



**HAL**  
open science

# Métabolomique par chromatographie en phase gazeuse couplée à la spectrométrie de masse et désordres génomiques détectés par CGH array : biomarqueurs ou mécanismes

Clément Pontoizeau

## ► To cite this version:

Clément Pontoizeau. Métabolomique par chromatographie en phase gazeuse couplée à la spectrométrie de masse et désordres génomiques détectés par CGH array : biomarqueurs ou mécanismes. Médecine humaine et pathologie. 2016. dumas-01631647

**HAL Id: dumas-01631647**

**<https://dumas.ccsd.cnrs.fr/dumas-01631647>**

Submitted on 9 Nov 2017

**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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

# AVERTISSEMENT

Cette thèse d'exercice est le fruit d'un travail approuvé par le jury de soutenance et réalisé dans le but d'obtenir le diplôme d'Etat de docteur en médecine. Ce document est mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt toute poursuite pénale.

*Code de la Propriété Intellectuelle. Articles L 122.4*

*Code de la Propriété Intellectuelle. Articles L 335.2-L 335.10*

UNIVERSITÉ PARIS DESCARTES  
Faculté de Médecine PARIS DESCARTES

Année 2016

N° 207

THÈSE  
POUR LE DIPLÔME D'ÉTAT  
DE  
DOCTEUR EN MÉDECINE

Métabolomique par chromatographie en phase gazeuse couplée  
à la spectrométrie de masse et désordres génomiques détectés  
par CGH array : biomarqueurs ou mécanismes ?

Présentée et soutenue publiquement  
le 15 septembre 2016

Par

**Clément PONTOIZEAU**

Né le 28 août 1984 à Limoges (87)

Dirigée par M. Le Docteur Chris Ottolenghi, MCU-PH

Jury :

M. Le Professeur Robert Barouki, PU-PH ..... Président

M. Le Professeur Serge Romana, PU-PH

M. Le Professeur Jean Louis Baudeau, PU-PH

Mme Le Docteur Julie Steffann, PH

Mme Le Docteur Agnès Rotig, Directeur de recherche Inserm



Except where otherwise noted, this work is licensed under  
<http://creativecommons.org/licenses/by-nc-nd/3.0/>

## Remerciements

Je remercie le Pr. Robert BAROUKI pour avoir accepté de présider le jury de cette thèse. Je tiens à le remercier particulièrement pour m'avoir guidé vers le métabolisme, m'avoir accueilli dans son laboratoire et me faire confiance pour la suite.

Je remercie les membres du jury, le Dr Julie STEFFANN, le Dr Agnès ROTIG, le Pr. Jean Louis BEAUDEUX et le Pr Serge ROMANA pour avoir accepté d'évaluer ce travail de thèse.

Je remercie le Dr Chris OTTOLENGHI pour avoir encadré ce travail de thèse avec exigence et précision dans un souci d'exhaustivité et pour toutes les discussions très fructueuses que nous avons eues.

Je remercie tous les acteurs qui ont rendu ce travail possible, tout particulièrement le Dr Valérie MALAN, le Pr Serge ROMANA et le Pr Michel VEKEMANS du service de cytogénétique. Je remercie également le Pr. Stanislas LYONNET, le Pr. Jeanne AMIEL et le Dr. Anne PHILIPPE du service de génétique clinique de l'hôpital Necker pour leur aide dans la validation à venir des résultats de ce travail de thèse.

Je remercie Bernadette pour sa bienveillance, sa disponibilité, ses conseils et son encadrement au quotidien. Je remercie également Florence, qui a souvent été dérangée par des discussions intempestives, pour sa disponibilité et sa patience.

Je remercie les membres du secteur acides aminés/acides organiques qui ont conduit les analyses de routine exploitées dans ce travail : Jacqueline, Marie, Odile, Sabine, Martine et Morgan.

Je remercie l'ensemble du laboratoire de biochimie métabolomique et protéomique de l'hôpital Necker pour l'accueil qui m'a été réservé.

Je remercie mes proches et ma famille pour leur patience et leur soutien.

Je remercie Mathilde pour son soutien indéfectible.

