictive performances of the classifiers. We also performed pathway analysis using these targeted analysis data which yielded as the main impaired metabolisms : aspartate metabolism, glutamate metabolism, valine, leucine and isoleucine degradation, tyrosine metabolism, ammonia recycling and purine metabolism. Salazar et al. [29] generated an in-silico MPS VI model by silencing the *ARSB* gene. The metabolic effects have been analyzed and the altered pathways have been reported [29]. In an attempt to get a deeper insight in our data, we compared the altered pathways identified through our model using experimental metabolomics data to those of Salazar et al.'s [29] in-silico model. As illustrated in detail (Table

S2) and Venn diagram (Figure 3B), this comparison showed three main common metabolic pathways : glutathione, histidine and arginine-proline metabolism. The overall results are shown in Figure 3.

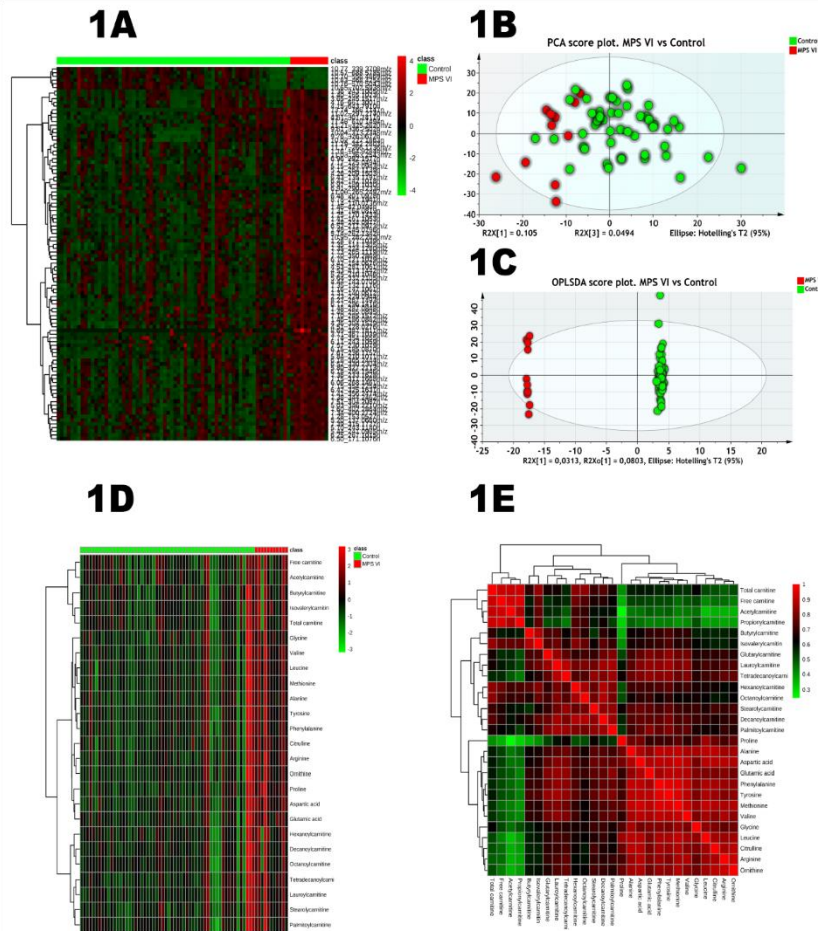


Figure 1. (A) Hierarchical cluster analysis and heat map visualization of the top 100 variables (x-axis) ranked by *t*-test. The urine sample classes are represented along the x-axis. The color code was used to represent log-scaled intensities of features between -4 (green) and $+4$ (red), showing the features' relative abundance according to the groups. (B) PCA scores' plot of the normalized dataset. The two groups are represented by different colors. A clear separation is observed between the groups with a clear clustering of the control group. (C) OPLSDA scores' plot ($R^2 = 0.99$, $Q^2 = 0.60$) shows a clear separation between the MPS VI and controls. Detailed model characteristics and validation are given in the Supplementary Materials. (D) Heat map representing the clustering of amino acids, free and total carnitine along with acylcarnitines across the two groups : Mucopolysaccharidosis (MPS) VI and controls. Columns represent individual samples and rows refer to amino acid. Shades of red or green represent elevation or decrease, respectively, of an amino acid. (E) Spearman rank-order correlation matrix assessed targeted metabolites based on their concentration profiles across all samples. Shades of green or red represent low-to-high correlation coefficient between metabolites.

Table 1. Some discriminant putatively annotated features and related statistical metrics.

HMDB	Putative Annotation	Formula	<i>M</i>	<i>m/z</i>	Adduct	$\Delta m/z$ (ppm)	<i>t_R</i> (min)	<i>t_D</i> (ms)	CCS (Å ²)	%RSD	VIP	FDR	AUC
HMDB0003464	4-guanidinobutanoic acid	C ₅ H ₁₁ N ₃ O ₂	145.0851	146.0932	M + H	5	1.48	1.89	124.4	4.60	0.97	5.18×10^{-4}	0.88
HMDB0001276	<i>N</i> -acetylspermidine	C ₉ H ₂₁ N ₃ O	187.1685	188.1774	M + H	9	1.25	2.38	139.3	8.58	0.53	1.74×10^{-2}	0.83
HMDB00062	Carnitine	C ₈ H ₁₈ N ₄ O ₂	202.1430	203.1518	M + can + H	0.48	1.41	2.43	140.4	9.97	0.70	3.52×10^{-3}	0.85
HMDB0015444	Phenylalaninylalanine	C ₁₂ H ₁₆ N ₂ O ₃	236.1161	237.1225	M + H	4	7.67	2.7	147.8	10.51	1.55	1.95×10^{-4}	0.94
HMDB0002012	Ubiquinone-1	C ₁₄ H ₁₈ O ₄	250.1205	251.1291	M + H	5	7.17	2.86	152.4	5.69	0.16	2.74×10^{-2}	0.80
HMDB0000145	Estrone	C ₁₈ H ₂₂ O ₂	270.1620	271.1675	M + H	6	6.50	3.19	161.5	4.26	0.27	2.83×10^{-2}	0.79

M : monoisotopic mass, ppm : parts per million; *t_R* : retention time; *t_D* : drift time; CCS : cross collision section; VIP : variable importance in projection; RSD : relative standard deviation; FDR; false discovery rate; AUC : area under curve.

Table 2. Significantly dysregulated pathways.

Pathway	Overlap Size	p-Value (FDR = 5%)
Vitamin B9 (folate) metabolism	5	2.87×10^{-4}
Glycine, serine, alanine and threonine metabolism	7	3.36×10^{-4}
Alanine and Aspartate metabolism	4	4.68×10^{-4}
Histidine metabolism	4	1.29×10^{-3}
Vitamin E metabolism	5	2.21×10^{-3}
Carnitine shuttle	5	2.21×10^{-3}
Glycosphingolipid metabolism	3	3.61×10^{-3}
Vitamin B3 (nicotinate and nicotinamide) metabolism	3	3.61×10^{-3}
Selenoamino acid metabolism	2	4.15×10^{-03}
Glutathione Metabolism	2	4.15×10^{-3}
CoA Catabolism	2	4.15×10^{-3}
Electron transport chain	2	4.15×10^{-3}
Vitamin B5–CoA biosynthesis from pantothenate	2	4.15×10^{-3}
Methionine and cysteine metabolism	6	4.66×10^{-3}
Aspartate and asparagine metabolism	7	8.25×10^{-3}
Purine metabolism	5	1.01×10^{-2}
Arginine and proline metabolism	4	1.20×10^{-2}
Lysine metabolism	4	1.70×10^{-2}
Linoleate metabolism	4	1.70×10^{-2}
Aminosugar metabolism	3	2.29×10^{-2}
Porphyrin metabolism	3	2.29×10^{-2}
Pyruvate metabolism	2	2.63×10^{-2}

FDR : False discovery rate.

Table 3. The normalized concentrations of dermatan sulfate, free amino acids, carnitine (total and free), and acylcarnitines in urine samples of the MPS VI and control groups. ($\mu\text{M}/\text{mM}$ creatinine).

Control vs. MPS VI				
	AUC	p-Value (FDR)	Fold Change	Effect in MPS VI
Dermatan sulfate	0.90	1.23×10^{-3}	10.0	Increased
Aspartic acid	0.85	6.41×10^{-3}	1.61	Increased
Valine	0.83	6.41×10^{-3}	1.74	Increased
Glutamic acid	0.79	2.50×10^{-2}	1.42	Increased
Leucine	0.79	2.50×10^{-2}	1.44	Increased
Tetradecanoylcarnitine	0.78	2.50×10^{-2}	1.36	Increased
Alanine	0.75	2.58×10^{-2}	1.32	Increased
Lauroylcarnitine	0.75	2.58×10^{-2}	1.18	Increased
Methionine	0.77	2.58×10^{-2}	1.28	Increased
Phenylalanine	0.77	2.58×10^{-2}	1.28	Increased
Proline	0.79	2.58×10^{-2}	1.48	Increased
Stearoylcarnitine	0.76	2.58×10^{-2}	1.07	Increased
Tyrosine	0.76	2.58×10^{-2}	1.30	Increased
Isovalerylcarnitine	0.71	2.70×10^{-2}	1.42	Increased
Citrulline	0.79	3.11×10^{-2}	1.26	Increased
Hexanoylcarnitine	0.71	3.32×10^{-2}	1.20	Increased
Arginine	0.80	3.95×10^{-2}	1.52	Increased
Palmitoylcarnitine	0.68	6.31×10^{-2}	0.00	/
Butyrylcarnitine	0.68	9.16×10^{-2}	1.15	Increased
Free carnitine	0.67	9.46×10^{-2}	1.27	Increased
Decanoylcarnitine	0.67	1.09×10^{-1}	0.90	Decreased
Ornithine	0.75	1.17×10^{-1}	1.05	Increased
Glycine	0.70	1.20×10^{-1}	0.95	Decreased
Glutaryl carnitine	0.66	1.59×10^{-1}	0.95	Decreased
Octanoylcarnitine	0.65	1.63×10^{-1}	0.79	Decreased
Acetylcarnitine	0.62	1.82×10^{-1}	1.04	Increased
Total carnitine	0.62	2.02×10^{-1}	0.89	Decreased

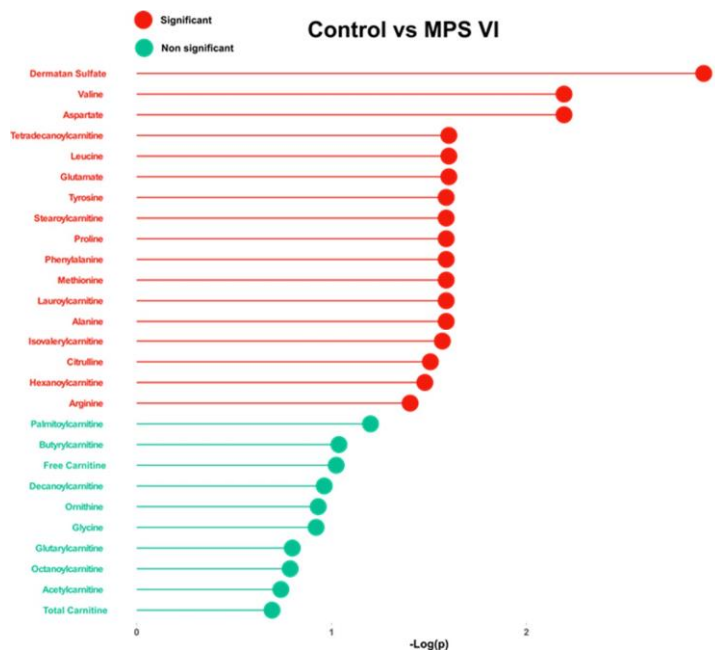


Figure 2. Bar plot showing of dermatan sulfate, total carnitine, free carnitine, and the 13 amino acids and their related $-\log(p)$ values between MPS VI and controls. Cut-off is set to FDR = 0.05. Corresponding p -values are presented in Table 3.

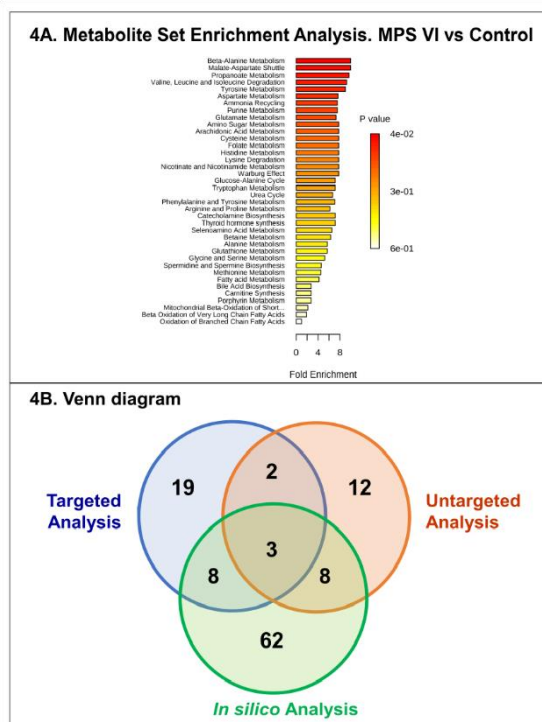


Figure 3. (A) Pathway analysis using the assessed free carnitine, acylcarnitines, and the 24 amino-acid concentrations. **(B)** Venn diagram of the significant pathways retrieved from untargeted, targeted approaches, and in silico systems biology approach from Salazar DA et al. [25]. The diagram shows three common metabolisms : glutathione metabolism, histidine metabolism, arginine and proline metabolism. Detailed pathway information is given in the Supplementary Material Table S4.

3. Discussion

This study aimed to investigate MPS VI urine metabolic patterns to unveil the putative biological differences compared to control individuals. The adopted strategy, untargeted metabolomics analysis, enabled us to build a predictive model showing two clusters corresponding to patient and control groups. These differential metabolic patterns are underlined

by the major metabolic remodeling in MPS VI group with amino acid-related metabolism being the most affected. The second confirmatory step consisted in quantifying amino acids and acylcarnitines using a targeted approach. Therefore, we performed an integrative analysis between the pathway analysis results of this present study and those of Salazar et al. [29] based on an in-silico strategy by silencing *ARSB* gene. As shown in Figure 3 and Table S5, three main metabolisms are common : arginine-proline, histidine, and glutathione metabolism.

It is now admitted that the clinical manifestations of lysosomal storage diseases are only partly explained by the adverse effects of the lysosomal non-degraded material and that other common pathogenic effects underlie LSDs [30]. In the case of MPSs, the GAG storage affects key cellular processes, and major signaling pathways such as Sonic hedgehog (Shh) and Wnt/ β -catenin signaling. These pathways are involved in both morphogenetic processes and tissue homeostasis [31–33]. This GAG-related effect is worsened by pathophysiological disruption common to all LSDs such as impairment of autophagy and mitophagy processes responsible for cell damages and death. The subsequent increase of altered mitochondria and reactive oxygen species may represent the main pathogenic mechanism of LSD diseases [34]. In this context, it is instructive to note that the main altered pathways characterized in this study, arginine-proline, histidine and glutathione metabolisms, are involved in regulating autophagy and protecting from oxidative stress. The mechanistic target of rapamycin complex 1 (mTORC1) is a key nutrient/energy sensor; the active form of mTORC1 promotes anabolic biosynthetic pathways and inhibits catabolic processes mainly autophagy. The activation/inactivation of mTORC1 is dependent on its intracellular localization. It has been well documented that amino acids are involved in mTORC1 regulation/localization and this amino acid-sensing localization occurs both in the lysosome and the cytoplasm. Arginine is the main amino acid contributor to this regulation. The alteration of arginine pathway in LSDs may represent a major cause of autophagy deregulation. Of note, metabolomics analysis of urine MPS I and MPSIII patients versus controls have also demonstrated a major alteration of arginine pathway [35,36]. Histidine and glutathione pathways are linked to oxidative stress process and have both protective effects toward reactive oxygen species (ROS). Histidine is the precursor of carnosine (β -alanyl-L-histidine). The latter is likely to exert a protective effect in oxidative diseases. Histidine and carnosine are active through the imidazole cycle of histidine molecule that helps in scavenging ROS [37,38]. Autophagy dysfunction exacerbates cellular stress generation and in turn, the increased ROS production triggers the alteration of both autophagy and apoptosis processes. Accordingly, recent studies unveiled the emerging common cellular pathways involved upon oxidative stress that link autophagy and apoptosis [39]. Mitochondria dysfunction and the subsequent programmed cell death disruptions may represent the most significant cellular deregulation underlying LSD pathogenesis. Indeed, oxidative damage contribution in the LSD pathophysiology in general and in MPSs in particular has recently been documented [34,40]. These data suggest metabolic remodeling and provide a functional snapshot profile of metabolite abundances. More mechanistic studies are needed for deeper exploration of these disturbances. Furthermore, this study presents some limits regarding : (i) the small number of patients mainly related to the rarity of the disease, and (ii) age distribution heterogeneity related to the stringent conditions for control of pediatric cohorts; this has been addressed using a covariate correction method (lrima R-package). Given the broader scope of mass spectrometry-based untargeted metabolomics regarding metabolome coverage, the presented results are still exploratory, yet encouraging, and require further validation studies through more mechanistic experiments to go deeper in understanding the role of the unveiled metabolisms and their interactions with dermatan sulfate metabolism.