# Table des matières

<b><u>PARTIE 1 INTRODUCTION.....</u></b>	<b><u>4</u></b>
<b><u>PARTIE 2 MATERIEL ET METHODES .....</u></b>	<b><u>9</u></b>
2.1 SUJETS.....	9
2.2 METHODOLOGIES.....	10
<b><u>PARTIE 3 RESULTATS.....</u></b>	<b><u>13</u></b>
3.1 MISE EN EVIDENCE DE BIOMARQUEURS METABOLIQUES .....	13
3.1.1. AUGMENTATION DE L'URACILE, DE LA THYMINE ET DELETION DE <i>TYMP</i> .....	13
3.1.2. DIMINUTION DE L'ACIDE N-ACETYLASPARTIQUE ET DELETION DE <i>NAT8L</i> .....	17
3.1.3. AUTRES GENES.....	19
3.2 COMPARAISONS DES PROFILS METABOLIQUES ENTRE PATIENTS PRESENTANT DES DESORDRES GENOMIQUES COMPARABLES OU OPPOSES .....	20
3.3 INTERPRETATION DES PROFILS METABOLIQUE EN TERME DE RETARD DE DEVELOPPEMENT ET VIEILLISSEMENT PRECOCE .....	21
3.4 MARQUEURS METABOLIQUES POUR LA DETECTION DE DELETIONS AFFECTANT DES GENES HAPLOINSUFFISANTS .....	24
3.5 GENES METABOLIQUES ASSOCIES A DES VARIATIONS QUANTITATIVES .....	24
<b><u>PARTIE 4 DISCUSSION .....</u></b>	<b><u>25</u></b>
4.1 SIGNATURES METABOLIQUES DES DELETIONS DE <i>TYMP</i> ET <i>NAT8L</i> .....	25
4.2 LA DELETION DE <i>TYMP</i> EST-ELLE PATHOGENE ? .....	26
4.3 AGE, PATHOLOGIE ET METABOLISME .....	28
4.4 PERSPECTIVES .....	29
<b><u>PARTIE 5 CONCLUSION.....</u></b>	<b><u>32</u></b>
<b><u>PARTIE 6 BIBLIOGRAPHIE.....</u></b>	<b><u>33</u></b>
<b><u>PARTIE 7 ANNEXES.....</u></b>	<b><u>38</u></b>

# Partie 1 Introduction

---

Le métabolisme est à la croisée de différentes influences, intégrant à la fois génétique et environnement (régime alimentaire, flore microbienne, toxique...). L'évaluation des contributions de chacun de ces facteurs représente un enjeu important. Les maladies héréditaires du métabolisme constituent l'expression majeure, asymptotique, de l'influence de la génétique sur le métabolisme. Elles correspondent à des événements génétiques rares et délétères qui impliquent des gènes clés du métabolisme, conduisant à des phénotypes métaboliques extrêmes à l'origine de la maladie. L'étude de populations humaines avec les moyens modernes de la génétique commence à mettre en évidence des corrélations significatives entre variants génétiques fréquents et variations modérées du métabolisme, étoffant le spectre du contrôle génétique du métabolisme (1).

Les analyses métaboliques constituent réciproquement un outil de choix pour la compréhension des mécanismes physiopathologiques associés aux maladies génétiques ou aux déséquilibres génomiques. En effet, elles peuvent fournir une caractérisation moléculaire fonctionnelle aux événements génétiques dont les conséquences sont souvent difficiles à prédire : nouvelles mutations homozygotes, associations de mutations à l'état hétérozygote dont l'effet combinatoire est difficile à évaluer, déséquilibres génomiques, souvent retrouvés à l'état hétérozygote, dont le caractère pathogène est souvent difficile à estimer sur la seule connaissance de la séquence du segment chromosomique impliqué. En pratique clinique, les analyses métaboliques permettent la détection de marqueurs, potentiellement d'intérêt diagnostique, qui peuvent orienter vers un mécanisme physiopathologique et d'éventuelles pistes thérapeutiques.

Dans le cadre de cette étude, nous allons nous intéresser plus particulièrement à l'impact sur le métabolisme d'un certain type d'anomalies génétiques : les déséquilibres génomiques, détectés par hybridation génomique comparative sur réseau d'ADN (CGH array). Ces anomalies correspondent à des variations quantitatives du génome, appelées « copy variant number » (CNV). Ces variations du nombre de copies de l'ADN, d'une taille d'un kilobase à plusieurs mégabases correspondent à des délétions, duplications, insertions et remaniements complexes (2). La signification d'un CNV identifié dans la pathogenèse d'un phénotype est

parfois difficile à établir. Un CNV est considéré comme pathogène lorsqu'il est jugé responsable des signes cliniques observés chez le patient. Un CNV pathogène avec une pénétrance incomplète et expressivité variable est un déséquilibre génomique nécessaire mais non suffisant à la survenue de la pathologie. Les CNV considérés comme bénins ne présentent pas de conséquences phénotypiques connues à ce jour. Enfin, d'autres CNV (variants de signification indéterminée, VOUS) présentent des difficultés d'interprétation, leur lien de causalité avec le phénotype ne pouvant être établi (2). Ces anomalies peuvent être récurrentes ou sporadiques. Les CNV induisent approximativement les variations attendues au niveau des dosages géniques, avec diminution et augmentation de moitié respectivement pour les délétions et les duplications (3, 4). La comparaison d'individus présentant des anomalies récurrentes réciproques, délétions ou duplications du même locus, a pu mettre en évidence des phénotypes « opposés ». Les délétions et duplications au locus 1q21.1 entraînent par exemple microcéphalie ou macrocéphalie (5). Lacaria et al. ont montré dans des modèles animaux que des réarrangements de la région 17p11.2 associés aux syndromes de Smith-Magenis (délétion) et Potocki-Lupski (duplication) causent des altérations opposées dans le métabolisme du cholestérol et des lipoprotéines (6).