4. Material and Methods

4.1. Urine Samples

Since the MPS VI diagnosis is mainly based on urine analysis, urine samples have been widely included in expert centers collections. Random morning urine samples were collected from MPS VI patients with confirmed diagnosis by demonstrating marked enzyme deficiency in leucocytes and/or by molecular analysis. Urine samples were collected in four expert centers for inherited metabolic diseases in France. Sixteen MPS VI patients were evaluated : 10 males (age range from 3.1 to 14.7 years, mean age : 6.9 years) and six females (age range from 1.6 to 11.7 years, mean age : 3.9 years). Control urine samples were from 66 healthy subjects, 27 males and 39 females (age range from 5.5 to 70 years,

mean age : 40.8 years). The samples were retrieved from biological collections authorized by the French Ministry of Research. The samples were obtained in accordance with Good Clinical Practices. This project was approved by the Research Ethics Board of Rouen University Hospital (CERNI 27-04-2016 E2016-21).

4.2. Reagents and Chemicals

Acetonitrile was purchased from VWR Chemicals (Fontenay-sous-Bois, France), ultrapure water (18 MX) from Millipore (Molsheim, France) and formic acid from Fluka (Saint Quentin Fallavier, France). The chemicals used were of analytical grade. Leucine Enkephalin (Saint-Quentin Fallavier, Sigma–Aldrich, France) at a concentration of 2 ng/mL (in acetonitrile/water, 50/50) was used as reference for mass measurements. Poly-DL-alanine was prepared in 50 :50 (v/v) water/acetonitrile at 10 mg/L and used for ion mobility cell calibration. MassChrom® Amino Acid and Acylcarnitine kit (Ref. 55000) was purchased from Chromsystems (Gräfelfing Germany). HPLC gradient grade methanol was purchased from VWR Chemicals. Dermatan sulfate calibration standard was from Sigma–Aldrich (Saint-Quentin Fallavier, France).

4.3. Untargeted Metabolic Phenotyping

The protocol used in this study was previously described [35] and Figure 4 presents an overview of the implemented metabolomics workflow.

4.3.1. Sample Handling

The sample handling component was a Waters 2777C sample manager (Waters Corp., Milford, MA, USA) equipped with a 25 µL Hamilton syringe, a 2 µL loop used for full loop injections of prepared sample, and a 2-drawer sample chamber thermo-stated at 4 °C with a constant flow of dry nitrogen gas to prevent the buildup of condensation.

4.3.2. Chromatographic Conditions

The chromatography was performed on a Waters NanoAcquity UPLC module (Saint Quentin en Yvelines, France) upgraded to work with a 1-mm column and composed with a binary solvent manager and column heater/cooler module. Separation was carried out at 45 °C using a 1.0 × 100 mm, Acquity UPLC HSS T3 column (Waters), with a particle size of 1.8 µm, equipped with a 0.2-µm prefilter. Urine was eluted from the LC column using the following linear gradient (curve number 6) : 0–1 min : 99% A; 1–3 min, 99–85% A; 3–6 min, 85–50% A; 6–9 min, 50–0% A; 9–12 min, 100% B, 12–16 min, 99% A for re-equilibration. Solvent A was water and solvent B was acetonitrile, both solvents contained 0.1% formic acid. Sample analysis order has been randomized to avoid potential for confounding critical variables with analytical run order effects.

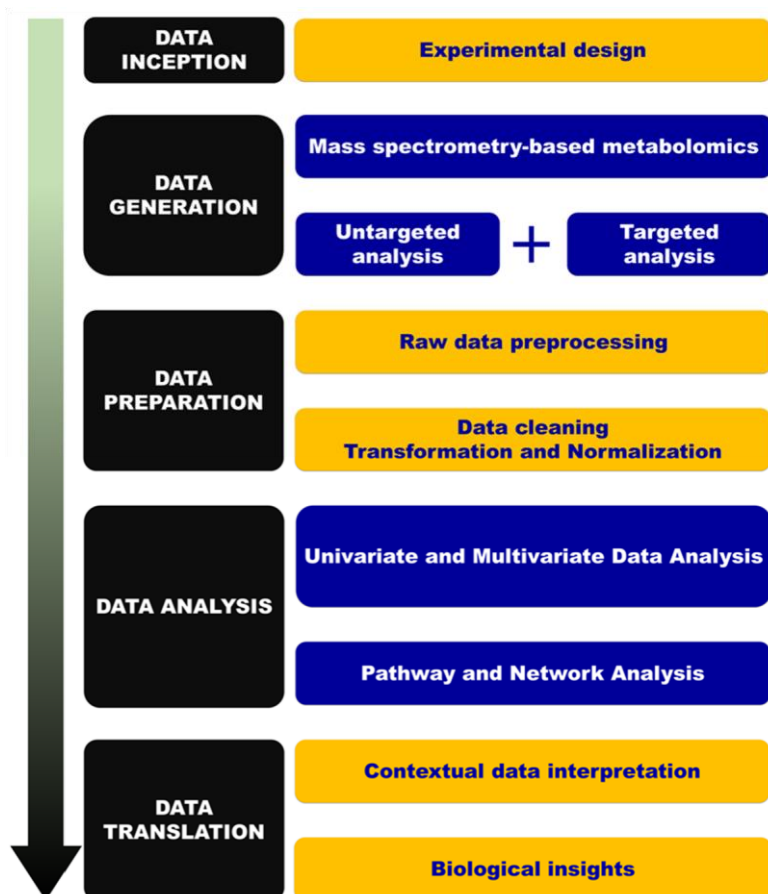


Figure 4. Illustration of the metabolomics workflow spanning from experimental design to pathway analysis and biological interpretation. HMDB : Human Metabolome Database. KEGG : Kyoto Encyclopedia of Genes and Genomes. MetCCS : Metabolite CCS database. MSEA : Metabolite Set Enrichment Analysis. RSD : relative standard deviation.

4.3.3. Ion Mobility and Mass Spectrometry

The U-HPLC system was coupled to a hybrid quadrupole orthogonal time-of-flight (TOF) mass spectrometer (SYNAPT G2 HDMS, Waters MS Technologies, Manchester, UK). The mass spectrometer was operated in positive electrospray ionization mode. A mass range of m/z 50–1200 was used in both modes. The sample cone voltage, extraction cone voltage, source temperature, desolvation temperature, desolvation gas flow, and cone gas flow were optimized and were as follows respectively : 25 V, 5 V, 120 °C, 500 °C, 400 L/h, 50 L/h. Leucine enkephalin was used as the lock mass $[M + H]^+$ at m/z 556.2771. Sodium formate solution was used for external instrument calibration. The Synapt G2 HDMS was equipped with a traveling wave “Triwave™” geometry in which the ion mobility cell (IMS T-wave). The helium cell gas flow, wave height, Trap Bias, and IMS wave delay were set at 180 mL/min, 40 V, 45 V, and 450 μ s, respectively. The TOF analyzer was operated in the V resolution mode with an average mass resolution of $m/\Delta m$ 20,000 (full-width at half-maximum definition). Data acquisition of an ion mobility experiment consisted of 200 bins. The CCS values, obtained in nitrogen, were experimentally determined using singly charged Poly-DL-alanine oligomers as the TWIM calibrant species for ESI+. The CCS values were derived according to previously reported procedures [41]. The ion mobility resolution was $\sim 40 \Omega/\Delta\Omega$ (fwhm). The N2 CCS values reported were determined at the apex of the ion-mobility peak.

4.3.4. Raw Data Processing

All LC-IM-MS raw data files data processing, peak detection and peak matching across samples using retention time (t_R) correction and chromatographic alignment along with drift time and CCS calculation were performed using Progenesis Q1 (Waters MS Technologies, Manchester, UK) to yield a data matrix containing retention times, accurate masses, CCS, and peak intensities. The preprocessing step resulted in an X-matrix where t_R , CCS and m/z values were concatenated into “ t_R_m/z_CCS ” features (in columns) present in each sample (in rows) with corresponding peak areas.

4.3.5. Quality Control

Aliquoted 10 μ L of each urine sample were mixed together to generate pooled quality control samples (QCs). The QCs and solvent blank samples (mobile phase) were injected sequentially in-between the urine samples. In addition, a dilution series of QC samples (6%, 12.5%, 25%, 50% and 100% original concentration) were used to assess the quality of the extracted features. In this study, we used a filter strategy in which the features intensity must be correlated to the matrix concentrations in a series of diluted QC samples in order to be included in further analysis. Feature groups with correlation coefficient of less than 0.70 were removed from the dataset. Furthermore, datasets were refined by removal of feature groups that did not meet the threshold of peak area measurement precision prior to data analysis. This approach uses RSD values derived from repeated measurements of a pooled QC sample. The threshold was set to RSD < 25% to enhance the biological interpretation of metabolomics data.

4.4. Targeted Analysis

4.4.1. Amino Acids and Acylcarnitines Profiling

Amino acids, free carnitine, and acylcarnitines were semi-quantified using a flow injection tandem mass spectrometry method based on a Masschrom[®] amino acids and acylcarnitines kit (chromsystems, gräfelfing germany). The analyses were performed following the instructions of the kit manufacturer using liquid chromatography instrument prominence Shimadzu UFLC System (Shimadzu, Prominence, Kyoto, Japan) coupled to the 4000 Qtrap mass spectrometer (Sciex, Framingham, MA, USA) with an electrospray ion source. A system suitability test was conducted before each batch of the samples (analysis of a standard mixture) to warm up the LS-MS/MS system and check the inter-day performance of the system. Data acquisition and processing were performed using the Analyst 1.5 software (Sciex, Framingham, MA, USA). The list of measured MRM transitions is presented in Tables S3 and S4.

4.4.2. Dermatan Sulfate Assessment

Dermatan sulfate was assessed as previously described [42]. Briefly, 25 μ L of homogenized urine sample was evaporated under a nitrogen stream. Then, 500 μ L of a commercial methanolic HCl 3 N solution was added and vortexed. The samples were incubated at 65 °C for 1 h, then immediately evaporated under a nitrogen stream and then suspended in 200 μ L of the resuspension solution containing dermatan sulfate (9 μ g/mL) 90 :10 acetonitrile :water solution. Samples were transferred to vials and centrifuged prior to the injection of 2 μ L in the HPLC-MS/MS system.

4.5. Data Analysis and Modeling

Support vector regression normalization method was applied using the MetNormalizer R package [43] before any data analysis of untargeted metabolomics data to remove the unwanted intra- and inter-batch measurement analytical variations. Then the normalized data matrix was log-transformed and pareto-scaled. All data analyses and modeling were done using SIMCA 14.0 (MKS DAS, Umeå, Sweden). First, hierarchical cluster analysis was applied to the dataset to get an overview of the clustering trends of samples with similar profiles of variable intensity. Furthermore, multivariate data analysis and modeling was performed using principal component analysis (PCA) as an unsupervised method where the variations in correlated variables were summarized into a smaller number of latent variables, so-called principal components. The latent variables were used to describe the observations. The relationship between observations that are characterized by many variables can be visualized in low dimensional plots. This is what makes PCA a dimension reduction method to better visualize high numbers of variables. Principal component analysis was first

applied to get an overview of the data and identify potential severe outliers which were defined as observations whose scores mapped outside the Hotelling's T2 ellipse (confidence interval = 0.95) in a cross-validated seven-component model [44]. Orthogonal partial least-squares-discriminant analysis (OPLS-DA) was used as a supervised method. This method aims to maximize the covariance between the independent variables (metabolites) and the corresponding dependent variable Y (groups, i.e., Control vs Disease) by finding a linear subspace of the explanatory variables. This space allows the prediction of the Y variable (groups) based on a reduced number of factors (PLS components). This method provides several statistical metrics such as variable importance on projection (VIP) which is a parameter used for calculating the cumulative measure of the influence of individual X-variables on the model, and therefore the discriminative power of each variable (metabolite). To select the most relevant features, the training group was repeatedly been split into a training set and a test set. A permutation test (999 iterations) was performed to prevent the OPLS-DA over fitting of the model by comparing diagnostic statistic metrics of the generated model with those of randomly generated models. R2X is the cumulative modeled variation in X (X = features), R2Y is the cumulative modeled variation in Y (Y = sample groups), and Q2Y is the cumulative predicted variation in Y, based on the cross-validation. The range of these parameters was between 0 and 1, where 1 indicates a perfect fit. Furthermore, cross-validated analysis of variance (CV-ANOVA) was systematically performed based on the cross-validated model [45]. The X matrix was 82 × 854 variables. For all differential analysis, confounding factor effects (gender and age) correction was made using the limma package [46]. Given the high number of tested metabolites, the probability of finding a random association between a given metabolite and the group increases. This accumulation of false positives is called the multiple testing problem when a statistical test is applied across multiple features. Therefore, retrieved *p*-values from multiple tests performed in parallel across the metabolites should be corrected. Different methods are proposed to handle this issue. We used the false discovery rate method with 0.05% as the cut-off. Sensitivity and specificity are two basic measures of diagnostic accuracy. Sensitivity, or true-positive probability (TPP), is the probability of test results being positive for actually abnormal cases. Specificity, or true-negative probability (TNP), is the probability of test results being negative for actually control cases. The receiver operating characteristic (ROC) curve is a plot of TPP versus its false-positive fraction (FPF), or 1-specificity. Thus, the ROC curve is a valuable tool to assess and compare the diagnostic accuracy between biomarkers [47]. The area under the entire curve (AUC) is widely used as a summary metrics of the ROC curve. The AUC might be seen as an average value of sensitivity for all possible values of specificity. Thus, the greater the AUC, the better is the test. A perfect biomarker would have 100% sensitivity and 100% specificity. The point on the ROC curve which corresponds to this perfect case would be at the upper left-hand corner (0, 1) which lead to an AUC = 1 while a useless biomarker would have AUC = 0.5. Most of biomarkers fall between these two values.

4.6. Feature Selection, Annotation, and Network Analysis

To select the most discriminant variables for the separation of groups we use the covariances and correlations between the X matrix and OPLS scores. The covariance values give the magnitude of the contribution of a variable while the correlation values reflect the effect and reliability of the variable for the model component scores. Variables with both very high correlation and covariance are important for the model's explanation. Furthermore, selection of discriminant variables was achieved using the VIP score procedures for each validated OPLS-DA model [48]. Putative annotation of detected features was performed using accurate mass comparison using freely available metabolite databases HMDB, LipidBlast, and Metlin. Furthermore, CCS values were also compared to the MetCCS database [49]. In order to provide a broader understanding of metabolic changes in MPS VI, we also explored the biochemical pathways using a network analysis approach using Mummichog (v.1.0.5) which allows pathway enrichment analyses. The idea behind this metabolic network prediction strategy assumes that metabolite concentration alterations are more likely to occur within a metabolic connected network rather than in a random fashion. This Mummichog Python package highlights pathways that are significantly impacted in the studied groups. Significantly impacted biochemical pathways are those exhibiting an adjusted *p*-value < 0.05. For this comparison, we focused on features that significantly changed (511 features with *q*-values = 0.05 and FDR = 5%). Mummichog annotates metabolites based on accurate mass *m/z* (5 ppm mass error was used) and tests significant pathway enrichment within a reference metabolic network using a Fisher's exact test [50]. The matched candidates were then mapped to reference human metabolic networks from the KEGG, MetaCyc, Recon, and Edinburgh Human Metabolic Network. The null distribution in pathway analysis was obtained from

1000 set of randomly permuted m/z lists draw from all features detected in the whole metabolomic dataset and modeled by Gamma distribution. To protect against incorrect pathway selection, redundant pathways or those enriched by fewer than two metabolites were excluded. MetaboAnalyst [51] has been used for metabolite set enrichment analysis using the amino acid concentration matrix.

5. Conclusions

Profound metabolic remodeling beyond the targeted deficient pathway is unveiled using an integrative untargeted and targeted metabolomics approach to investigate urine samples originating from MPS VI patients. This study revealed the alteration of different metabolisms seemingly without any apparent link with the metabolism of dermatan sulfate, although there are some studies showing a link between glycosaminoglycans and lipid metabolism, including lipid deposits in the case of GAG accumulation [52]. This shows the importance of such global omics studies to reveal possible links between different nodes in a metabolic network that could not be revealed with reductionist strategies. This crucial information is needed for a holistic understanding of functional mechanisms underlying this rare condition. More targeted validation studies are required to explore the impaired metabolism at a mechanistic level. This may end up to a better management of these patients by tuning these altered pathways through autophagy modulation and antioxidant adjuvants which might potentiate MPS specific therapies.

Supplementary Materials : Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/2/446/s1>.

Author Contributions : Conceptualization, S.B.; Data curation, A.T., L.A.-D., I.S.-A., C.P., and T.P.; Formal analysis, A.T.; Methodology, A.T., C.A., and S.B.; Project administration, S.B.; Resources, M.P., J.A., F.Z., S.M., C.A., and S.B.; Supervision, S.B.; Visualization, A.T.; Writing—original draft, A.T.; Writing—review & editing, S.M., C.A., and S.B. All authors read and approved the final manuscript.

Funding : The authors gratefully acknowledge the Region Haute-Normandie, the European Regional Development Fund (ERDF 31708), Sanofi-Genzyme, Shire and Vaincre les Maladies Lysosomales for their support.

Acknowledgments : The authors are grateful to Thomas Plichet (Rouen University Hospital, France) for his technical assistance.

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

Abbreviations

IEM	Inborn errors of metabolism
LSD	Lysosomal storage diseases
MPS	Mucopolysaccharidoses
GAGs	Glycosaminoglycans
MPS VI	Mucopolysaccharidosis type VI
ERT	Enzyme replacement therapy
HS	Heparan sulfate
CCS	Collision cross section
DS	Dermatan sulfate
KS	Keratan sulfate
ROC	Receiver operating characteristic
FDR	False discovery rate
PCA	Principal component analysis
OPLS-DA	Orthogonal partial least-squares-discriminant analysis
VIP	Variable influence in projection
AUC	Area under curve

References

1. Bekri, S. The role of metabolomics in precision medicine. *Expert Rev. Precis. Med. Drug Dev.* **2016**, *1*, 517–532. [[CrossRef](#)]
2. Tebani, A.; Afonso, C.; Marret, S.; Bekri, S. Omics-based strategies in precision medicine : Toward a paradigm shift in inborn errors of metabolism investigations. *Int. J. Mol. Sci.* **2016**, *17*, 1555. [[CrossRef](#)]
3. Nicholson, J.K.; Lindon, J.C.; Holmes, E. ‘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.* **1999**, *29*, 1181–1189. [[CrossRef](#)] [[PubMed](#)]
4. Fiehn, O. Metabolomics—The link between genotypes and phenotypes. *Plant Mol. Biol.* **2002**, *48*, 155–171. [[CrossRef](#)] [[PubMed](#)]
5. Tebani, A.; Abily-Donval, L.; Afonso, C.; Marret, S.; Bekri, S. Clinical metabolomics : The new metabolic window for inborn errors of metabolism investigations in the post-genomic era. *Int. J. Mol. Sci.* **2016**, *17*, 1167. [[CrossRef](#)]
6. Auray-Blais, C.; Maranda, B.; Lavoie, P. High-throughput tandem mass spectrometry multiplex analysis for newborn urinary screening of creatine synthesis and transport disorders, triple h syndrome and otc deficiency. *Clin. Chim. Acta Int. J. Clin. Chem.* **2014**, *436*, 249–255. [[CrossRef](#)] [[PubMed](#)]
7. Pitt, J.J. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin. Biochem. Rev. Aust. Assoc. Clin. Biochem.* **2009**, *30*, 19–34.
8. Pitt, J.J. Newborn screening. *Clin. Biochem. Rev. Aust. Assoc. Clin. Biochem.* **2010**, *31*, 57–68.
9. Pitt, J.J.; Eggington, M.; Kahler, S.G. Comprehensive screening of urine samples for inborn errors of metabolism by electrospray tandem mass spectrometry. *Clin. Chem.* **2002**, *48*, 1970–1980.
10. Spacil, Z.; Tatipaka, H.; Barcenas, M.; Scott, C.R.; Turecek, F.; Gelb, M.H. High-throughput assay of 9 lysosomal enzymes for newborn screening. *Clin. Chem.* **2013**, *59*, 502–511. [[CrossRef](#)]
11. Therrell, B.L.; Padilla, C.D.; Loeber, J.G.; Kneisser, I.; Saadallah, A.; Borrajo, G.J.; Adams, J. Current status of newborn screening worldwide : 2015. *Semin. Perinatol.* **2015**, *39*, 171–187. [[CrossRef](#)] [[PubMed](#)]
12. Tebani, A.; Afonso, C.; Bekri, S. Advances in metabolome information retrieval : Turning chemistry into biology. Part ii : Biological information recovery. *J. Inherit. Metab. Dis.* **2018**, *41*, 393–406. [[CrossRef](#)] [[PubMed](#)]
13. Tebani, A.; Afonso, C.; Bekri, S. Advances in metabolome information retrieval : Turning chemistry into biology. Part i : Analytical chemistry of the metabolome. *J. Inherit. Metab. Dis.* **2018**, *41*, 379–391. [[CrossRef](#)] [[PubMed](#)]
14. Ballabio, A.; Gieselmann, V. Lysosomal disorders : From storage to cellular damage. *Biochim. Biophys. Acta* **2009**, *1793*, 684–696. [[CrossRef](#)] [[PubMed](#)]
15. Wraith, J.E. The mucopolysaccharidoses : A clinical review and guide to management. *Arch. Dis. Child.* **1995**, *72*, 263–267. [[CrossRef](#)]
16. Neufeld, E.F.; Muenzer, J. The mucopolysaccharidoses. In *The Metabolic and Molecular Basis of Inherited Disease*; Scriver, C., Ballabio, A., Sly, W., Vaele, D., Eds.; Mc Graw-Hill : New York, NY, USA, 2001; pp. 3421–3452.

17. Stewart, F.J.; Bentley, A.; Burton, B.K.; Guffon, N.; Hale, S.L.; Harmatz, P.R.; Kircher, S.G.; Kochhar, P.K.; Mitchell, J.J.; Plöckinger, U.; et al. Pregnancy in patients with mucopolysaccharidosis : A case series. *Mol. Genet. Metab. Rep.* **2016**, *8*, 111–115. [[CrossRef](#)]
18. Montano, A.M.; Lock-Hock, N.; Steiner, R.D.; Graham, B.H.; Szlago, M.; Greenstein, R.; Pineda, M.; Gonzalez-Meneses, A.; Coker, M.; Bartholomew, D.; et al. Clinical course of sly syndrome (mucopolysaccharidosis type vii). *J. Med. Genet.* **2016**, *53*, 403–418. [[CrossRef](#)]
19. Parenti, G.; Andria, G.; Ballabio, A. Lysosomal storage diseases : From pathophysiology to therapy. *Annu. Rev. Med.* **2015**, *66*, 471–486. [[CrossRef](#)]
20. Maroteaux, P.; Leveque, B.; Marie, J.; Lamy, M. A new dysostosis with urinary elimination of chondroitin sulfate b. *La Presse Med.* **1963**, *71*, 1849–1852.
21. Muenzer, J. Overview of the mucopolysaccharidoses. *Rheumatology* **2011**, *50* (Suppl. 5), v4–v12. [[CrossRef](#)]
22. Valayannopoulos, V.; Nicely, H.; Harmatz, P.; Turbeville, S. Mucopolysaccharidosis vi. *Orphanet J. Rare Dis.* **2010**, *5*, 5. [[CrossRef](#)] [[PubMed](#)]
23. Hendriksz, C.J.; Giugliani, R.; Harmatz, P.; Lampe, C.; Martins, A.M.; Pastores, G.M.; Steiner, R.D.; Leão Teles, E.; Valayannopoulos, V. Design, baseline characteristics, and early findings of the MPS VI (mucopolysaccharidosis vi) clinical surveillance program (csp). *J. Inherit. Metab. Dis.* **2013**, *36*, 373–384. [[CrossRef](#)]
24. Quartel, A.; Hendriksz, C.J.; Parini, R.; Graham, S.; Lin, P.; Harmatz, P. Growth charts for individuals with mucopolysaccharidosis vi (maroteaux–lamy syndrome). In *JIMD Reports*; Zschocke, J., Baumgartner, M., Morava, E., Patterson, M., Rahman, S., Peters, V., Eds.; Springer : Berlin/Heidelberg, Germany, 2015; Volume 18, pp. 1–11.
25. Scarpa, M.; Buffone, E.; Marca, P.L.; Campello, M.; Rampazzo, A. Difficulties in diagnosing slowly progressive mucopolysaccharidosis vi : A case series. *J. Pediatr. Rehabil. Med.* **2010**, *3*, 71–75. [[PubMed](#)]
26. Harmatz, P.; Hendriksz, C.J.; Lampe, C.; McGill, J.J.; Parini, R.; Leao-Teles, E.; Valayannopoulos, V.; Cole, T.J.; Matousek, R.; Graham, S.; et al. The effect of galsulfase enzyme replacement therapy on the growth of patients with mucopolysaccharidosis vi (maroteaux–lamy syndrome). *Mol. Genet. Metab.* **2017**, *122*, 107–112. [[CrossRef](#)]
27. Furujo, M.; Kosuga, M.; Okuyama, T. Enzyme replacement therapy attenuates disease progression in two japanese siblings with mucopolysaccharidosis type vi : 10-year follow up. *Mol. Genet. Metab. Rep.* **2017**, *13*, 69–75. [[CrossRef](#)] [[PubMed](#)]
28. Jester, S.; Larsson, J.; Eklund, E.A.; Papadopoulou, D.; Månsson, J.-E.; Békássy, A.N.; Turkiewicz, D.; Toporski, J.; Øra, I. Haploidentical stem cell transplantation in two children with mucopolysaccharidosis vi : Clinical and biochemical outcome. *Orphanet J. Rare Dis.* **2013**, *8*, 134. [[CrossRef](#)] [[PubMed](#)]
29. Salazar, D.A.; Rodriguez-Lopez, A.; Herrero, A.; Barbosa, H.; Herrera, J.; Ardila, A.; Barreto, G.E.; Gonzalez, J.; Almeciga-Diaz, C.J. Systems biology study of mucopolysaccharidosis using a human metabolic reconstruction network. *Mol. Genet. Metab.* **2016**, *117*, 129–139. [[CrossRef](#)]
30. Fiorenza, M.T.; Moro, E.; Erickson, R.P. The pathogenesis of lysosomal storage disorders : Beyond the engorgement of lysosomes to abnormal development and neuroinflammation. *Hum. Mol. Genet.* **2018**, *27*, R119–R129. [[CrossRef](#)]
31. Petrova, R.; Joyner, A.L. Roles for hedgehog signaling in adult organ homeostasis and repair. *Development* **2014**, *141*, 3445–3457. [[CrossRef](#)]
32. Whalen, D.M.; Malinauskas, T.; Gilbert, R.J.; Siebold, C. Structural insights into proteoglycan-shaped hedgehog signaling. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16420–16425. [[CrossRef](#)]
33. Chang, S.C.; Mulloy, B.; Magee, A.I.; Couchman, J.R. Two distinct sites in sonic hedgehog combine for heparan sulfate interactions and cell signaling functions. *J. Biol. Chem.* **2011**, *286*, 44391–44402. [[CrossRef](#)] [[PubMed](#)]
34. Donida, B.; Jacques, C.E.D.; Mescka, C.P.; Rodrigues, D.G.B.; Marchetti, D.P.; Ribas, G.; Giugliani, R.; Vargas, C.R. Oxidative damage and redox in lysosomal storage disorders : Biochemical markers. *Clin. Chim. Acta Int. J. Clin. Chem.* **2017**, *466*, 46–53. [[CrossRef](#)] [[PubMed](#)]
35. Tebani, A.; Schmitz-Afonso, I.; Abily-Donval, L.; Heron, B.; Piraud, M.; Ausseil, J.; Brassier, A.; De Lonlay, P.; Zerimech, F.; Vaz, F.M.; et al. Urinary metabolic phenotyping of mucopolysaccharidosis type i combining untargeted and targeted strategies with data modeling. *Clin. Chim. Acta Int. J. Clin. Chem.* **2017**, *475*, 7–14. [[CrossRef](#)] [[PubMed](#)]
36. Tebani, A.; Abily-Donval, L.; Schmitz-Afonso, I.; Héron, B.; Piraud, M.; Ausseil, J.; Zerimech, F.; Gonzalez, B.; Marret, S.; Afonso, C.; et al. Unveiling metabolic remodeling in mucopolysaccharidosis type iii through integrative metabolomics and pathway analysis. *J. Transl. Med.* **2018**, *16*, 248. [[CrossRef](#)] [[PubMed](#)]

37. Nagesh Babu, G.; Gupta, M.; Paliwal, V.K.; Singh, S.; Chatterji, T.; Roy, R. Serum metabolomics study in a group of parkinson's disease patients from northern india. *Clin. Chim. Acta Int. J. Clin. Chem.* **2018**, *480*, 214–219. [[CrossRef](#)] [[PubMed](#)]
38. Bellia, F.; Vecchio, G.; Cuzzocrea, S.; Calabrese, V.; Rizzarelli, E. Neuroprotective features of carnosine in oxidative driven diseases. *Mol. Aspects Med.* **2011**, *32*, 258–266. [[CrossRef](#)] [[PubMed](#)]
39. Cooper, K.F. Till death do us part : The marriage of autophagy and apoptosis. *Oxid. Med. Cell. Longev.* **2018**, *2018*, 4701275. [[CrossRef](#)]
40. Donida, B.; Marchetti, D.P.; Jacques, C.E.D.; Ribas, G.; Deon, M.; Manini, P.; da Rosa, H.T.; Moura, D.J.; Saffi, J.; Giugliani, R.; et al. Oxidative profile exhibited by mucopolysaccharidosis type iva patients at diagnosis : Increased keratan urinary levels. *Mol. Genet. Metab. Rep.* **2017**, *11*, 46–53. [[CrossRef](#)]
41. Paglia, G.; Angel, P.; Williams, J.P.; Richardson, K.; Olivos, H.J.; Thompson, J.W.; Menikarachchi, L.; Lai, S.; Walsh, C.; Moseley, A.; et al. Ion mobility-derived collision cross section as an additional measure for lipid fingerprinting and identification. *Anal. Chem.* **2015**, *87*, 1137–1144. [[CrossRef](#)]
42. Auray-Blais, C.; Lavoie, P.; Tomatsu, S.; Valayannopoulos, V.; Mitchell, J.J.; Raiman, J.; Beaudoin, M.; Maranda, B.; Clarke, J.T.R. Uplc-ms/ms detection of disaccharides derived from glycosaminoglycans as biomarkers of mucopolysaccharidoses. *Anal. Chim. Acta* **2016**, *936*, 139–148. [[CrossRef](#)]
43. Shen, X.; Gong, X.; Cai, Y.; Guo, Y.; Tu, J.; Li, H.; Zhang, T.; Wang, J.; Xue, F.; Zhu, Z.-J. Normalization and integration of large-scale metabolomics data using support vector regression. *Metabolomics* **2016**, *12*, 89. [[CrossRef](#)]
44. Eriksson, L.; Trygg, J.; Wold, S. A chemometrics toolbox based on projections and latent variables. *J. Chemom.* **2014**, *28*, 332–346. [[CrossRef](#)]
45. Eriksson, L.; Trygg, J.; Wold, S. Cv-anova for significance testing of pls and opl^s models. *J. Chemom.* **2008**, *22*, 594–600. [[CrossRef](#)]
46. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [[CrossRef](#)] [[PubMed](#)]
47. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology* **1982**, *143*, 29–36. [[CrossRef](#)] [[PubMed](#)]
48. Galindo-Prieto, B.; Eriksson, L.; Trygg, J. Variable influence on projection (vip) for orthogonal projections to latent structures (opls). *J. Chemom.* **2014**, *28*, 623–632. [[CrossRef](#)]
49. Zhou, Z.; Xiong, X.; Zhu, Z.J. Metccs predictor : A web server for predicting collision cross-section values of metabolites in ion mobility-mass spectrometry based metabolomics. *Bioinformatics* **2017**, *33*, 2235–2237. [[CrossRef](#)] [[PubMed](#)]
50. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F.H.; Khan, N.; Soltow, Q.A.; Jones, D.P.; Pulendran, B. Predicting network activity from high throughput metabolomics. *PLoS Comput. Biol.* **2013**, *9*, e1003123. [[CrossRef](#)]
51. Xia, J.; Sinelnikov, I.V.; Han, B.; Wishart, D.S. Metaboanalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Res.* **2015**, *43*, W251–W257. [[CrossRef](#)]
52. Wight, T.N. A role for proteoglycans in vascular disease. *Matrix Biol.* **2018**, *71–72*, 396–420. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).

1.3 Application des sciences « omiques » dans la maladie de Fabry : Projet METALYS - étude Fabomics

1.3.1 Objectifs de l'étude

Le projet FABOMICS avait pour objectif d'appliquer plusieurs techniques de sciences « omiques » (transcriptomique, métabolomique, protéomique) à la maladie de Fabry. L'utilisation de ces outils avait pour but de rechercher un profil « omique » spécifique de la maladie de Fabry en comparaison à un groupe contrôle. Cela pourrait permettre d'identifier des biomarqueurs pour le diagnostic et le suivi de cette maladie notamment chez les patients pauci ou asymptomatiques et/ou chez les femmes chez lesquels le diagnostic est parfois difficile et long à obtenir. Le projet avait également pour but de préciser les mécanismes physiopathologiques de cette maladie en visualisant les voies métaboliques modulées.

1.3.2 Collecte des échantillons

Quatre-vingt-cinq patients Fabry issus de la cohorte FFabry (310) (Dr Wladimir Mauhin, Dr Olivier Lidove, Pr. Olivier Benveniste) et 83 sujets sains ont été inclus. Pour valider nos résultats, nous avons eu accès à des échantillons provenant de collections déclarées d'échantillons d'autres maladies lysosomales : 44 échantillons Pompe (Pr. Pascal Laforêt), 20 échantillons Niemann-Pick type C (Dr Yann Nadjar et Dr Foudil Lamari – Pitié Salpêtrière) et 48 échantillons Gaucher (Pr. Marc Berger - CETG)

1.3.3 Partenaires

Ce travail a bénéficié de l'expertise de :

- Tous les centres experts qui ont adressé les collections biologiques et les métadonnées clinico-biologiques.
- Service de Biochimie Métabolique - Pr Soumeya Bekri - CHU de Rouen
- Laboratoire COBRA - UMR 6014 -IRCOF Equipe 1 – Université des sciences Mont-Saint-Aignan
- Equipe 4 INSERM U1245.- Dr Bruno Gonzalez- Université de Rouen

1.3.4 Résultats et publication

Quarante protéines plasmatiques ont été quantifiées par électro chimiluminescence (Meso Scale Discovery [MSD]). L'analyse différentielle des résultats obtenus montre que des cytokines et des facteurs d'angiogenèse (famille IL7 et VEGF) sous-tendent la séparation entre le groupe Fabry et le groupe contrôle. L'analyse des échantillons issus de patients atteints d'autres maladies lysosomales a permis de montrer une surexpression significative de ces molécules chez les patients Fabry.