Les désordres génomiques ont évidemment un impact majeur sur les gènes du métabolisme lorsqu'il s'agit de délétions homozygotes ou hémizygotiques qui induisent des déficits enzymatiques complets à l'origine de la maladie. Par exemple, des délétions du chromosome X chez des individus masculins sont impliquées dans des déficits en glycérol kinase et des déficits en ornithine transcarbamylase (7, 8). La particularité des désordres génomiques est alors de donner lieu à des phénotypes complexes résultant de la délétion simultanée de plusieurs gènes, appelés syndromes délétionnels. Le déficit de glycérol kinase peut ainsi être associé à la dystrophie musculaire de Duchenne et à l'hypoplasie congénitale des surrénales (8). Les anomalies métaboliques peuvent donc être considérées comme des marqueurs de syndromes délétionnels et doivent systématiquement faire rechercher l'existence d'autres signes cliniques pouvant évoquer l'inactivation de « gènes contigus ».

Au-delà de ces anomalies majeures en rapport avec des déficits enzymatiques complets, il est concevable que les CNVs puissent induire des altérations plus modérées du métabolisme. Par exemple, une duplication impliquant un gène codant pour une enzyme pourrait en augmenter l'activité spécifique (activité normalisée à la quantité de tissu ou de protéine). Siggberg et al. ont mis en évidence une activité spécifique augmentée de l'enzyme succinate semialdéhyde

dehydrogenase (SSADH) impliquée dans le métabolisme de l'acide  $\gamma$ -amino-butyrique chez des patients porteurs de duplications au locus 6p22.2, affectant le gène *ALDH5A1* codant pour la SSADH (9). À l'inverse, une délétion impliquant un gène métabolique peut diminuer l'activité de l'enzyme associée dans la mesure où sa concentration est déterminée en *cis*. La question qui nous intéresse particulièrement est de savoir si une réduction modérée de l'activité d'une enzyme ou d'un transporteur peut suffisamment altérer un profil métabolique (généralement observé à l'état stationnaire) pour aider à évoquer un diagnostic. En effet, d'assez nombreuses études ont mis en évidence des altérations modérées des profils métaboliques chez des individus hétérozygotes pourtant asymptomatiques : *PAH*/phenylcétonurie, *SLC3A1* et *SLC7A9*/cystinurie (10), *ASS1*/citrullinémie de type 1 (données non publiées), *MCCCI*/déficit en 3-méthylcrotonyl-CoA carboxylase (11)...

Par ailleurs, des études « génome - métabolome - association studies, mGWAS » ont identifié un grand nombre d'associations significatives entre polymorphismes nucléotidiques (SNP) et concentrations de métabolites dans le plasma ou l'urine (1, 12, 13).

Ainsi, il est légitime de supposer que des CNV touchant des gènes métaboliques aient souvent une expression biochimique, certes modérée et sans conséquences cliniques évidentes, mais potentiellement utiles à des fins diagnostiques (marqueurs de CNV).

D'autre part, il n'est pas exclu que de telles altérations biochimiques puissent parfois, *conditionnellement à d'autres facteurs*, s'exprimer par des effets cliniques, et donc une haploinsuffisance dans le cas des délétions ou autres causes de déficits à l'état hétérozygote.

La théorie de la MCA (« Metabolic Control Analysis ») a permis d'expliquer d'une manière simple pourquoi la grande majorité des mutations affectant le métabolisme sont récessives. En effet, dans des conditions assez souvent retrouvées à l'état métabolique basal, le flux métabolique dépendant de multiples enzymes d'une voie ne sera altéré que par des variations massives (souvent >90%) dans l'activité de chaque enzyme individuelle (14). Néanmoins, les modélisations issues de la MCA donnent des résultats moins univoques pour des conditions impliquant des états métaboliques extrêmes (e.g. , saturation en substrat) ou des paramètres métaboliques complexes (cinétique, régulation, etc.). Ainsi, il a même été proposé (15) que la prépondérance de la récessivité parmi les mutations observées semble constituer, en fait, un résultat adaptatif, c.à.d., que l'évolution biologique ait conduit à un choix opportun de paramètres métaboliques permettant la situation typique décrite par Kacser & Burns. Le modèle central de la MCA serait donc à considérer comme typique, utile pour l'organisme, facile à s'expliquer, mais pas une nécessité.



Ainsi, il ne faut peut-être pas s'étonner qu'il existe maintenant plusieurs exemples de maladies héréditaires dues à l'haploinsuffisance d'un gène métabolique : déficit en GLUT1 (16), déficit en transporteurs MCT1 (17) et possiblement certains déficits en P5C synthase (18). Ho et al. ont postulé qu'une haploinsuffisance du gène *GPR172B* codant pour un transporteur de la riboflavine puisse être à l'origine de carence en cette vitamine (19) et Dervent et al. ont observé que des mutations à l'état hétérozygote du gène *ALDH5HA1* sont associées à des phénotypes modérés et à des activités de la SSADH intermédiaires entre patients homozygotes et individus normaux ce qui suggère un mécanisme d'haploinsuffisance (20). Hung et al. ont suggéré sur les bases de l'étude d'un modèle murin que les porteurs de mutations pathogènes à l'état hétérozygote du gène *NPCI*, impliqué dans la maladie de Niemann-Pick C1 puissent être vulnérables à des troubles neuropsychiatriques (troubles moteurs et anxieux) (21).

Il est tentant de suggérer que de telles anomalies sans effets dans des conditions habituelles puissent néanmoins contribuer à la pathologie dans des contextes spécifiques, par exemple en se présentant comme des facteurs de gravité pour d'autres maladies.

De plus, les effets des CNV sur le métabolisme peuvent être complexes en combinant les effets directs et indirects de plusieurs gènes. Ainsi, Napoli et al. ont observé que la délétion 22q11.2 s'associe à l'haploinsuffisance de deux gènes métaboliques compris dans la région, *PRODH*, codant la proline oxydase, et *SLC25A1*, codant pour le transporteur mitochondrial du citrate, et induit une signature métabolique plasmatique qui, en partie, ressemble à celles associées aux inactivations homozygotes de chacun de ces gènes, mais qui reflète également une redirection globale du métabolisme central. Notamment, la phosphorylation oxydative diminue au profit d'un métabolisme glycolytique en impliquant peut-être la dérégulation (en *trans*) de facteurs de transcriptions aux multiples effets métaboliques comme c-MYC et HIF1-alpha (22). A notre connaissance, il n'existe pas d'études associant CNV détectés par CGH array et marqueurs métaboliques, contrairement au nombre croissant d'études mGWAS (cf ci-dessus), qui ont mis en évidence des associations significatives entre polymorphismes nucléotidiques (SNP) et marqueurs métaboliques (1, 12, 13).

L'analyse fonctionnelle de désordres génomiques par des analyses métaboliques pourrait se révéler particulièrement pertinente en clinique dans le contexte des déficiences intellectuelles. En effet, des dysfonctionnements métaboliques sont impliqués dans de nombreuses pathologies neurologiques, en rapport aux besoins énergétiques importants du cerveau et aux rôles de neurotransmetteurs de nombreux métabolites (glutamate, aspartate, glycine,...).

Certaines maladies héréditaires du métabolisme s'expriment sous la forme d'une déficience intellectuelle isolée (déficit en adenylosuccinate lyase, déficit de la glycosylation (CDG), déficit en créatine, certaines formes de déficits du cycle de l'urée) (23). Au total, les étiologies métaboliques représentent environ 1% des cas de déficience intellectuelle isolée, mais des investigations métaboliques poussées et répétées peuvent augmenter le rendement diagnostique à 3% (24). Les analyses métaboliques font donc partie intégrante du bilan étiologique des déficiences intellectuelles (25). Par ailleurs, les anomalies chromosomiques et déséquilibres génomiques représentent presque 30% des étiologies de déficiences intellectuelles (2, 26). La CGH array s'impose comme examen de première intention dans ces pathologies (2).

Il serait donc intéressant de croiser les données métaboliques et les résultats de CGH array obtenus lors de bilans étiologiques de déficiences intellectuelles à la recherche d'anomalies métaboliques modérées induites par les désordres génomiques et potentiellement impliquées dans la pathogénèse de la maladie. D'un point de vue pratique, cela permettrait d'aider à la classification des CNV et fournirait des marqueurs métaboliques diagnostiques.

L'objectif principal de cette étude est de mettre en évidence des marqueurs métaboliques urinaires de déséquilibres génomiques de type CNV, chez des patients présentant une déficience intellectuelle ou un syndrome polymalformatif. Ce projet s'appuie sur la mise au point récente du dosage de nombreux métabolites urinaires par chromatographie en phase gazeuse couplée à la spectrométrie de masse (GC-MS ou GC-MS/MS). Nous avons également exploré certaines hypothèses physiopathologiques en rapport avec le développement, et généré des listes d'enzymes ou transporteurs associés à des métabolites pouvant servir de marqueurs de CNV, qu'ils aient ou non à un rôle pathologique à l'état hétérozygote.

## Partie 2 Matériel et Méthodes

---

### 2.1 Sujets

Entre avril 2014 et avril 2015, 1925 dosages d'acides organiques urinaires ont été réalisés pour 1741 patients par GC-MS dans le Laboratoire de Biochimie Métabolomique et Protéomique de l'hôpital Necker. Pendant cette même période, le Laboratoire de Cytogénétique a effectué 2208 CGH arrays à la recherche de désordres génomiques. 238 patients ont été étudiés simultanément par les deux approches. Parmi ces 238 patients, seul 30 présentaient un désordre génomique : variants pathogènes connus (N = 15), variants prédisposant à des troubles neurodéveloppementaux (N = 4), variants de signification inconnue (VOUS, N = 6) et variants bénins (N=5) (cf. Table Annexe 1). Les remaniements étaient de tailles variables, compris entre 0,11 Mb et 9,6 Mb (médiane = 1.31 Mb).