Ces données sont détaillées dans une publication en cours de préparation.

Tebani A, Mauhin W, **Abily-Donval L**, Nadjar Y, Lamari F, Laforêt P, Marret S, Lidove O, Bekri S. Proteomics analysis in Fabry disease reveals predictive biological patterns

II. Partie 2 : modulation de l'autophagie dans la maladie de Fabry

La maladie de Fabry est une pathologie lysosomale liée au déficit en α -galactosidase A (181, 311). Cette maladie est caractérisée par l'accumulation progressive de glycosphingolipides (en particulier le globotriaosylcéramide Gb3) dans les cellules de nombreux tissus (312). La physiopathologie de la maladie de Fabry est complexe et les symptômes ne peuvent pas être seulement expliqués par l'accumulation de Gb3 (189, 312). Un traitement symptomatique par enzymothérapie substitutive existe depuis 2001 mais les effets sont variables selon les patients.

L'autophagie est une fonction biologique permettant le maintien de l'homéostasie cellulaire. Elle a un rôle de dégradation et de recyclage indispensable à la vie cellulaire. Ce processus semble altéré dans plusieurs maladies lysosomales (73, 107, 110). Dans la maladie de Fabry, une accumulation de vacuoles autophagiques a été observée dans le rein. Ces vacuoles diminuent après trois ans de traitement par enzymothérapie substitutive. L'étude in vitro de fibroblastes et de lymphoblastes de patients atteints de la maladie de Fabry a montré une diminution du flux autophagique (98). Une dysfonction de l'autophagie pourrait donc expliquer la physiopathologie de la maladie de Fabry. Notre étude s'est donc intéressée à l'impact de la modulation du processus autophagique sur un modèle cellulaire Fabry en ciblant l'étude sur l'adressage de l'enzyme substitutive au lysosome.

2.1 Objectifs de l'étude

L'objectif de ce travail était d'évaluer les effets de la modulation du processus autophagique sur le devenir et l'activité de l'enzyme substitutive dans la cellule. L'hypothèse de notre travail était que dans les cellules Fabry, l'accumulation de substrat dans les lysosomes a un impact sur la maturation des autophagosomes et pourrait, ainsi, perturber le trafic intracellulaire et diminuer alors l'adressage de l'enzyme substitutive vers son lieu d'action le lysosome. L'adressage de l'enzyme exogène aux lysosomes a été étudié dans des fibroblastes issus de patients atteints de la maladie de Fabry et de témoins sains et sur une lignée podocytaire KO pour le gène *GLA*. Les effets de la modulation de l'autophagie sur l'adressage de l'enzyme exogène aux lysosomes et sur sa capacité à dégrader le Gb3 ont été évalués.

2.2 Partenaires

Ce travail a bénéficié d'une collaboration avec le Dr Frédérique Barbey du CHU de Lausanne et de l'expertise de l'équipe 4 de l'unité Inserm U1245.

2.3 Résultats et publication associée

L'internalisation de l'enzyme exogène dans l'endosome précoce dans les cellules de Fabry n'est pas altérée. Cependant, son adressage aux lysosomes est significativement retardé dans les fibroblastes Fabry comparativement aux fibroblastes contrôles. Une inhibition de l'autophagie améliore l'adressage de l'enzyme exogène aux lysosomes et permet ainsi d'augmenter significativement la dégradation du substrat Gb3.

Les résultats de cette étude sont présentés sous forme d'article en cours de préparation :

L. Abily-Donval, F. Barbey, S. Torre, C. Lesueur, M. Chevrier, T. Pereira, C. Pilon-Tardivel, B.J. Gonzalez, S. Marret, A. Tebani, S. Bekri. Autophagy modulation enhances ERT response in Fabry disease.

DISCUSSION ET PERSPECTIVES

I. Médecine de précision et maladies héréditaires du métabolisme

Le concept de la médecine de précision a émergé ces dernières années et vise une prise en charge personnalisée du patient adaptée à ses caractéristiques individuelles ; et donc un diagnostic et une prise en charge optimisés (313). La médecine de précision utilise à la fois des données individuelles propres au patient mais également des données populationnelles. Ces stratégies offrent la possibilité d'explorer et de mieux comprendre un système biologique sans hypothèse préalable.

Le challenge de la biologie moderne est d'appréhender une maladie comme un processus intégré, dynamique et complexe, en étudiant les changements et les interactions entre les différents niveaux d'information biologique. Cette médecine, grâce à l'avènement de technologies avancées permettant l'acquisition de données multiplex à haut débit, permet d'explorer ainsi tous les supports de l'information biologiques : l'ADN (génomique), l'ARN (transcriptomique), les protéines (protéomique), et les métabolites (métabolomique). Les technologies « omiques » et l'intégration des données permettront à terme d'augmenter le nombre de diagnostics et de réduire leur délai. Cette approche contraste avec les techniques conventionnelles qui sont séquentielles et qui se basent sur des hypothèses préalables.

La médecine de précision a d'abord été développée pour des disciplines telles que la pharmacologie (314-316) ou la cancérologie (313, 317, 318). Des variants génétiques peuvent en effet modifier la pharmacocinétique d'un médicament chez un patient (319). L'étude génétique des cellules tumorales permet une meilleure compréhension des mécanismes physiopathologiques du cancer et une adaptation du traitement avec des thérapies ciblées (320). En effet, l'exploration du génome permet d'étudier les caractéristiques génétiques individuelles des patients ce qui peut préciser le pronostic d'une pathologie ou prédire la réponse à un traitement. Néanmoins, les processus biologiques physiologiques et pathologiques sont la résultante d'une interaction complexe entre les gènes et leur environnement. Certaines informations biologiques fonctionnelles ne sont pas accessibles par la seule étude du génome et les autres techniques « omiques » sont donc complémentaires et indispensables pour une meilleure compréhension des mécanismes physiopathologiques moléculaires des maladies. L'utilisation de ces différents niveaux de l'information biologique permet de préciser au mieux le phénotype clinique du patient (253, 275).

Les études et les publications sur la médecine de précision se sont multipliées dans la littérature sur les dernières années. Néanmoins, la revue de la littérature a retrouvé peu d'études sur l'exploration des maladies héréditaires du métabolisme (MHM) par ces approches.

Pourtant, le domaine des maladies héréditaires du métabolisme semble être particulièrement intéressant pour l'application d'études « omiques » (253). Ces maladies sont caractérisées par le dysfonctionnement d'une voie métabolique dans l'organisme le plus souvent lié au déficit quantitatif ou qualitatif d'une enzyme, d'un cofacteur ou d'un transporteur. Le métabolisme humain est un processus complexe dans lequel les principales voies métaboliques des protéines, des glucides, des lipides en association avec des vitamines et des oligoéléments interagissent pour fournir à l'organisme l'énergie et les molécules indispensables au fonctionnement cellulaire. Les MHM ont été définies comme "monogéniques" et sont en majorité de transmission autosomique récessive. En 1902, Sir Archibald Garrod, à travers la description de l'alcaptonurie, a émis le paradigme « un gène - une pathologie » (321). Cependant, plusieurs phénotypes de sévérité variables sont possibles à partir de la même anomalie génétique (322). La réponse au traitement notamment l'enzymothérapie substitutive dans les maladies lysosomales est également variable selon les patients. Les mécanismes physiopathologiques des MHM sont donc complexes et non complètement connus. La poursuite de travaux de recherche dans ce domaine est donc pertinente pour améliorer le pronostic des patients. Les technologies « omiques » pourraient être d'un grand apport pour décrypter cette complexité.

Le diagnostic des MHM se fait classiquement grâce à des techniques biochimiques conventionnelles segmentées et séquentielles sur plusieurs plateformes d'analyse distinctes. Ces analyses peuvent être longues et coûteuses nécessitant plusieurs techniques d'étude. Le pronostic des MHM est conditionné par la rapidité de mise en place des traitements disponibles et un diagnostic le plus précoce possible est donc primordial. Les études « omiques » pourraient permettre de mettre en évidence de nouveaux profils de biomarqueurs simples spécifiques de la pathologie et ainsi améliorer la prise en charge diagnostique, pronostique et thérapeutique. Elles pourraient également permettre une meilleure compréhension de la physiopathologie complexe des MHM en étudiant les voies métaboliques perturbées (322). Ces technologies sont initialement laborieuses à mettre en place devant le peu d'études disponibles et l'absence de protocoles standardisés. Des stratégies de validation doivent donc être réalisées dès que possible. Des bases de données permettent également de vérifier la cohérence des données obtenues. Le développement de stratégies d'analyse et de nouvelles méthodes « omiques » facilite par la suite l'analyse dynamique des systèmes biologiques étudiés.

L'objet de ce travail de thèse était donc d'appliquer des techniques « omiques » dans un groupe de maladies lysosomales, les mucopolysaccharidoses, et dans la maladie de Fabry afin de rechercher de nouveaux biomarqueurs spécifiques de ces maladies afin d'optimiser leur diagnostic et leur suivi, de développer d'éventuelles stratégies thérapeutiques. Nous avons, en collaboration avec le Laboratoire COBRA (Pr. Carlos Afonso) développé une analyse métabolomique non ciblée par la spectrométrie de

masse à haute résolution couplée à la chromatographie liquide ultra-haute performance et la mobilité ionique (306).

L'analyse des échantillons urinaires issus de patients atteints de mucopolysaccharidoses a permis de mettre en évidence des altérations du métabolisme de plusieurs biomolécules dont certains acides aminés. L'intégration des données issues des analyses non ciblées et des analyses ciblées des biomolécules discriminantes d'une part et des données *in silico* d'autre part (323) a permis de définir une liste de biomolécules qui sont différentiellement exprimées. L'analyse des voies métaboliques a montré que ces biomolécules sont associées à l'autophagie et au stress oxydatif. Par conséquent, l'altération de ces voies métaboliques participe à la physiopathologie des mucopolysaccharidoses. L'étude protéomique ciblée des plasmas issus de patients atteints de la maladie de Fabry versus des plasmas contrôles a permis de caractériser une augmentation des concentrations de protéines relatives à l'angiogenèse et à l'inflammation. Ces nouvelles données pourraient ouvrir la voie à de nouveaux biomarqueurs spécifiques mais également permettre une meilleure compréhension des mécanismes moléculaires et des voies métaboliques impliquées. Des études ultérieures ciblées sont à prévoir et permettront d'explorer ces données intéressantes. Une des perspectives de ce travail est de valider ces données chez l'animal et/ou dans un modèle cellulaire.

II. Etude de l'adressage de l'enzyme au lysosome et modulation de l'autophagie dans la maladie de Fabry

Notre étude a mis en évidence une altération du flux autophagique avec accumulation des autophagosomes immatures et retard d'adressage de l'enzyme au lysosome dans la maladie de Fabry (98). Le bloc autophagique semble diminuer l'efficacité du transport de l'enzyme au lysosome et donc diminuer les fonctions de dégradation du lysosome. D'autre part, l'expression de Rab 7 est significativement augmentée dans les fibroblastes de patients Fabry. Enfin, l'inhibition de l'autophagie par la wortmannine dans les cultures cellulaires augmente l'efficacité de l'enzymothérapie substitutive en diminuant l'accumulation du Gb3.

Comme nous l'avons vu précédemment, les lysosomes sont des organelles ayant de multiples fonctions dans la cellule (autophagie, dégradation de molécules, homéostasie lipidique,...) et leur dysfonctionnement est à l'origine de multiples pathologies (111, 112). Ces maladies ont un phénotype extrêmement variable et leurs mécanismes physiopathologiques ne sont pas précisément connus.

L'altération du flux autophagique a déjà été décrite dans d'autres maladies lysosomales avec un défaut de maturation des autophagosomes et une anomalie de la fusion autophagosome lysosome (103, 324).

Ainsi, dans le déficit multiple en sulfatases, il existe également une altération du flux autophagique avec accumulation de vacuoles autophagiques et défaut de maturation des autophagosomes (103). Le défaut de fonctionnement de l'autophagie induit notamment l'accumulation de molécules potentiellement délétères pour le fonctionnement cellulaire, par exemple l'ubiquitine (103). Néanmoins, la physiopathologie des maladies lysosomales semble beaucoup plus complexe qu'une simple accumulation de macromolécules. Les phénomènes de transport et de mouvement des molécules et du système endosome lysosome semble y jouer un rôle important.

Les lysosomes se déplacent dans la cellule et leur position varie selon l'environnement et selon leur rôle prédominant (325). Ainsi, selon la localisation du lysosome dans la cellule, le pH intraluminal et ses fonctions sont différents (326). Les lysosomes de la région périnucléaire ont un pH plus acide que ceux en périphérie et leur mobilité est limitée (326, 327). Les lysosomes se déplacent de façon bidirectionnelle le long des microtubules sous l'action de protéines motrices (328). La dynéine permet le mouvement vers la région périnucléaire alors que la kinésine guide vers le mouvement vers la périphérie cellulaire (329). A l'état basal, la majorité des lysosomes se concentrent dans la région périnucléaire (*perinuclear cloud*) qui entoure le centre d'organisation des microtubules (327). Les composés endocytés en périphérie vont passer dans l'endosome précoce qui va se diriger vers le nuage périnucléaire. L'endosome précoce va maturer en endosome tardif et fusionner avec le lysosome (327)

Le positionnement du lysosome dans la cellule varie en fonction de l'environnement cellulaire. Il peut être modifié dans certaines pathologies. La localisation périnucléaire des lysosomes pourrait être la conséquence de l'accumulation du cholestérol retrouvé dans de nombreuses maladies lysosomales (330, 331). Le cholestérol pourrait être impliqué dans la formation des vésicules intraluminales dans l'endosome tardif. L'accumulation intra lysosomale du cholestérol va activer la protéine Rab 7 via la protéine ORP1L (*oxysterol-binding protein-related protein*) localisée dans le lysosome (332). Rab 7 est une petite GTPase localisée dans les endosomes tardifs ou les lysosomes. Elle va interagir avec la dynéine et la kinésine et a un rôle dans le transport du système endosome lysosome dans la cellule. Elle permet également la fusion de l'autophagosome avec le lysosome (333-335). L'augmentation de l'expression de Rab7 dans les cellules Fabry pourrait donc être en lien avec le retard de maturation et l'accumulation des autophagosomes. Le système de transport via la dynéine pourrait dysfonctionner ce qui est cohérent avec le retard d'adressage de l'enzyme au lysosome.

La position du lysosome dans la cellule est également dépendante de l'environnement nutritionnel et donc de l'activité du complexe mTORC1, véritable chef d'orchestre du processus d'autophagie. En situation nutritionnelle riche, les lysosomes tendent à se localiser en périphérie de la cellule. Le complexe mTORC1 est recruté au niveau membranaire pour évaluer la disponibilité en nutriments ;

l'autophagie est inhibée. En situation de carence nutritionnelle, les lysosomes se localisent plutôt en région périnucléaire. mTORC1 est libéré du lysosome et son activité est diminuée (336). La fusion endosome/lysosome et l'autophagie sont augmentées. Le système de transport est régulé par l'intermédiaire de mTORC1. Un complexe nommé BORG (*BLOC-1-related complex*) favorise également la position périnucléaire du lysosome dans la cellule en fonction de l'environnement nutritionnel (337).

Les mécanismes physiopathologiques précis ne sont pas complètement connus. Notre étude a montré que l'action de la wortmannine diminue l'accumulation du Gb3 dans la cellule Fabry. La wortmannine est un inhibiteur de la PI3K (phosphatidylinositol 3-kinase) qui inhibe la formation des autophagosomes (338). Le complexe PI3K est capable d'interagir avec la dynéine via une interaction avec DLC1 (*Dynein light chain 1*) lorsqu'il est lié à Ambra1 (47). Il existe donc un lien entre l'autophagie et le fonctionnement du cytosquelette impliqué dans le transport au sein du système endosome lysosome. La modulation de l'autophagie pourrait avoir une action sur le transport de l'enzyme déficiente dans la cellule et améliorer son adressage au lysosome. Les mécanismes moléculaires précis impliqués dans l'adressage de l'enzyme au lysosome et sur son activité seront à préciser dans des études futures et sont probablement beaucoup plus complexes que les processus décrits jusqu'à présent.

III. Conclusion

La recherche dans le domaine des maladies lysosomales par des techniques aussi variées que les techniques « omiques » ou des techniques mécanistiques plus conventionnelles permet une meilleure compréhension du fonctionnement du lysosome. Cette organelle qui paraissait être un simple lieu de dégradation se révèle en réalité très complexe. Ces nombreuses fonctions sont finement régulées et sont soumises à l'influence de l'environnement cellulaire. De multiples molécules et gènes y interagissent créant un véritable système biologique élaboré. Cette complexité permet d'expliquer la difficulté à développer des traitements efficaces mais ouvre également la possibilité à de multiples stratégies thérapeutiques et diagnostiques porteuses d'espoir pour les patients.

BIBLIOGRAPHIE

1. De Duve C. The lysosome. *Sci Am.* 1963;208 :64-72.
2. Bainton DF. The discovery of lysosomes. *J Cell Biol.* 1981;91(3 Pt 2) :66s-76s.
3. Saftig P, Klumperman J. Lysosome biogenesis and lysosomal membrane proteins : trafficking meets function. *Nat Rev Mol Cell Biol.* 2009;10(9) :623-35.
4. Darmoise A, Maschmeyer P, Winau F. The immunological functions of saposins. *Adv Immunol.* 2010;105 :25-62.
5. Staudt C, Puissant E, Boonen M. Subcellular Trafficking of Mammalian Lysosomal Proteins : An Extended View. *Int J Mol Sci.* 2016;18(1).
6. Saftig P. Physiology of the lysosome. In : Mehta A, Beck M, Sunder-Plassmann G, editors. *Fabry Disease : Perspectives from 5 Years of FOS.* Oxford2006.
7. Luzio JP, Pryor PR, Bright NA. Lysosomes : fusion and function. *Nat Rev Mol Cell Biol.* 2007;8(8) :622-32.
8. Lefrancois S, Zeng J, Hassan AJ, Canuel M, Morales CR. The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin. *EMBO J.* 2003;22(24) :6430-7.
9. Kudo M, Brem MS, Canfield WM. Mucopolidosis II (I-cell disease) and mucopolidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha / beta -subunits precursor gene. *Am J Hum Genet.* 2006;78(3) :451-63.
10. Blanz J, Zunke F, Markmann S, Damme M, Braulke T, Saftig P, et al. Mannose 6-phosphate-independent Lysosomal Sorting of LIMP-2. *Traffic.* 2015;16(10) :1127-36.
11. Ghosh P, Dahms NM, Kornfeld S. Mannose 6-phosphate receptors : new twists in the tale. *Nat Rev Mol Cell Biol.* 2003;4(3) :202-12.
12. Ghosh P, Griffith J, Geuze HJ, Kornfeld S. Mammalian GGAs act together to sort mannose 6-phosphate receptors. *J Cell Biol.* 2003;163(4) :755-66.
13. Park SY, Guo X. Adaptor protein complexes and intracellular transport. *Biosci Rep.* 2014;34(4).
14. Tong PY, Tollefsen SE, Kornfeld S. The cation-independent mannose 6-phosphate receptor binds insulin-like growth factor II. *J Biol Chem.* 1988;263(6) :2585-8.
15. Kang JX, Li Y, Leaf A. Mannose-6-phosphate/insulin-like growth factor-II receptor is a receptor for retinoic acid. *Proc Natl Acad Sci U S A.* 1997;94(25) :13671-6.
16. Motyka B, Korbitt G, Pinkoski MJ, Heibin JA, Caputo A, Hobman M, et al. Mannose 6-phosphate/insulin-like growth factor II receptor is a death receptor for granzyme B during cytotoxic T cell-induced apoptosis. *Cell.* 2000;103(3) :491-500.
17. Eskelinen EL. Roles of LAMP-1 and LAMP-2 in lysosome biogenesis and autophagy. *Mol Aspects Med.* 2006;27(5-6) :495-502.
18. Nishino I, Fu J, Tanji K, Yamada T, Shimojo S, Koori T, et al. Primary LAMP-2 deficiency causes X-linked vacuolar cardiomyopathy and myopathy (Danon disease). *Nature.* 2000;406(6798) :906-10.
19. Chi C, Leonard A, Knight WE, Beussman KM, Zhao Y, Cao Y, et al. LAMP-2B regulates human cardiomyocyte function by mediating autophagosome-lysosome fusion. *Proc Natl Acad Sci U S A.* 2019;116(2) :556-65.
20. Dibbens L, Schwake M, Saftig P, Rubboli G. SCARB2/LIMP2 deficiency in action myoclonus-renal failure syndrome. *Epileptic Disord.* 2016;18(S2) :63-72.
21. Gonzalez A, Valeiras M, Sidransky E, Tayebi N. Lysosomal integral membrane protein-2 : a new player in lysosome-related pathology. *Mol Genet Metab.* 2014;111(2) :84-91.
22. van Meel E, Klumperman J. Imaging and imagination : understanding the endo-lysosomal system. *Histochem Cell Biol.* 2008;129(3) :253-66.
23. De Duve C, Wattiaux R. Functions of lysosomes. *Annu Rev Physiol.* 1966;28 :435-92.

24. Dice JF. Peptide sequences that target cytosolic proteins for lysosomal proteolysis. *Trends Biochem Sci.* 1990;15(8) :305-9.
25. Rout AK, Strub MP, Piszczek G, Tjandra N. Structure of transmembrane domain of lysosome-associated membrane protein type 2a (LAMP-2A) reveals key features for substrate specificity in chaperone-mediated autophagy. *J Biol Chem.* 2014;289(51) :35111-23.
26. Orenstein SJ, Cuervo AM. Chaperone-mediated autophagy : molecular mechanisms and physiological relevance. *Semin Cell Dev Biol.* 2010;21(7) :719-26.
27. Agarraberes FA, Dice JF. A molecular chaperone complex at the lysosomal membrane is required for protein translocation. *J Cell Sci.* 2001;114(Pt 13) :2491-9.
28. Cuervo AM, Dice JF. Regulation of lamp2a levels in the lysosomal membrane. *Traffic.* 2000;1(7) :570-83.
29. Cuervo AM, Knecht E, Terlecky SR, Dice JF. Activation of a selective pathway of lysosomal proteolysis in rat liver by prolonged starvation. *Am J Physiol.* 1995;269(5 Pt 1) :C1200-8.
30. Kaushik S, Cuervo AM. Autophagy as a cell-repair mechanism : activation of chaperone-mediated autophagy during oxidative stress. *Mol Aspects Med.* 2006;27(5-6) :444-54.
31. Kiffin R, Kaushik S, Zeng M, Bandyopadhyay U, Zhang C, Massey AC, et al. Altered dynamics of the lysosomal receptor for chaperone-mediated autophagy with age. *J Cell Sci.* 2007;120(Pt 5) :782-91.
32. Zhou D, Li P, Lin Y, Lott JM, Hislop AD, Canaday DH, et al. Lamp-2a facilitates MHC class II presentation of cytoplasmic antigens. *Immunity.* 2005;22(5) :571-81.
33. Cuervo AM, Hildebrand H, Bomhard EM, Dice JF. Direct lysosomal uptake of alpha 2-microglobulin contributes to chemically induced nephropathy. *Kidney Int.* 1999;55(2) :529-45.
34. Sooparb S, Price SR, Shaoguang J, Franch HA. Suppression of chaperone-mediated autophagy in the renal cortex during acute diabetes mellitus. *Kidney Int.* 2004;65(6) :2135-44.
35. Ho PW, Leung CT, Liu H, Pang SY, Lam CS, Xian J, et al. Age-dependent accumulation of oligomeric SNCA/alpha-synuclein from impaired degradation in mutant LRRK2 knockin mouse model of Parkinson disease : role for therapeutic activation of chaperone-mediated autophagy (CMA). *Autophagy.* 2019 :1-24.
36. Nascimbeni AC, Fanin M, Angelini C, Sandri M. Autophagy dysregulation in Danon disease. *Cell Death Dis.* 2017;8(1) :e2565.
37. Venugopal B, Mesires NT, Kennedy JC, Curcio-Morelli C, Laplante JM, Dice JF, et al. Chaperone-mediated autophagy is defective in mucopolipidosis type IV. *J Cell Physiol.* 2009;219(2) :344-53.
38. Mijaljica D, Prescott M, Devenish RJ. Microautophagy in mammalian cells : revisiting a 40-year-old conundrum. *Autophagy.* 2011;7(7) :673-82.
39. Farre JC, Subramani S. Peroxisome turnover by micropexophagy : an autophagy-related process. *Trends Cell Biol.* 2004;14(9) :515-23.
40. Krick R, Muehe Y, Prick T, Bremer S, Schlotterhose P, Eskelinen EL, et al. Piecemeal microautophagy of the nucleus requires the core macroautophagy genes. *Mol Biol Cell.* 2008;19(10) :4492-505.
41. Lemasters JJ. Variants of mitochondrial autophagy : Types 1 and 2 mitophagy and micromitophagy (Type 3). *Redox Biol.* 2014;2 :749-54.
42. Wada Y, Sun-Wada GH, Kawamura N. Microautophagy in the visceral endoderm is essential for mouse early development. *Autophagy.* 2013;9(2) :252-4.
43. Takikita S, Myerowitz R, Schreiner C, Baum R, Raben N, Plotz PH. The values and limits of an in vitro model of Pompe disease : the best laid schemes o' mice an' men. *Autophagy.* 2009;5(5) :729-31.
44. Tsuji T, Fujimoto M, Tatematsu T, Cheng J, Orii M, Takatori S, et al. Niemann-Pick type C proteins promote microautophagy by expanding raft-like membrane domains in the yeast vacuole. *Elife.* 2017;6.

45. Levine B, Kroemer G. Biological Functions of Autophagy Genes : A Disease Perspective. *Cell*. 2019;176(1-2) :11-42.
46. Choi AM, Ryter SW, Levine B. Autophagy in human health and disease. *N Engl J Med*. 2013;368(7) :651-62.
47. Di Bartolomeo S, Corazzari M, Nazio F, Oliverio S, Lisi G, Antonioli M, et al. The dynamic interaction of AMBRA1 with the dynein motor complex regulates mammalian autophagy. *J Cell Biol*. 2010;191(1) :155-68.
48. Yu L, McPhee CK, Zheng L, Mardones GA, Rong Y, Peng J, et al. Termination of autophagy and reformation of lysosomes regulated by mTOR. *Nature*. 2010;465(7300) :942-6.
49. Pacheco CD, Lieberman AP. The pathogenesis of Niemann-Pick type C disease : a role for autophagy? *Expert Rev Mol Med*. 2008;10 :e26.
50. Ganley IG, Lam du H, Wang J, Ding X, Chen S, Jiang X. ULK1.ATG13.FIP200 complex mediates mTOR signaling and is essential for autophagy. *J Biol Chem*. 2009;284(18) :12297-305.
51. Chan EY. mTORC1 phosphorylates the ULK1-mAtg13-FIP200 autophagy regulatory complex. *Sci Signal*. 2009;2(84) :pe51.
52. Wong PM, Puente C, Ganley IG, Jiang X. The ULK1 complex : sensing nutrient signals for autophagy activation. *Autophagy*. 2013;9(2) :124-37.
53. Nazarko VY, Zhong Q. ULK1 targets Beclin-1 in autophagy. *Nat Cell Biol*. 2013;15(7) :727-8.
54. Fimia GM, Stoykova A, Romagnoli A, Giunta L, Di Bartolomeo S, Nardacci R, et al. Ambra1 regulates autophagy and development of the nervous system. *Nature*. 2007;447(7148) :1121-5.
55. Itakura E, Mizushima N. Atg14 and UVRAG : mutually exclusive subunits of mammalian Beclin 1-PI3K complexes. *Autophagy*. 2009;5(4) :534-6.
56. Pattingre S, Tassa A, Qu X, Garuti R, Liang XH, Mizushima N, et al. Bcl-2 antiapoptotic proteins inhibit Beclin 1-dependent autophagy. *Cell*. 2005;122(6) :927-39.
57. Luo S, Rubinsztein DC. Apoptosis blocks Beclin 1-dependent autophagosome synthesis : an effect rescued by Bcl-xL. *Cell Death Differ*. 2010;17(2) :268-77.
58. Zhong Y, Wang QJ, Li X, Yan Y, Backer JM, Chait BT, et al. Distinct regulation of autophagic activity by Atg14L and Rubicon associated with Beclin 1-phosphatidylinositol-3-kinase complex. *Nat Cell Biol*. 2009;11(4) :468-76.
59. Matsunaga K, Morita E, Saitoh T, Akira S, Ktistakis NT, Izumi T, et al. Autophagy requires endoplasmic reticulum targeting of the PI3-kinase complex via Atg14L. *J Cell Biol*. 2010;190(4) :511-21.
60. Mizushima N, Yoshimori T, Ohsumi Y. The role of Atg proteins in autophagosome formation. *Annu Rev Cell Dev Biol*. 2011;27 :107-32.
61. Mizushima N, Kuma A, Kobayashi Y, Yamamoto A, Matsubae M, Takao T, et al. Mouse Apg16L, a novel WD-repeat protein, targets to the autophagic isolation membrane with the Apg12-Apg5 conjugate. *J Cell Sci*. 2003;116(Pt 9) :1679-88.
62. Kabeya Y, Mizushima N, Ueno T, Yamamoto A, Kirisako T, Noda T, et al. LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *EMBO J*. 2000;19(21) :5720-8.
63. Tanida I, Ueno T, Kominami E. Human light chain 3/MAP1LC3B is cleaved at its carboxyl-terminal Met121 to expose Gly120 for lipidation and targeting to autophagosomal membranes. *J Biol Chem*. 2004;279(46) :47704-10.
64. Ward C, Martinez-Lopez N, Otten EG, Carroll B, Maetzel D, Singh R, et al. Autophagy, lipophagy and lysosomal lipid storage disorders. *Biochim Biophys Acta*. 2016;1861(4) :269-84.
65. Hyttinen JM, Niittykoski M, Salminen A, Kaarniranta K. Maturation of autophagosomes and endosomes : a key role for Rab7. *Biochim Biophys Acta*. 2013;1833(3) :503-10.
66. Feng Y, He D, Yao Z, Klionsky DJ. The machinery of macroautophagy. *Cell Res*. 2014;24(1) :24-41.
67. Mizushima N, Levine B. Autophagy in mammalian development and differentiation. *Nat Cell Biol*. 2010;12(9) :823-30.

68. Luzio JP, Hackmann Y, Dieckmann NM, Griffiths GM. The biogenesis of lysosomes and lysosome-related organelles. *Cold Spring Harb Perspect Biol.* 2014;6(9) :a016840.
69. Kim SG, Buel GR, Blenis J. Nutrient regulation of the mTOR complex 1 signaling pathway. *Mol Cells.* 2013;35(6) :463-73.
70. Sengupta S, Peterson TR, Sabatini DM. Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. *Mol Cell.* 2010;40(2) :310-22.
71. Yan Y, Zhou XE, Xu HE, Melcher K. Structure and Physiological Regulation of AMPK. *Int J Mol Sci.* 2018;19(11).
72. Kim J, Kundu M, Viollet B, Guan KL. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nat Cell Biol.* 2011;13(2) :132-41.
73. Choi AM, Ryter SW, Levine B. Autophagy in human health and disease. *N Engl J Med.* 2013;368(19) :1845-6.
74. Wang K, Klionsky DJ. Mitochondria removal by autophagy. *Autophagy.* 2011;7(3) :297-300.
75. Lamark T, Johansen T. Aggrephagy : selective disposal of protein aggregates by macroautophagy. *Int J Cell Biol.* 2012;2012 :736905.
76. Tsukamoto S, Kuma A, Murakami M, Kishi C, Yamamoto A, Mizushima N. Autophagy is essential for preimplantation development of mouse embryos. *Science.* 2008;321(5885) :117-20.
77. Yue Z, Jin S, Yang C, Levine AJ, Heintz N. Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor. *Proc Natl Acad Sci U S A.* 2003;100(25) :15077-82.
78. Antonoli M, Albiero F, Fimia GM, Piacentini M. AMBRA1-regulated autophagy in vertebrate development. *Int J Dev Biol.* 2015;59(1-3) :109-17.
79. Kuma A, Hatano M, Matsui M, Yamamoto A, Nakaya H, Yoshimori T, et al. The role of autophagy during the early neonatal starvation period. *Nature.* 2004;432(7020) :1032-6.
80. Yen WL, Klionsky DJ. How to live long and prosper : autophagy, mitochondria, and aging. *Physiology (Bethesda).* 2008;23 :248-62.
81. Simonsen A, Cumming RC, Brech A, Isakson P, Schubert DR, Finley KD. Promoting basal levels of autophagy in the nervous system enhances longevity and oxidant resistance in adult *Drosophila*. *Autophagy.* 2008;4(2) :176-84.
82. Booth LA, Tavallai S, Hamed HA, Cruickshanks N, Dent P. The role of cell signalling in the crosstalk between autophagy and apoptosis. *Cell Signal.* 2014;26(3) :549-55.
83. Yu L, Lenardo MJ, Baehrecke EH. Autophagy and caspases : a new cell death program. *Cell Cycle.* 2004;3(9) :1124-6.
84. Yu L, Alva A, Su H, Dutt P, Freundt E, Welsh S, et al. Regulation of an ATG7-beclin 1 program of autophagic cell death by caspase-8. *Science.* 2004;304(5676) :1500-2.
85. Shimizu S, Kanaseki T, Mizushima N, Mizuta T, Arakawa-Kobayashi S, Thompson CB, et al. Role of Bcl-2 family proteins in a non-apoptotic programmed cell death dependent on autophagy genes. *Nat Cell Biol.* 2004;6(12) :1221-8.
86. Lepine S, Allegood JC, Edmonds Y, Milstien S, Spiegel S. Autophagy induced by deficiency of sphingosine-1-phosphate phosphohydrolase 1 is switched to apoptosis by calpain-mediated autophagy-related gene 5 (Atg5) cleavage. *J Biol Chem.* 2011;286(52) :44380-90.
87. Liu Y, Levine B. Autosis and autophagic cell death : the dark side of autophagy. *Cell Death Differ.* 2015;22(3) :367-76.
88. Patterson NL, Mintern JD. Intersection of autophagy with pathways of antigen presentation. *Protein Cell.* 2012;3(12) :911-20.
89. Sharma V, Verma S, Seranova E, Sarkar S, Kumar D. Selective Autophagy and Xenophagy in Infection and Disease. *Front Cell Dev Biol.* 2018;6 :147.
90. Levine B, Deretic V. Unveiling the roles of autophagy in innate and adaptive immunity. *Nat Rev Immunol.* 2007;7(10) :767-77.
91. Jellinger KA. Basic mechanisms of neurodegeneration : a critical update. *J Cell Mol Med.* 2010;14(3) :457-87.