Les patients 9 et 12 présentaient une duplication comparable de 115 kb au locus 1q21.1. Le patient 12 possédait également une duplication au locus 17q25.1. Le patient 5 présentait une délétion de 2.7 Mb incluant le locus 1q21.1 des deux patients précédents, ainsi qu'une duplication 16p11.2. Le patient 8 présentait une duplication de 2 Mb au locus 1q21.1 différente des patients 9 et 12 mais touchant une zone commune à la zone délétée chez le patient 5. Nous avons recherché des similarités entre les patients 9 et 12 et des variations inverses chez les patients 5 et 8. Le patient 5 présentait également une duplication 16p11.2 alors que le patient 4 présentait une délétion au locus 16p11.2 incluant une partie commune. Nous avons recherché des variations inverses chez ces deux patients. Les patients 26, 28, 29 et 30 présentaient une délétion aux loci 15q11.2q13.2 ou 15q11.2q13.1 alors que le patient 27 présentait une duplication au locus 15q11.2q13.1. Des variations communes au trois premiers patients et inverses du quatrième patient ont été recherchées. Les 20 patients restant présentaient des désordres génomiques différents ne permettant pas de les regrouper en sous-groupes.

## 2.2 Méthodologies

Le dosage des acides organiques urinaires a été réalisé par chromatographie en phase gazeuse couplée à la spectrométrie de masse (GC-MS), selon le protocole de routine hospitalière. Cette méthode d'analyse automatique permet la détection semi-quantitative de 300 pics correspondant à plus de 100 acides organiques. En pratique courante, 74 pics correspondant à des métabolites d'intérêt diagnostique sont systématiquement contrôlés manuellement. La quantification des acides organiques est faite par rapport à un étalon interne (phenylbutyrate) et rapportée à la concentration urinaire de créatinine.

Nous avons établi des intervalles de variation normale en fonction de l'âge pour 39 métabolites d'intérêt diagnostique présents dans plus de 85% de des 1925 profils acquis. Les patients présentant une maladie métabolique connue, une cétonurie significative (acide 3-hydroxybutanoïque > 15 mmol/mol de créatinine) ou une supplémentation en triglycérides à chaînes moyennes (acide 3-hydroxysébacique > 5 mmol/mol de créatinine) ont été exclus, conduisant à une sélection de 1493 profils pour la normalisation. Le jeu de données a été transformé afin d'obtenir des distributions normales pour les âges et concentrations en métabolites. Les âges sont transformés par :

$$\text{age\_norm} = (\text{age} + 0.01)^{0.2}$$

et les concentrations en métabolites :

$$\text{concentration\_norm} = \log(\text{concentration} + 0.5)$$

Le jeu de données a été ordonné par âges des patients croissants et divisé en 30 groupes de même nombre d'individus. La valeur médiane de la concentration de chaque métabolite a été calculée pour chacun de ces groupes ainsi que la médiane de l'écart absolu à la médiane (MAD, ou déviation standard « robuste »). Les variations de concentrations en fonction de l'âge ont ensuite été extrapolées pour chacun des métabolites par une régression locale polynomiale pondérée (loess) afin d'établir des normes continues en fonction de l'âge. Un exemple de résultats est montré Figure 2 pour l'uracile et la thymine. Pour la comparaison des profils, les concentrations en métabolites ont été exprimées en nombres de médianes des écarts absolus à la médiane (ou nombre de MAD), et donc normalisées par rapport à l'âge. Une valeur est considérée comme déviante si elle est supérieure à + 2 MAD ou inférieure à - 2 MAD. Les résultats sont présentés Figure 1.

Afin d'interpréter de manière systématique les altérations des profils métaboliques à l'aune des gènes « métaboliques » impliqués dans les désordres génomiques, nous avons établi à partir d'une analyse de la littérature, une liste de 221 gènes, enzymes ou transporteurs, connus pour être associés aux maladies héréditaires du métabolisme traditionnelles. Les anomalies de glycosylation et les maladies des lipides complexes n'ont pas été prises en compte dans cette analyse car elles n'entraînent pas de modifications détectables par GC-MS. De plus, afin d'augmenter l'exhaustivité de l'analyse, nous avons utilisé une liste de gènes proposée par Possemato et al. compilant l'ensemble des gènes, enzymes ou transporteurs impliqués dans le métabolisme humain obtenus par croisement de réseaux métaboliques et de la base de données KEGG (27). Nous avons identifié de manière systématique les gènes « métaboliques » localisés dans les régions chromosomiques remaniées chez les patients afin d'évaluer si ces désordres génomiques pouvaient être mises en relation avec d'éventuels métabolites présentant des valeurs déviantes chez ces mêmes patients. Les résultats sont présentés dans la Table 1 et la Table Annexe 2.

Nous avons également évalué si les régions chromosomiques remaniées chez nos patients comprenaient des polymorphismes ponctuels ou SNP (Single Nucleotide Polymorphism) associés à des variations de concentrations en métabolites, décrits dans des études d'association à l'échelle du génome (Genome-Wide Association Study, GWAS) (12, 13).

Si un métabolite présente une valeur déviante pour un désordre génomique, la présence dans la région de ce désordre, d'un gène métabolique en lien avec ce métabolite et de SNPs associés à la concentration de ce métabolite est un argument fort pour suspecter le désordre génomique comme responsable de cette variation.

Nous nous sommes également demandé si les variations des concentrations en acides organiques urinaires pouvaient nous renseigner sur d'éventuelles anomalies de développement en évaluant si les profils présentés par les patients ressemblaient plutôt à ceux d'individus plus jeunes ou plus âgés. Pour chaque métabolite nous avons évalué un âge « métabolique » théorique correspondant à l'âge pour lequel la valeur médiane de la distribution des concentrations du métabolite en fonction de l'âge est la plus proche de la concentration du métabolite observée chez le patient (Figure Annexe 1). Une différence positive entre l'âge « métabolique » théorique et l'âge réel correspondraient à une avance de maturation et une valeur négative à un retard. Cette analyse a été effectuée chez les patients présentant des désordres génomiques (Figure 4) ainsi que chez 900 individus contrôles. Chacun est

caractérisé par une distribution de différences d'âges (39 valeurs, correspondant à chacun des métabolites) dont nous avons calculé la valeurs moyenne et la déviation standard. La Figure 5 représente les histogrammes des distributions des valeurs moyennes et des déviations standards pour les patients porteurs de désordres génomiques et individus contrôles. Les analyses statistiques ont été effectuées dans l'environnement R (<https://cran.r-project.org>). Les intervalles de confiance des valeurs de kurtosis ont été calculés par bootstrap.

## Partie 3 Résultats

---

### 3.1 Mise en évidence de biomarqueurs métaboliques

Nous avons obtenus les profils métaboliques urinaires de 30 patients porteurs de désordres génomiques par GC-MS. La Figure 1 présente ces résultats sous forme de heatmap. Le code couleur correspond à la concentration des métabolites exprimée en nombre de MAD, normalisée par rapport à l'âge.

La Table 1 liste les patients (6/30) présentant des désordres génomiques qui affectent des gènes métaboliques connus pour être associés à des maladies héréditaires du métabolisme, obtenue par croisement avec une liste de 221 gènes (cf. Méthodes). Les altérations des profils métaboliques sont également rapportées pour chacun des patients. De façon similaire, la Table 2 présente les résultats du croisement avec une liste de plus de 2000 gènes métaboliques du génome humain (cf Méthodes), dont 86 sont impliqués chez 21 patients.

Deux patients (n°1 et n°2 de la Table 1) nous sont apparus particulièrement intéressants.

#### 3.1.1. Augmentation de l'uracile, de la thymine et délétion de *TYMP*

La patiente 1, âgée de 2 ans, présentait un retard de développement psychomoteur et une ataxie. L'analyse par CGH array a mis en évidence une délétion de 8.2 Mb au locus 22q13.2qter affectant deux gènes associés à des métaboliques *ARSA* et *TYMP*. Le gène *ARSA* code pour l'enzyme lysosomale arylsulfatase A dont le déficit est à l'origine d'une maladie lysosomale, la leucodystrophie metachromatique. Cette maladie entraîne une augmentation des sulfatides urinaires, non détectables par notre approche de GC-MS. Le second gène, *TYMP*, code pour la thymidine phosphorylase dont les mutations homozygotes sont à l'origine d'une encéphalopathie myo-neuro-gastro intestinale mitochondriale (MNGIE de type 1) (28). Cette maladie peut être suspectée devant une augmentation dans l'urine de métabolites impliqués dans la biosynthèse des pyrimidines : uracile, thymine, thymidine et deoxyuridine (29).

Table 1. Patients présentant des désordres génomiques impliquant des gènes responsables de maladies métaboliques.

Patient	Age	Remaniement	Caractère	Chr	Borne.Inf	Borne.Sup	Symbole	OMIM	Nom complet	Clinique	Métabolites chez patients	
1	2.3	del 22q13.2qter	Pathogène fréquent limites variables	22	del	43006204	51193680	ARSA	250100	arylsulfatase A	leucodystrophie métachromatique	uracil++, thymine++, 3-methylglutaric/glutaconic++, pyroglutamic++, erythronic+, tiglylglycine+, orotic+, methylcitric+, vanillic-, vanilpyruvic+
								TYMP	131222	thymidine phosphorylase	MNGIE type 1	
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	Pathogène récurrent	4	del	1	9628003	IDUA	607014	a L iduronidase	mucopolysaccharidose de type I (maladie de Hurler-Scheie)	N-acétylaspartate-, 3OH-propionic+, ethylmalonic -, succinic-, 3methylgutacon+, 2-hydroxyglutaric-, hydroxyméthylglutaric-
								NAT8L	614063	N-acetyltransferase 8-like	Déficit en N-acétylaspartate	
3	0.7	del Xp21.3p21.1	Pathogène	X	del	28919849	35296972	GK	300474	Glycérol kinase	Néoglucogénèse	glycérol+++
4	17.0	del 16p11.2	Prédisposition	16	del	29673954	30198600	ALDOA	103850	aldolase A, fructose-bisphosphate	glycogénose de type XII	Aucun
5	2.2	del 1q21.1/ dup 16p11.2	Prédisposition	16	dup	29674848	30198600	ALDOA	103850	aldolase A, fructose-bisphosphate	glycogénose de type XII	3-methylglutaric+
6	5.7	del 16q23.2q24.1	Pathogène	16	del	81228359	85111857	MLYCD	606761	malonyl-CoA decarboxylase	acidurie organique (malonique et méthylmalonique)	pyroglutamic-

del : délétion ; dup : duplication ; tr. neuro. dev. : troubles neuro-développementaux ; VOUS : variant de signification inconnue ; Borne Inf/Borne sup : localisation chromosomique (GRCh37)