92. Wong E, Cuervo AM. Autophagy gone awry in neurodegenerative diseases. *Nat Neurosci.* 2010;13(7) :805-11.
93. Ma JF, Huang Y, Chen SD, Halliday G. Immunohistochemical evidence for macroautophagy in neurones and endothelial cells in Alzheimer's disease. *Neuropathol Appl Neurobiol.* 2010;36(4) :312-9.
94. Komatsu M, Waguri S, Chiba T, Murata S, Iwata J, Tanida I, et al. Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature.* 2006;441(7095) :880-4.
95. Schaeffer V, Lavenir I, Ozcelik S, Tolnay M, Winkler DT, Goedert M. Stimulation of autophagy reduces neurodegeneration in a mouse model of human tauopathy. *Brain.* 2012;135(Pt 7) :2169-77.
96. Ravikumar B, Vacher C, Berger Z, Davies JE, Luo S, Oroz LG, et al. Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nat Genet.* 2004;36(6) :585-95.
97. Oh TS, Cho H, Cho JH, Yu SW, Kim EK. Hypothalamic AMPK-induced autophagy increases food intake by regulating NPY and POMC expression. *Autophagy.* 2016;12(11) :2009-25.
98. Chevrier M, Brakch N, Celine L, Genty D, Ramdani Y, Moll S, et al. Autophagosome maturation is impaired in Fabry disease. *Autophagy.* 2010;6(5) :589-99.
99. Kinghorn KJ, Asghari AM, Castillo-Quan JI. The emerging role of autophagic-lysosomal dysfunction in Gaucher disease and Parkinson's disease. *Neural Regen Res.* 2017;12(3) :380-4.
100. Sun Y, Grabowski GA. Impaired autophagosomes and lysosomes in neuronopathic Gaucher disease. *Autophagy.* 2010;6(5) :648-9.
101. Fukuda T, Roberts A, Ahearn M, Zaal K, Ralston E, Plotz PH, et al. Autophagy and lysosomes in Pompe disease. *Autophagy.* 2006;2(4) :318-20.
102. Takamura A, Higaki K, Kajimaki K, Otsuka S, Ninomiya H, Matsuda J, et al. Enhanced autophagy and mitochondrial aberrations in murine G(M1)-gangliosidosis. *Biochem Biophys Res Commun.* 2008;367(3) :616-22.
103. Settembre C, Fraldi A, Jahreiss L, Spampinato C, Venturi C, Medina D, et al. A block of autophagy in lysosomal storage disorders. *Hum Mol Genet.* 2008;17(1) :119-29.
104. Settembre C, Fraldi A, Rubinsztein DC, Ballabio A. Lysosomal storage diseases as disorders of autophagy. *Autophagy.* 2008;4(1) :113-4.
105. Jennings JJ, Jr., Zhu JH, Rbaibi Y, Luo X, Chu CT, Kiselyov K. Mitochondrial aberrations in mucopolipidosis Type IV. *J Biol Chem.* 2006;281(51) :39041-50.
106. Koike M, Shibata M, Waguri S, Yoshimura K, Tanida I, Kominami E, et al. Participation of autophagy in storage of lysosomes in neurons from mouse models of neuronal ceroid-lipofuscinoses (Batten disease). *Am J Pathol.* 2005;167(6) :1713-28.
107. Cao Y, Espinola JA, Fossale E, Massey AC, Cuervo AM, MacDonald ME, et al. Autophagy is disrupted in a knock-in mouse model of juvenile neuronal ceroid lipofuscinosis. *J Biol Chem.* 2006;281(29) :20483-93.
108. Canonico B, Cesarini E, Salucci S, Luchetti F, Falcieri E, Di Sario G, et al. Defective Autophagy, Mitochondrial Clearance and Lipophagy in Niemann-Pick Type B Lymphocytes. *PLoS One.* 2016;11(10) :e0165780.
109. Ko DC, Milenkovic L, Beier SM, Manuel H, Buchanan J, Scott MP. Cell-autonomous death of cerebellar purkinje neurons with autophagy in Niemann-Pick type C disease. *PLoS Genet.* 2005;1(1) :81-95.
110. Tanaka Y, Guhde G, Suter A, Eskelinen EL, Hartmann D, Lullmann-Rauch R, et al. Accumulation of autophagic vacuoles and cardiomyopathy in LAMP-2-deficient mice. *Nature.* 2000;406(6798) :902-6.
111. Platt FM, d'Azzo A, Davidson BL, Neufeld EF, Tiffet CJ. Lysosomal storage diseases. *Nat Rev Dis Primers.* 2018;4(1) :27.
112. Ballabio A, Gieselmann V. Lysosomal disorders : from storage to cellular damage. *Biochim Biophys Acta.* 2009;1793(4) :684-96.

113. Fuller M, Meikle PJ, Hopwood JJ. Epidemiology of lysosomal storage diseases : an overview. In : Mehta A, Beck M, Sunder-Plassmann G, editors. *Fabry Disease : Perspectives from 5 Years of FOS*. Oxford 2006.
114. Gahl WA. Disorders of lysosomal membrane transport--cystinosis and Salla disease. *Enzyme*. 1987;38(1-4) :154-60.
115. Feldhammer M, Durand S, Mrazova L, Boucher RM, Laframboise R, Steinfeld R, et al. Sanfilippo syndrome type C : mutation spectrum in the heparan sulfate acetyl-CoA : alpha-glucosaminide N-acetyltransferase (HGSNAT) gene. *Hum Mutat*. 2009;30(6) :918-25.
116. Miedel MT, Rbaibi Y, Guerriero CJ, Colletti G, Weixel KM, Weisz OA, et al. Membrane traffic and turnover in TRP-ML1-deficient cells : a revised model for mucopolysaccharidosis type IV pathogenesis. *J Exp Med*. 2008;205(6) :1477-90.
117. Platt FM, Boland B, van der Spoel AC. The cell biology of disease : lysosomal storage disorders : the cellular impact of lysosomal dysfunction. *J Cell Biol*. 2012;199(5) :723-34.
118. Nixon RA, Yang DS, Lee JH. Neurodegenerative lysosomal disorders : a continuum from development to late age. *Autophagy*. 2008;4(5) :590-9.
119. Vitner EB, Platt FM, Futerman AH. Common and uncommon pathogenic cascades in lysosomal storage diseases. *J Biol Chem*. 2010;285(27) :20423-7.
120. Valayannopoulos V, Wijburg FA. Therapy for the mucopolysaccharidoses. *Rheumatology (Oxford)*. 2011;50 Suppl 5 :v49-59.
121. Parenti G, Andria G, Ballabio A. Lysosomal storage diseases : from pathophysiology to therapy. *Annu Rev Med*. 2015;66 :471-86.
122. Valayannopoulos V, Brassier A, Chabli A, Caillaud C, Lemoine M, Odent T, et al. [Enzyme replacement therapy for lysosomal storage disorders]. *Arch Pediatr*. 2011;18(10) :1119-23.
123. Muenzer J. Early initiation of enzyme replacement therapy for the mucopolysaccharidoses. *Mol Genet Metab*. 2014;111(2) :63-72.
124. Pastores GM. Recombinant glucocerebrosidase (imiglucerase) as a therapy for Gaucher disease. *BioDrugs*. 2010;24(1) :41-7.
125. Aldenhoven M, Wynn RF, Orchard PJ, O'Meara A, Veys P, Fischer A, et al. Long-term outcome of Hurler syndrome patients after hematopoietic cell transplantation : an international multicenter study. *Blood*. 2015;125(13) :2164-72.
126. Kim SU. Lysosomal storage diseases : Stem cell-based cell- and gene-therapy. *Cell Transplant*. 2014.
127. Giraldo P, Andrade-Campos M, Alfonso P, Irun P, Atutxa K, Acedo A, et al. Twelve years of experience with miglustat in the treatment of type 1 Gaucher disease : The Spanish ZAGAL project. *Blood Cells Mol Dis*. 2018;68 :173-9.
128. Patterson MC, Vecchio D, Prady H, Abel L, Wraith JE. Miglustat for treatment of Niemann-Pick C disease : a randomised controlled study. *Lancet Neurol*. 2007;6(9) :765-72.
129. Belmatoug N, Burlina A, Giraldo P, Hendriksz CJ, Kuter DJ, Mengel E, et al. Gastrointestinal disturbances and their management in miglustat-treated patients. *J Inher Metab Dis*. 2011;34(5) :991-1001.
130. Hoseki J, Ushioda R, Nagata K. Mechanism and components of endoplasmic reticulum-associated degradation. *J Biochem*. 2010;147(1) :19-25.
131. Suzuki Y, Ichinomiya S, Kurosawa M, Matsuda J, Ogawa S, Iida M, et al. Therapeutic chaperone effect of N-octyl 4-epi-beta-valienamine on murine G(M1)-gangliosidosis. *Mol Genet Metab*. 2012;106(1) :92-8.
132. Lukas J, Knospe AM, Seemann S, Citro V, Cubellis MV, Rolfs A. In Vitro Enzyme Measurement to Test Pharmacological Chaperone Responsiveness in Fabry and Pompe Disease. *J Vis Exp*. 2017(130).
133. Muntze J, Gensler D, Maniuc O, Liu D, Cairns T, Oder D, et al. Oral Chaperone Therapy Migalastat for Treating Fabry Disease : Enzymatic Response and Serum Biomarker Changes After 1 Year. *Clin Pharmacol Ther*. 2019;105(5) :1224-33.

134. Suzuki Y. Chaperone therapy update : Fabry disease, GM1-gangliosidosis and Gaucher disease. *Brain Dev.* 2013;35(6) :515-23.
135. Scarpa M, Orchard PJ, Schulz A, Dickson PI, Haskins ME, Escolar ML, et al. Treatment of brain disease in the mucopolysaccharidoses. *Mol Genet Metab.* 2017;122S :25-34.
136. Mole SE, Anderson G, Band HA, Berkovic SF, Cooper JD, Kleine Holthaus SM, et al. Clinical challenges and future therapeutic approaches for neuronal ceroid lipofuscinosis. *Lancet Neurol.* 2019;18(1) :107-16.
137. Muenzer J. Overview of the mucopolysaccharidoses. *Rheumatology (Oxford).* 2011;50 Suppl 5 :v4-12.
138. Bach G, Friedman R, Weissmann B, Neufeld EF. The defect in the Hurler and Scheie syndromes : deficiency of -L-iduronidase. *Proc Natl Acad Sci U S A.* 1972;69(8) :2048-51.
139. Clarke LA. Mucopolysaccharidosis Type I. In : Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LH, Stephens K, et al., editors. *GeneReviews*((R)). Seattle (WA)1993.
140. Scheie HG, Hambrick GW, Jr., Barness LA. A newly recognized forme fruste of Hurler's disease (gargoylism). *Am J Ophthalmol.* 1962;53 :753-69.
141. Maroteaux P. [Diagnosis of Hurler's disease in the laboratory]. *Ann Genet.* 1966;9(1) :42-6.
142. Giugliani R, Federhen A, Rojas MV, Vieira T, Artigalás O, Pinto LL, et al. Mucopolysaccharidosis I, II, and VI : Brief review and guidelines for treatment. *Genet Mol Biol.* 2010;33(4) :589-604.
143. Aldenhoven M, Boelens JJ, de Koning TJ. The clinical outcome of Hurler syndrome after stem cell transplantation. *Biol Blood Marrow Transplant.* 2008;14(5) :485-98.
144. Peters C, Shapiro EG, Krivit W. Neuropsychological development in children with Hurler syndrome following hematopoietic stem cell transplantation. *Pediatr Transplant.* 1998;2(4) :250-3.
145. Lum SH, Stepien KM, Ghosh A, Broomfield A, Church H, Mercer J, et al. Long term survival and cardiopulmonary outcome in children with Hurler syndrome after haematopoietic stem cell transplantation. *J Inher Metab Dis.* 2017;40(3) :455-60.
146. Brooks DA. Alpha-L-iduronidase and enzyme replacement therapy for mucopolysaccharidosis I. *Expert Opin Biol Ther.* 2002;2(8) :967-76.
147. Jameson E, Jones S, Wraith JE. Enzyme replacement therapy with laronidase (Aldurazyme) for treating mucopolysaccharidosis type I. *Cochrane Database Syst Rev.* 2013(9) :CD009354.
148. Begley DJ, Pontikis CC, Scarpa M. Lysosomal storage diseases and the blood-brain barrier. *Curr Pharm Des.* 2008;14(16) :1566-80.
149. Muenzer J, Beck M, Eng CM, Giugliani R, Harmatz P, Martin R, et al. Long-term, open-labeled extension study of idursulfase in the treatment of Hunter syndrome. *Genet Med.* 2011;13(2) :95-101.
150. Martin R, Beck M, Eng C, Giugliani R, Harmatz P, Munoz V, et al. Recognition and diagnosis of mucopolysaccharidosis II (Hunter syndrome). *Pediatrics.* 2008;121(2) :e377-86.
151. Muenzer J, Bodamer O, Burton B, Clarke L, Frenking GS, Giugliani R, et al. The role of enzyme replacement therapy in severe Hunter syndrome-an expert panel consensus. *Eur J Pediatr.* 2012;171(1) :181-8.
152. Muenzer J, Wraith JE, Beck M, Giugliani R, Harmatz P, Eng CM, et al. A phase II/III clinical study of enzyme replacement therapy with idursulfase in mucopolysaccharidosis II (Hunter syndrome). *Genet Med.* 2006;8(8) :465-73.
153. da Silva EM, Strufaldi MW, Andriolo RB, Silva LA. Enzyme replacement therapy with idursulfase for mucopolysaccharidosis type II (Hunter syndrome). *Cochrane Database Syst Rev.* 2016;2 :CD008185.
154. Giugliani R, Villarreal ML, Valdez CA, Hawilou AM, Guelbert N, Garzon LN, et al. Guidelines for diagnosis and treatment of Hunter Syndrome for clinicians in Latin America. *Genet Mol Biol.* 2014;37(2) :315-29.
155. Bradley LA, Haddow HRM, Palomaki GE. Treatment of mucopolysaccharidosis type II (Hunter syndrome) : results from a systematic evidence review. *Genet Med.* 2017;19(11) :1187-201.

156. Muenzer J, Beck M, Eng CM, Escolar ML, Giugliani R, Guffon NH, et al. Multidisciplinary management of Hunter syndrome. *Pediatrics*. 2009;124(6) :e1228-39.
157. Andrade F, Aldamiz-Echevarria L, Llarena M, Couce ML. Sanfilippo syndrome : Overall review. *Pediatr Int*. 2015;57(3) :331-8.
158. Cross EM, Hare DJ. Behavioural phenotypes of the mucopolysaccharide disorders : a systematic literature review of cognitive, motor, social, linguistic and behavioural presentation in the MPS disorders. *J Inherit Metab Dis*. 2013;36(2) :189-200.
159. Valstar MJ, Ruijter GJ, van Diggelen OP, Poorthuis BJ, Wijburg FA. Sanfilippo syndrome : a mini-review. *J Inherit Metab Dis*. 2008;31(2) :240-52.
160. Lavery C, Hendriksz CJ, Jones SA. Mortality in patients with Sanfilippo syndrome. *Orphanet J Rare Dis*. 2017;12(1) :168.
161. Fedele AO. Sanfilippo syndrome : causes, consequences, and treatments. *Appl Clin Genet*. 2015;8 :269-81.
162. The classics : Chondro-osteo-dystrophy. roentgenographic and clinical features of a child with dislocation of vertebrae, James F. Brailsford, M.D. : *Am. J. Surg.* 7 :404, 1929. *Clin Orthop Relat Res*. 1976(114) :4-9.
163. Hendriksz CJ, Harmatz P, Beck M, Jones S, Wood T, Lachman R, et al. Review of clinical presentation and diagnosis of mucopolysaccharidosis IVA. *Mol Genet Metab*. 2013;110(1-2) :54-64.
164. van der Horst GT, Kleijer WJ, Hoogeveen AT, Huijmans JG, Blom W, van Diggelen OP. Morquio B syndrome : a primary defect in beta-galactosidase. *Am J Med Genet*. 1983;16(2) :261-75.
165. Baujat G, Valayannopoulos V. [Natural history of Morquio A disease]. *Arch Pediatr*. 2014;21 Suppl 1 :S32-8.
166. Caciotti A, Garman SC, Rivera-Colon Y, Procopio E, Catarzi S, Ferri L, et al. GM1 gangliosidosis and Morquio B disease : an update on genetic alterations and clinical findings. *Biochim Biophys Acta*. 2011;1812(7) :782-90.
167. Hendriksz CJ, Berger KI, Giugliani R, Harmatz P, Kampmann C, Mackenzie WG, et al. International guidelines for the management and treatment of Morquio A syndrome. *Am J Med Genet A*. 2015;167A(1) :11-25.
168. Hendriksz CJ. Elosulfase alfa (BMN 110) for the treatment of mucopolysaccharidosis IVA (Morquio A Syndrome). *Expert Rev Clin Pharmacol*. 2016;9(12) :1521-32.
169. Hendriksz CJ, Burton B, Fleming TR, Harmatz P, Hughes D, Jones SA, et al. Efficacy and safety of enzyme replacement therapy with BMN 110 (elosulfase alfa) for Morquio A syndrome (mucopolysaccharidosis IVA) : a phase 3 randomised placebo-controlled study. *J Inherit Metab Dis*. 2014;37(6) :979-90.
170. Jones SA, Bialer M, Parini R, Martin K, Wang H, Yang K, et al. Safety and clinical activity of elosulfase alfa in pediatric patients with Morquio A syndrome (mucopolysaccharidosis IVA) less than 5 y. *Pediatr Res*. 2015;78(6) :717-22.
171. Maroteaux P, Leveque B, Marie J, Lamy M. [a New Dysostosis with Urinary Elimination of Chondroitin Sulfate B]. *Presse Med*. 1963;71 :1849-52.
172. Baehner F, Schmiedeskamp C, Krummenauer F, Miebach E, Bajbouj M, Whybra C, et al. Cumulative incidence rates of the mucopolysaccharidoses in Germany. *J Inherit Metab Dis*. 2005;28(6) :1011-7.
173. Malm G, Lund AM, Mansson JE, Heiberg A. Mucopolysaccharidoses in the Scandinavian countries : incidence and prevalence. *Acta Paediatr*. 2008;97(11) :1577-81.
174. Valayannopoulos V, Nicely H, Harmatz P, Turbeville S. Mucopolysaccharidosis VI. *Orphanet J Rare Dis*. 2010;5 :5.
175. Giugliani R, Harmatz P, Wraith JE. Management guidelines for mucopolysaccharidosis VI. *Pediatrics*. 2007;120(2) :405-18.
176. Harmatz P, Yu ZF, Giugliani R, Schwartz IV, Guffon N, Teles EL, et al. Enzyme replacement therapy for mucopolysaccharidosis VI : evaluation of long-term pulmonary function in patients

- treated with recombinant human N-acetylgalactosamine 4-sulfatase. *J Inher Metab Dis.* 2010;33(1) :51-60.
177. Decker C, Yu ZF, Giugliani R, Schwartz IV, Guffon N, Teles EL, et al. Enzyme replacement therapy for mucopolysaccharidosis VI : Growth and pubertal development in patients treated with recombinant human N-acetylgalactosamine 4-sulfatase. *J Pediatr Rehabil Med.* 2010;3(2) :89-100.
178. Krivit W, Pierpont ME, Ayaz K, Tsai M, Ramsay NK, Kersey JH, et al. Bone-marrow transplantation in the Maroteaux-Lamy syndrome (mucopolysaccharidosis type VI). Biochemical and clinical status 24 months after transplantation. *N Engl J Med.* 1984;311(25) :1606-11.
179. Wang CC, Hwu WL, Lin KH. Long-term follow-up of a girl with Maroteaux-Lamy syndrome after bone marrow transplantation. *World J Pediatr.* 2008;4(2) :152-4.
180. J F. Ein Eintrag zur Kenntnis der Purpura haemorrhagica nodularis (Purpura papulosa haemorrhagica Hebrae). *Arch Derm Syphil.* 1898;43 :187-200.
181. Anderson W. A case of angio-keratoma. *Brit J Derm.* 1898;10 :113-7.
182. Meikle PJ, Hopwood JJ, Clague AE, Carey WF. Prevalence of lysosomal storage disorders. *JAMA.* 1999;281(3) :249-54.
183. Spada M, Pagliardini S, Yasuda M, Tukul T, Thiagarajan G, Sakuraba H, et al. High incidence of later-onset fabry disease revealed by newborn screening. *Am J Hum Genet.* 2006;79(1) :31-40.
184. Garman SC, Garboczi DN. The molecular defect leading to Fabry disease : structure of human alpha-galactosidase. *J Mol Biol.* 2004;337(2) :319-35.
185. Oliveira JP, Ferreira S. Multiple phenotypic domains of Fabry disease and their relevance for establishing genotype- phenotype correlations. *Appl Clin Genet.* 2019;12 :35-50.
186. Clarke JT, Stoltz JM, Mulcahey MR. Neutral glycosphingolipids of serum lipoproteins in Fabry's disease. *Biochim Biophys Acta.* 1976;431(2) :317-25.
187. Moore DF, Scott LT, Gladwin MT, Altarescu G, Kaneski C, Suzuki K, et al. Regional cerebral hyperperfusion and nitric oxide pathway dysregulation in Fabry disease : reversal by enzyme replacement therapy. *Circulation.* 2001;104(13) :1506-12.
188. Moore DF, Ye F, Brennan ML, Gupta S, Barshop BA, Steiner RD, et al. Ascorbate decreases Fabry cerebral hyperperfusion suggesting a reactive oxygen species abnormality : an arterial spin tagging study. *J Magn Reson Imaging.* 2004;20(4) :674-83.
189. Schiffmann R, Ries M. Fabry Disease : A Disorder of Childhood Onset. *Pediatr Neurol.* 2016;64 :10-20.
190. Aerts JM, Groener JE, Kuiper S, Donker-Koopman WE, Strijland A, Ottenhoff R, et al. Elevated globotriaosylsphingosine is a hallmark of Fabry disease. *Proc Natl Acad Sci U S A.* 2008;105(8) :2812-7.
191. Elleder M. Sequelae of storage in Fabry disease--pathology and comparison with other lysosomal storage diseases. *Acta Paediatr Suppl.* 2003;92(443) :46-53; discussion 45.
192. Schiffmann R, Rapkiewicz A, Abu-Asab M, Ries M, Askari H, Tsokos M, et al. Pathological findings in a patient with Fabry disease who died after 2.5 years of enzyme replacement. *Virchows Arch.* 2006;448(3) :337-43.
193. Massi D, Martinelli F, Battini ML, Comin CE, Franchi A, Gioia O, et al. Angiokeratoma corporis diffusum (Anderson-Fabry's disease) : a case report. *J Eur Acad Dermatol Venereol.* 2000;14(2) :127-30.
194. Alroy J, Sabnis S, Kopp JB. Renal pathology in Fabry disease. *J Am Soc Nephrol.* 2002;13 Suppl 2 :S134-8.
195. Okeda R, Nishihara M. An autopsy case of Fabry disease with neuropathological investigation of the pathogenesis of associated dementia. *Neuropathology.* 2008;28(5) :532-40.
196. Nelson MP, Tse TE, O'Quinn DB, Percival SM, Jaimes EA, Warnock DG, et al. Autophagy-lysosome pathway associated neuropathology and axonal degeneration in the brains of alpha-galactosidase A-deficient mice. *Acta Neuropathol Commun.* 2014;2 :20.

197. Liebau MC, Braun F, Hopker K, Weitbrecht C, Bartels V, Muller RU, et al. Dysregulated autophagy contributes to podocyte damage in Fabry's disease. *PLoS One*. 2013;8(5) :e63506.
198. Yanagisawa H, Hossain MA, Miyajima T, Nagao K, Miyashita T, Eto Y. Dysregulated DNA methylation of GLA gene was associated with dysfunction of autophagy. *Mol Genet Metab*. 2019.
199. Biancini GB, Jacques CE, Hammerschmidt T, de Souza HM, Donida B, Deon M, et al. Biomolecules damage and redox status abnormalities in Fabry patients before and during enzyme replacement therapy. *Clin Chim Acta*. 2016;461 :41-6.
200. Shen JS, Meng XL, Moore DF, Quirk JM, Shayman JA, Schiffmann R, et al. Globotriaosylceramide induces oxidative stress and up-regulates cell adhesion molecule expression in Fabry disease endothelial cells. *Mol Genet Metab*. 2008;95(3) :163-8.
201. Chimenti C, Scopelliti F, Vulpis E, Tafani M, Villanova L, Verardo R, et al. Increased oxidative stress contributes to cardiomyocyte dysfunction and death in patients with Fabry disease cardiomyopathy. *Hum Pathol*. 2015;46(11) :1760-8.
202. Rozenfeld P, Feriozzi S. Contribution of inflammatory pathways to Fabry disease pathogenesis. *Mol Genet Metab*. 2017;122(3) :19-27.
203. De Francesco PN, Mucci JM, Ceci R, Fossati CA, Rozenfeld PA. Fabry disease peripheral blood immune cells release inflammatory cytokines : role of globotriaosylceramide. *Mol Genet Metab*. 2013;109(1) :93-9.
204. Mauhin W, Lidove O, Masat E, Mingozzi F, Mariampillai K, Ziza JM, et al. Innate and Adaptive Immune Response in Fabry Disease. *JIMD Rep*. 2015;22 :1-10.
205. Lee MH, Choi EN, Jeon YJ, Jung SC. Possible role of transforming growth factor-beta1 and vascular endothelial growth factor in Fabry disease nephropathy. *Int J Mol Med*. 2012;30(6) :1275-80.
206. Kofeler HC, Fauland A, Rechberger GN, Trotschmuller M. Mass spectrometry based lipidomics : an overview of technological platforms. *Metabolites*. 2012;2(1) :19-38.
207. Rombach SM, Twickler TB, Aerts JM, Linthorst GE, Wijburg FA, Hollak CE. Vasculopathy in patients with Fabry disease : current controversies and research directions. *Mol Genet Metab*. 2010;99(2) :99-108.
208. Rigante D, Cipolla C, Basile U, Gulli F, Savastano MC. Overview of immune abnormalities in lysosomal storage disorders. *Immunol Lett*. 2017;188 :79-85.
209. DeGraba T, Azhar S, Dignat-George F, Brown E, Boutiere B, Altarescu G, et al. Profile of endothelial and leukocyte activation in Fabry patients. *Ann Neurol*. 2000;47(2) :229-33.
210. Sakuraba H, Igarashi T, Shibata T, Suzuki Y. Effect of vitamin E and ticlopidine on platelet aggregation in Fabry's disease. *Clin Genet*. 1987;31(5) :349-54.
211. Vedder AC, Biro E, Aerts JM, Nieuwland R, Sturk G, Hollak CE. Plasma markers of coagulation and endothelial activation in Fabry disease : impact of renal impairment. *Nephrol Dial Transplant*. 2009;24(10) :3074-81.
212. Shen JS, Meng XL, Wight-Carter M, Day TS, Goetsch SC, Forni S, et al. Blocking hyperactive androgen receptor signaling ameliorates cardiac and renal hypertrophy in Fabry mice. *Hum Mol Genet*. 2015;24(11) :3181-91.
213. Bertelsen AK, Tondel C, Krohn J, Bull N, Aarseth J, Houge G, et al. Small fibre neuropathy in Fabry disease. *J Neurol*. 2013;260(3) :917-9.
214. Mehta A, Ricci R, Widmer U, Dehout F, Garcia de Lorenzo A, Kampmann C, et al. Fabry disease defined : baseline clinical manifestations of 366 patients in the Fabry Outcome Survey. *Eur J Clin Invest*. 2004;34(3) :236-42.
215. Politei JM, Pagano MA. [Peripheral neuropathy in Anderson-Fabry disease : its physiology, evaluation and treatment]. *Rev Neurol*. 2004;38(10) :979-83.
216. Zarate YA, Hopkin RJ. Fabry's disease. *Lancet*. 2008;372(9647) :1427-35.
217. Politei JM, Bouhassira D, Germain DP, Goizet C, Guerrero-Sola A, Hilz MJ, et al. Pain in Fabry Disease : Practical Recommendations for Diagnosis and Treatment. *CNS Neurosci Ther*. 2016;22(7) :568-76.

218. Koping M, Shehata-Dieler W, Schneider D, Cebulla M, Oder D, Muntze J, et al. Characterization of vertigo and hearing loss in patients with Fabry disease. *Orphanet J Rare Dis*. 2018;13(1) :137.
219. Mehta A, Ginsberg L, Investigators FOS. Natural history of the cerebrovascular complications of Fabry disease. *Acta Paediatr Suppl*. 2005;94(447) :24-7; discussion 9-10.
220. Sims K, Politei J, Banikazemi M, Lee P. Stroke in Fabry disease frequently occurs before diagnosis and in the absence of other clinical events : natural history data from the Fabry Registry. *Stroke*. 2009;40(3) :788-94.
221. Mitsias P, Levine SR. Cerebrovascular complications of Fabry's disease. *Ann Neurol*. 1996;40(1) :8-17.
222. Moore DF, Altarescu G, Barker WC, Patronas NJ, Herscovitch P, Schiffmann R. White matter lesions in Fabry disease occur in 'prior' selectively hypometabolic and hyperperfused brain regions. *Brain Res Bull*. 2003;62(3) :231-40.
223. Yousef Z, Elliott PM, Cecchi F, Escoubet B, Linhart A, Monserrat L, et al. Left ventricular hypertrophy in Fabry disease : a practical approach to diagnosis. *Eur Heart J*. 2013;34(11) :802-8.
224. Nakayama Y, Tsumura K, Yamashita N, Yoshimaru K. Dynamic left ventricular arterial pressure gradient and sick sinus syndrome with heterozygous Fabry's disease improved following implantation of a dual chamber pacemaker. *Pacing Clin Electrophysiol*. 1999;22(7) :1114-5.
225. Linhart A, Elliott PM. The heart in Anderson-Fabry disease and other lysosomal storage disorders. *Heart*. 2007;93(4) :528-35.
226. Nair V, Belanger EC, Veinot JP. Lysosomal storage disorders affecting the heart : a review. *Cardiovasc Pathol*. 2019;39 :12-24.
227. Lidove O, Bekri S, Goizet C, Khau Van Kien A, Aractingi S, Knebelmann B, et al. [Fabry disease : proposed guidelines from a French expert group for its diagnosis, treatment and follow-up]. *Presse Med*. 2007;36(7-8) :1084-97.
228. Stephan F, Haber R. [Fabry disease]. *Ann Dermatol Venereol*. 2017;144(2) :137-46.
229. Orteu CH, Jansen T, Lidove O, Jaussaud R, Hughes DA, Pintos-Morell G, et al. Fabry disease and the skin : data from FOS, the Fabry outcome survey. *Br J Dermatol*. 2007;157(2) :331-7.
230. Zampetti A, Orteu CH, Antuzzi D, Bongiorno MR, Manco S, Gnarra M, et al. Angiokeratoma : decision-making aid for the diagnosis of Fabry disease. *Br J Dermatol*. 2012;166(4) :712-20.
231. Lidove O, Ramaswami U, Jaussaud R, Barbey F, Maisonobe T, Caillaud C, et al. Hyperhidrosis : a new and often early symptom in Fabry disease. International experience and data from the Fabry Outcome Survey. *Int J Clin Pract*. 2006;60(9) :1053-9.
232. van der Tol L, Sminia ML, Hollak CE, Biegstraaten M. Cornea verticillata supports a diagnosis of Fabry disease in non-classical phenotypes : results from the Dutch cohort and a systematic review. *Br J Ophthalmol*. 2016;100(1) :3-8.
233. Nguyen TT, Gin T, Nicholls K, Low M, Galanos J, Crawford A. Ophthalmological manifestations of Fabry disease : a survey of patients at the Royal Melbourne Fabry Disease Treatment Centre. *Clin Exp Ophthalmol*. 2005;33(2) :164-8.
234. Sodi A, Ioannidis AS, Mehta A, Davey C, Beck M, Pitz S. Ocular manifestations of Fabry's disease : data from the Fabry Outcome Survey. *Br J Ophthalmol*. 2007;91(2) :210-4.
235. Pitz S, Kalkum G, Arash L, Karabul N, Sodi A, Larroque S, et al. Ocular signs correlate well with disease severity and genotype in Fabry disease. *PLoS One*. 2015;10(3) :e0120814.
236. Ortiz A, Oliveira JP, Waldek S, Warnock DG, Cianciaruso B, Wanner C, et al. Nephropathy in males and females with Fabry disease : cross-sectional description of patients before treatment with enzyme replacement therapy. *Nephrol Dial Transplant*. 2008;23(5) :1600-7.
237. Cebulla M, Kurschat C, West M, Nicholls K, Torras J, Sunder-Plassmann G, et al. Kidney transplantation and enzyme replacement therapy in patients with Fabry disease. *J Nephrol*. 2013;26(4) :645-51.
238. Ries M, Moore DF, Robinson CJ, Tiffet CJ, Rosenbaum KN, Brady RO, et al. Quantitative dysmorphology assessment in Fabry disease. *Genet Med*. 2006;8(2) :96-101.