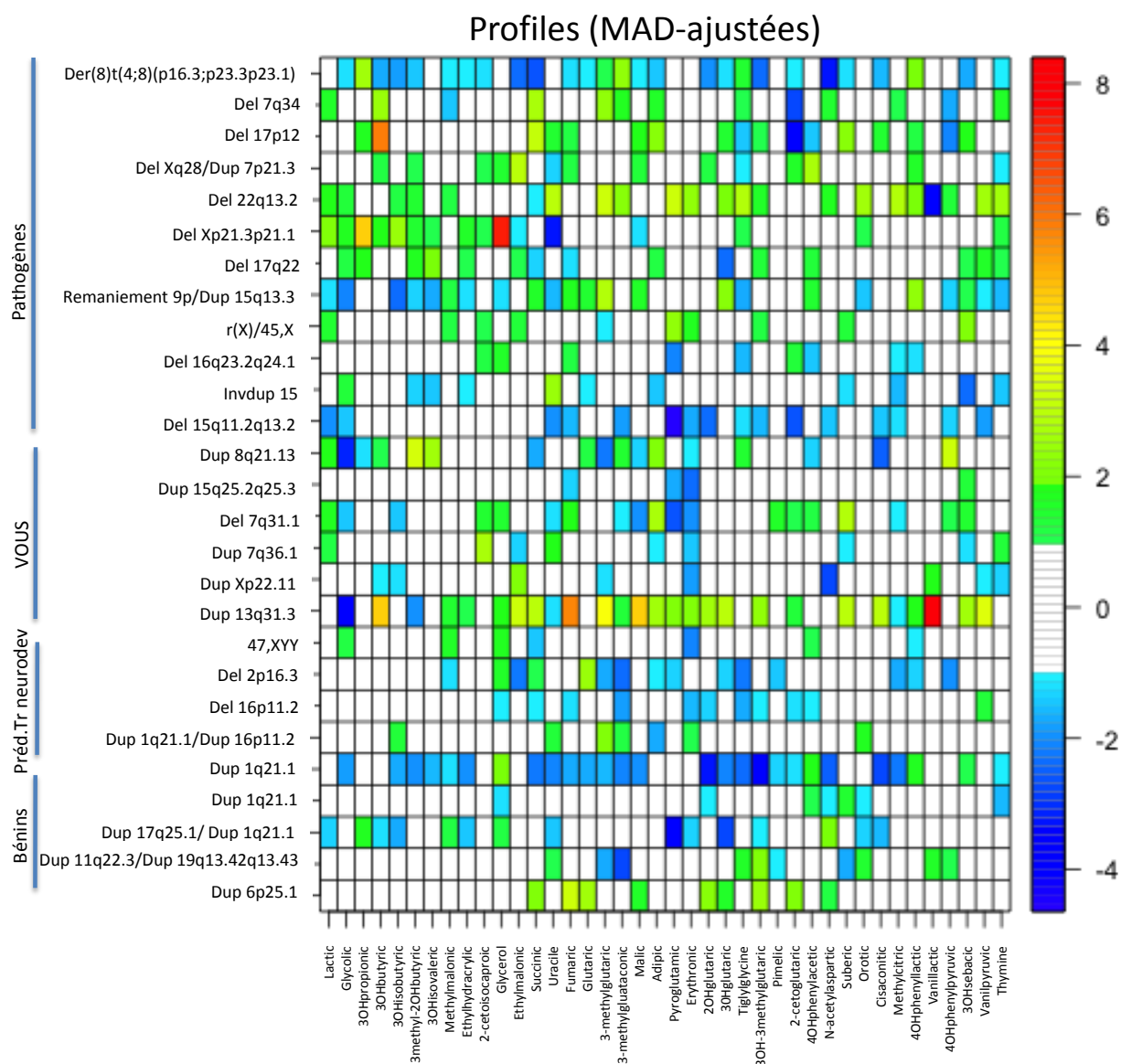


Figure 1. Heatmap des profils métaboliques de la cohorte de patients présentant des désordres génomiques. Chaque ligne correspondant à un patient, chaque colonne à un métabolite. Le code couleur correspond aux valeurs des métabolites exprimées en déviations standards (MAD ajustées par l'âge).