239. van der Tol L, Smid BE, Poorthuis BJ, Biegstraaten M, Deprez RH, Linthorst GE, et al. A systematic review on screening for Fabry disease : prevalence of individuals with genetic variants of unknown significance. *J Med Genet.* 2014;51(1) :1-9.
240. Juchniewicz P, Kloska A, Tylki-Szymanska A, Jakobkiewicz-Banecka J, Wegrzyn G, Moskot M, et al. Female Fabry disease patients and X-chromosome inactivation. *Gene.* 2018;641 :259-64.
241. Wilcox WR, Oliveira JP, Hopkin RJ, Ortiz A, Banikazemi M, Feldt-Rasmussen U, et al. Females with Fabry disease frequently have major organ involvement : lessons from the Fabry Registry. *Mol Genet Metab.* 2008;93(2) :112-28.
242. Maruyama H, Miyata K, Mikame M, Taguchi A, Guili C, Shimura M, et al. Effectiveness of plasma lyso-Gb3 as a biomarker for selecting high-risk patients with Fabry disease from multispecialty clinics for genetic analysis. *Genet Med.* 2019;21(1) :44-52.
243. *Maladie de Fabry : Protocole national de diagnostic et de soins (2010).*
244. Lidove O, Barbey F, Joly D. [Treatment of Fabry disease : Successes, failures, and expectations]. *Nephrol Ther.* 2016;12 Suppl 1 :S105-13.
245. Breunig F, Weidemann F, Strotmann J, Knoll A, Wanner C. Clinical benefit of enzyme replacement therapy in Fabry disease. *Kidney Int.* 2006;69(7) :1216-21.
246. Ries M, Clarke JT, Whybra C, Timmons M, Robinson C, Schlaggar BL, et al. Enzyme-replacement therapy with agalsidase alfa in children with Fabry disease. *Pediatrics.* 2006;118(3) :924-32.
247. Ramaswami U. Fabry disease during childhood : clinical manifestations and treatment with agalsidase alfa. *Acta Paediatr.* 2008;97(457) :38-40.
248. Schiffmann R, Pastores GM, Lien YH, Castaneda V, Chang P, Martin R, et al. Agalsidase alfa in pediatric patients with Fabry disease : a 6.5-year open-label follow-up study. *Orphanet J Rare Dis.* 2014;9 :169.
249. Germain DP, Nicholls K, Giugliani R, Bichet DG, Hughes DA, Barisoni LM, et al. Efficacy of the pharmacologic chaperone migalastat in a subset of male patients with the classic phenotype of Fabry disease and migalastat-amenable variants : data from the phase 3 randomized, multicenter, double-blind clinical trial and extension study. *Genet Med.* 2019.
250. Benjamin ER, Della Valle MC, Wu X, Katz E, Pruthi F, Bond S, et al. The validation of pharmacogenetics for the identification of Fabry patients to be treated with migalastat. *Genet Med.* 2017;19(4) :430-8.
251. Warnock DG, Thomas CP, Vujkovic B, Campbell RC, Charrow J, Laney DA, et al. Antiproteinuric therapy and Fabry nephropathy : factors associated with preserved kidney function during agalsidase-beta therapy. *J Med Genet.* 2015;52(12) :860-6.
252. Ries M, Mengel E, Kutschke G, Kim KS, Birklein F, Krummenauer F, et al. Use of gabapentin to reduce chronic neuropathic pain in Fabry disease. *J Inher Metab Dis.* 2003;26(4) :413-4.
253. Tebani A, Afonso C, Marret S, Bekri S. Omics-Based Strategies in Precision Medicine : Toward a Paradigm Shift in Inborn Errors of Metabolism Investigations. *Int J Mol Sci.* 2016;17(9).
254. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics : diagnosing rare disease in children. *Nat Rev Genet.* 2018;19(5) :253-68.
255. Metzker ML. Emerging technologies in DNA sequencing. *Genome Res.* 2005;15(12) :1767-76.
256. Tiffon C. The Impact of Nutrition and Environmental Epigenetics on Human Health and Disease. *Int J Mol Sci.* 2018;19(11).
257. Yi H, Cho YJ, Won S, Lee JE, Jin Yu H, Kim S, et al. Duplex-specific nuclease efficiently removes rRNA for prokaryotic RNA-seq. *Nucleic Acids Res.* 2011;39(20) :e140.
258. Chen Z, Duan X. Ribosomal RNA depletion for massively parallel bacterial RNA-sequencing applications. *Methods Mol Biol.* 2011;733 :93-103.
259. Hansen KD, Brenner SE, Dudoit S. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res.* 2010;38(12) :e131.
260. Ozsolak F, Milos PM. RNA sequencing : advances, challenges and opportunities. *Nat Rev Genet.* 2011;12(2) :87-98.

261. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*. 2009;460(7254) :479-86.
262. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*. 2010;141(1) :129-41.
263. Domon B, Aebersold R. Mass spectrometry and protein analysis. *Science*. 2006;312(5771) :212-7.
264. Fenn JB. Electrospray wings for molecular elephants (Nobel lecture). *Angew Chem Int Ed Engl*. 2003;42(33) :3871-94.
265. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science*. 1989;246(4926) :64-71.
266. Tanaka K. The origin of macromolecule ionization by laser irradiation (Nobel lecture). *Angew Chem Int Ed Engl*. 2003;42(33) :3860-70.
267. Geyer PE, Holdt LM, Teupser D, Mann M. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol*. 2017;13(9) :942.
268. Lequin RM. Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). *Clin Chem*. 2005;51(12) :2415-8.
269. Kurien BT, Scofield RH. Western blotting : an introduction. *Methods Mol Biol*. 2015;1312 :17-30.
270. Ma E, Vetter J, Bliss L, Lai HH, Mysorekar IU, Jain S. A multiplexed analysis approach identifies new association of inflammatory proteins in patients with overactive bladder. *Am J Physiol Renal Physiol*. 2016;311(1) :F28-34.
271. Belzeaux R, Lefebvre MN, Lazzari A, Le Carpentier T, Consoloni JL, Zendjidjian X, et al. How to : Measuring blood cytokines in biological psychiatry using commercially available multiplex immunoassays. *Psychoneuroendocrinology*. 2017;75 :72-82.
272. Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. *Trends Biotechnol*. 1998;16(9) :373-8.
273. Tebani A, Abily-Donval L, Afonso C, Marret S, Bekri S. Clinical Metabolomics : The New Metabolic Window for Inborn Errors of Metabolism Investigations in the Post-Genomic Era. *Int J Mol Sci*. 2016;17(7).
274. Jacob M, Lopata AL, Dasouki M, Abdel Rahman AM. Metabolomics toward personalized medicine. *Mass Spectrom Rev*. 2017.
275. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. *Nature*. 2012;491(7424) :384-92.
276. Suhre K, Raffler J, Kastenmuller G. Biochemical insights from population studies with genetics and metabolomics. *Arch Biochem Biophys*. 2016;589 :168-76.
277. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass Spectrom Rev*. 2007;26(1) :51-78.
278. Farriau JP, Dhondt JL. [Neonatal screening for phenylketonuria]. *J Genet Hum*. 1981;29(1) :3-11.
279. Vuckovic D. Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. *Anal Bioanal Chem*. 2012;403(6) :1523-48.
280. Lewis MR, Pearce JT, Spagou K, Green M, Dona AC, Yuen AH, et al. Development and Application of Ultra-Performance Liquid Chromatography-TOF MS for Precision Large Scale Urinary Metabolic Phenotyping. *Anal Chem*. 2016;88(18) :9004-13.
281. Milne SB, Mathews TP, Myers DS, Ivanova PT, Brown HA. Sum of the parts : mass spectrometry-based metabolomics. *Biochemistry*. 2013;52(22) :3829-40.
282. Weckwerth W, Wenzel K, Fiehn O. Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics*. 2004;4(1) :78-83.

283. Sitnikov DG, Monnin CS, Vuckovic D. Systematic Assessment of Seven Solvent and Solid-Phase Extraction Methods for Metabolomics Analysis of Human Plasma by LC-MS. *Sci Rep.* 2016;6 :38885.
284. Shearer GC, Harris WS, Pedersen TL, Newman JW. Detection of omega-3 oxylipins in human plasma and response to treatment with omega-3 acid ethyl esters. *J Lipid Res.* 2010;51(8) :2074-81.
285. Wolf C, Quinn PJ. Lipidomics : practical aspects and applications. *Prog Lipid Res.* 2008;47(1) :15-36.
286. Bales JR, Higham DP, Howe I, Nicholson JK, Sadler PJ. Use of high-resolution proton nuclear magnetic resonance spectroscopy for rapid multi-component analysis of urine. *Clin Chem.* 1984;30(3) :426-32.
287. Emwas A-HM, Salek RM, Griffin JL, Merzaban J. NMR-based metabolomics in human disease diagnosis : applications, limitations, and recommendations. *Metabolomics.* 2013;9(5) :1048-72.
288. Sheedy JR. Metabolite analysis of biological fluids and tissues by proton nuclear magnetic resonance spectroscopy. *Methods Mol Biol.* 2013;1055 :81-97.
289. Kuehnbaum NL, Britz-McKibbin P. New advances in separation science for metabolomics : resolving chemical diversity in a post-genomic era. *Chem Rev.* 2013;113(4) :2437-68.
290. Wolk DM, Clark AE. Matrix-Assisted Laser Desorption Time of Flight Mass Spectrometry. *Clin Lab Med.* 2018;38(3) :471-86.
291. Kebarle P. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J Mass Spectrom.* 2000;35(7) :804-17.
292. Tebani A.

Analyse métabolomique multidimensionnelle : application aux Erreurs Innées du Métabolisme 2017.

293. Lesur A, Domon B. Advances in high-resolution accurate mass spectrometry application to targeted proteomics. *Proteomics.* 2015;15(5-6) :880-90.
294. Eliuk S, Makarov A. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu Rev Anal Chem (Palo Alto Calif).* 2015;8 :61-80.
295. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R. The Orbitrap : a new mass spectrometer. *J Mass Spectrom.* 2005;40(4) :430-43.
296. Marshall AG, Hendrickson CL, Jackson GS. Fourier transform ion cyclotron resonance mass spectrometry : a primer. *Mass Spectrom Rev.* 1998;17(1) :1-35.
297. Schweikhard L, Alber GM, Marshall AG. High-frequency fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom.* 1993;4(2) :177-81.
298. Katajamaa M, Oresic M. Data processing for mass spectrometry-based metabolomics. *J Chromatogr A.* 2007;1158(1-2) :318-28.
299. Sugimoto M, Kawakami M, Robert M, Soga T, Tomita M. Bioinformatics Tools for Mass Spectroscopy-Based Metabolomic Data Processing and Analysis. *Curr Bioinform.* 2012;7(1) :96-108.
300. Misra BB, van der Hoft JJ. Updates in metabolomics tools and resources : 2014-2015. *Electrophoresis.* 2016;37(1) :86-110.
301. Roux A, Lison D, Junot C, Heilier JF. Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology : A review. *Clin Biochem.* 2011;44(1) :119-35.
302. Lafaye A, Junot C, Ramounet-Le Gall B, Fritsch P, Tabet JC, Ezan E. Metabolite profiling in rat urine by liquid chromatography/electrospray ion trap mass spectrometry. Application to the study of heavy metal toxicity. *Rapid Commun Mass Spectrom.* 2003;17(22) :2541-9.
303. Vinaixa M, Samino S, Saez I, Duran J, Guinovart JJ, Yanes O. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites.* 2012;2(4) :775-95.
304. Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML, et al. A tutorial review : Metabolomics and partial least squares-discriminant analysis--a marriage of convenience or a shotgun wedding. *Anal Chim Acta.* 2015;879 :10-23.

305. Westerhuis JA, van Velzen EJ, Hoefsloot HC, Smilde AK. Multivariate paired data analysis : multilevel PLS-DA versus OPLS-DA. *Metabolomics*. 2010;6(1) :119-28.
306. Tebani A, Schmitz-Afonso I, Rutledge DN, Gonzalez BJ, Bekri S, Afonso C. Optimization of a liquid chromatography ion mobility-mass spectrometry method for untargeted metabolomics using experimental design and multivariate data analysis. *Anal Chim Acta*. 2016;913 :55-62.
307. Tebani A, Schmitz-Afonso I, Abily-Donval L, Heron B, Piraud M, Ausseil J, et al. Urinary metabolic phenotyping of mucopolysaccharidosis type I combining untargeted and targeted strategies with data modeling. *Clin Chim Acta*. 2017;475 :7-14.
308. Tebani A, Abily-Donval L, Schmitz-Afonso I, Heron B, Piraud M, Ausseil J, et al. Unveiling metabolic remodeling in mucopolysaccharidosis type III through integrative metabolomics and pathway analysis. *J Transl Med*. 2018;16(1) :248.
309. Tebani A, Abily-Donval L, Schmitz-Afonso I, Piraud M, Ausseil J, Zerimech F, et al. Analysis of Mucopolysaccharidosis Type VI through Integrative Functional Metabolomics. *Int J Mol Sci*. 2019;20(2).
310. Mauhin W, Lidove O, Amelin D, Lamari F, Caillaud C, Mingozi F, et al. Deep characterization of the anti-drug antibodies developed in Fabry disease patients, a prospective analysis from the French multicenter cohort FFABRY. *Orphanet J Rare Dis*. 2018;13(1) :127.
311. Brady RO. Enzymatic abnormalities in diseases of sphingolipid metabolism. *Clin Chem*. 1967;13(7) :565-77.
312. Sillence DJ, Platt FM. Storage diseases : new insights into sphingolipid functions. *Trends Cell Biol*. 2003;13(4) :195-203.
313. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372(9) :793-5.
314. Dugger SA, Platt A, Goldstein DB. Drug development in the era of precision medicine. *Nat Rev Drug Discov*. 2018;17(3) :183-96.
315. Kaddurah-Daouk R, Weinshilboum RM, Pharmacometabolomics Research N. Pharmacometabolomics : implications for clinical pharmacology and systems pharmacology. *Clin Pharmacol Ther*. 2014;95(2) :154-67.
316. James LP. Metabolomics : integration of a new "omics" with clinical pharmacology. *Clin Pharmacol Ther*. 2013;94(5) :547-51.
317. Low SK, Zembutsu H, Nakamura Y. Breast cancer : The translation of big genomic data to cancer precision medicine. *Cancer Sci*. 2018;109(3) :497-506.
318. Mullane SA, Van Allen EM. Precision medicine for advanced prostate cancer. *Curr Opin Urol*. 2016;26(3) :231-9.
319. Relling MV, Evans WE. Pharmacogenomics in the clinic. *Nature*. 2015;526(7573) :343-50.
320. de Bono JS, Ashworth A. Translating cancer research into targeted therapeutics. *Nature*. 2010;467(7315) :543-9.
321. Garrod AE. The incidence of alkaptonuria : a study in chemical individuality. 1902 [classical article]. *Yale J Biol Med*. 2002;75(4) :221-31.
322. Lanpher B, Brunetti-Pierri N, Lee B. Inborn errors of metabolism : the flux from Mendelian to complex diseases. *Nat Rev Genet*. 2006;7(6) :449-60.
323. Salazar DA, Rodriguez-Lopez A, Herreno A, Barbosa H, Herrera J, Ardila A, et al. Systems biology study of mucopolysaccharidosis using a human metabolic reconstruction network. *Mol Genet Metab*. 2016;117(2) :129-39.
324. Seranova E, Connolly KJ, Zatyka M, Rosenstock TR, Barrett T, Tuxworth RI, et al. Dysregulation of autophagy as a common mechanism in lysosomal storage diseases. *Essays Biochem*. 2017;61(6) :733-49.
325. Ba Q, Raghavan G, Kiselyov K, Yang G. Whole-Cell Scale Dynamic Organization of Lysosomes Revealed by Spatial Statistical Analysis. *Cell Rep*. 2018;23(12) :3591-606.
326. Johnson DE, Ostrowski P, Jaumouille V, Grinstein S. The position of lysosomes within the cell determines their luminal pH. *J Cell Biol*. 2016;212(6) :677-92.

327. Jongsma ML, Berlin I, Wijdeven RH, Janssen L, Janssen GM, Garstka MA, et al. An ER-Associated Pathway Defines Endosomal Architecture for Controlled Cargo Transport. *Cell*. 2016;166(1) :152-66.
328. Neefjes J, Jongsma MML, Berlin I. Stop or Go? Endosome Positioning in the Establishment of Compartment Architecture, Dynamics, and Function. *Trends Cell Biol*. 2017;27(8) :580-94.
329. Soppina V, Rai AK, Ramaiya AJ, Barak P, Mallik R. Tug-of-war between dissimilar teams of microtubule motors regulates transport and fission of endosomes. *Proc Natl Acad Sci U S A*. 2009;106(46) :19381-6.
330. Ikonen E. Mechanisms for cellular cholesterol transport : defects and human disease. *Physiol Rev*. 2006;86(4) :1237-61.
331. Simons K, Ikonen E. Functional rafts in cell membranes. *Nature*. 1997;387(6633) :569-72.
332. Johansson M, Lehto M, Tanhuanpaa K, Cover TL, Olkkonen VM. The oxysterol-binding protein homologue ORP1L interacts with Rab7 and alters functional properties of late endocytic compartments. *Mol Biol Cell*. 2005;16(12) :5480-92.
333. Rocha N, Kuijl C, van der Kant R, Janssen L, Houben D, Janssen H, et al. Cholesterol sensor ORP1L contacts the ER protein VAP to control Rab7-RILP-p150 Glued and late endosome positioning. *J Cell Biol*. 2009;185(7) :1209-25.
334. Johansson M, Rocha N, Zwart W, Jordens I, Janssen L, Kuijl C, et al. Activation of endosomal dynein motors by stepwise assembly of Rab7-RILP-p150Glued, ORP1L, and the receptor betalll spectrin. *J Cell Biol*. 2007;176(4) :459-71.
335. Stenmark H. Rab GTPases as coordinators of vesicle traffic. *Nat Rev Mol Cell Biol*. 2009;10(8) :513-25.
336. Korolchuk VI, Saiki S, Lichtenberg M, Siddiqi FH, Roberts EA, Imarisio S, et al. Lysosomal positioning coordinates cellular nutrient responses. *Nat Cell Biol*. 2011;13(4) :453-60.
337. Jia R, Guardia CM, Pu J, Chen Y, Bonifacino JS. BORC coordinates encounter and fusion of lysosomes with autophagosomes. *Autophagy*. 2017;13(10) :1648-63.
338. Cho YS, Kwon HJ. Control of autophagy with small molecules. *Arch Pharm Res*. 2010;33(12) :1881-9.