Le profil métabolique urinaire de cette enfant présentait des anomalies particulièrement intéressantes dans ce contexte. La Figure 2 montre une augmentation modérée mais nette de l'uracile (3 SD), de la thymine (2.6 SD) chez la patiente 1 (cercle noir) par rapport aux autres patients de la cohorte (triangles bleus) ; un patient ayant un déficit homozygote de *TYMP* (cf ci-dessous) est également indiqué (cercle vert). Certains patients de la cohorte sans désordres génomiques atteignent des valeurs similaires, suggérant une spécificité incomplète de ces

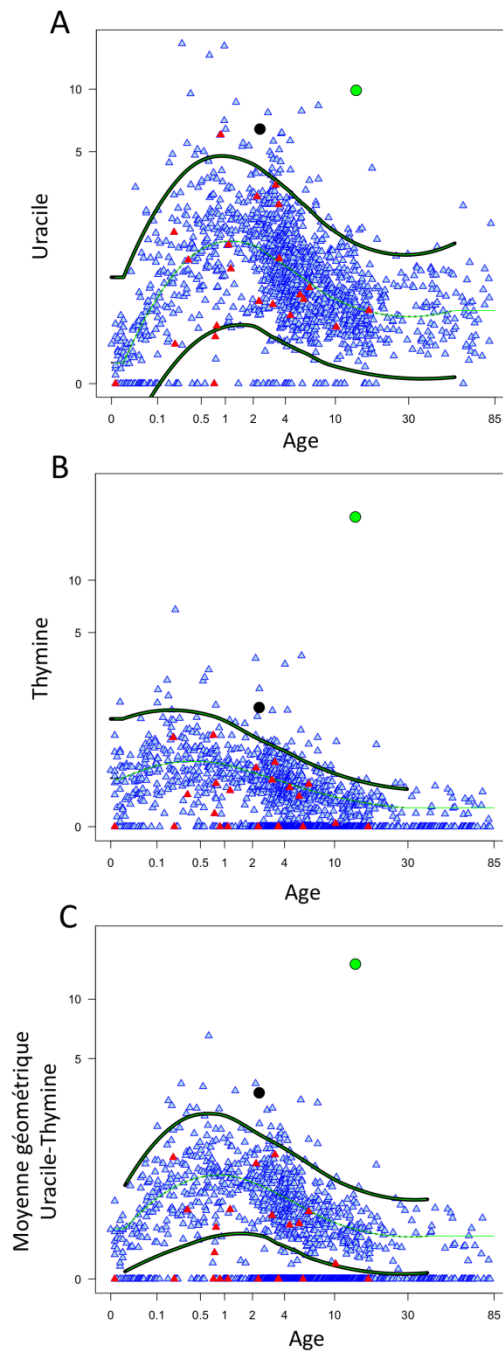


Figure 2. Concentrations urinaires A) d'uracile, B) de thymine et C) moyenne géométrique des concentrations d'uracile et de thymine, rapportées à la créatinine (unités arbitraires). Chaque triangle bleu correspond à un individu contrôlé, chaque triangle rouge à un patient présentant un désordre génomique non-bénin, le point noir à la patiente 1 présentant une délétion 22q13.2qter et le point vert à un patient atteint de MNGIE. La courbe verte centrale correspond à la valeur médiane des métabolites et les courbes extérieures à +2MAD et -2MAD.

marqueurs pour la détection de la délétion 22q13.2qter. Néanmoins, la combinaison des deux marqueurs paraît améliorer la performance des deux marqueurs (Figure 2C).

D'autre part, nous ne pouvons pas exclure que certains de ces patients soient porteurs de polymorphismes affectant l'activité de la thymidine phosphorylase.

En accord avec l'implication de la délétion de *TYMP* dans l'élévation de l'uracile et de la thymine chez notre patiente, Shin et al. ont rapporté dans une étude GWAS en population générale, l'existence d'une association très significative entre un SNP (rs131794) situé à proximité du gène *TYMP* et la concentration sanguine de l'uridine (13).

Au total, ces résultats suggèrent fortement que l'élévation de l'uracile et de la thymine dans l'urine puisse servir de signature de délétion 22qter affectant *TYMP*, même à l'état hétérozygote. De manière intéressante, le gène *SHANK3* est situé dans cette même région à seulement 145 kb de *TYMP*. Des anomalies de *SHANK3* ont été impliquées de manière significative chez des patients présentant des troubles envahissant du développement (30, 31). L'haploinsuffisance de ce gène est considérée comme responsable du phénotype de déficience intellectuelle présentée par la patiente. L'élévation de l'uracile et de la thymine dans l'urine pourrait donc être un marqueur indirect d'haploinsuffisance de *SHANK3* par délétion 22qter affectant ce gène et *TYMP*.

### 3.1.2. Diminution de l'acide N-acétylaspartique et délétion de *NAT8L*

Le patient 2, nouveau-né, présentait un syndrome de polymalformatif associant retard de croissance intra-utérin, microcéphalie, fente palatine et cardiopathie. L'analyse par CGH array a mis en évidence un remaniement chromosomique der(8)t(4;8)(p16.3;p23.3p23.1) généré par la translocation récurrente t(4;8)(p16.3p23.1) qui induit une délétion de 9.6 Mb en 4p16.3 et une duplication de 7.5 Mb en 8p23. La délétion 4p16.3 affecte deux gènes connus pour être associés à des maladies héréditaires du métabolisme, *IDUA* et *NAT8L* (Table 1). Le gène *IDUA* code l'enzyme lysosomale  $\alpha$ -L-iduronidase dont le déficit est à l'origine de la mucopolysaccharidose de type I (maladie de Hurler-Scheie) (32). Au niveau biochimique, cette maladie entraîne une augmentation de certains glucosaminoglycanes comme le dermatan sulfate et l'héparan sulfate dans l'urine. Ces métabolites, non détectables

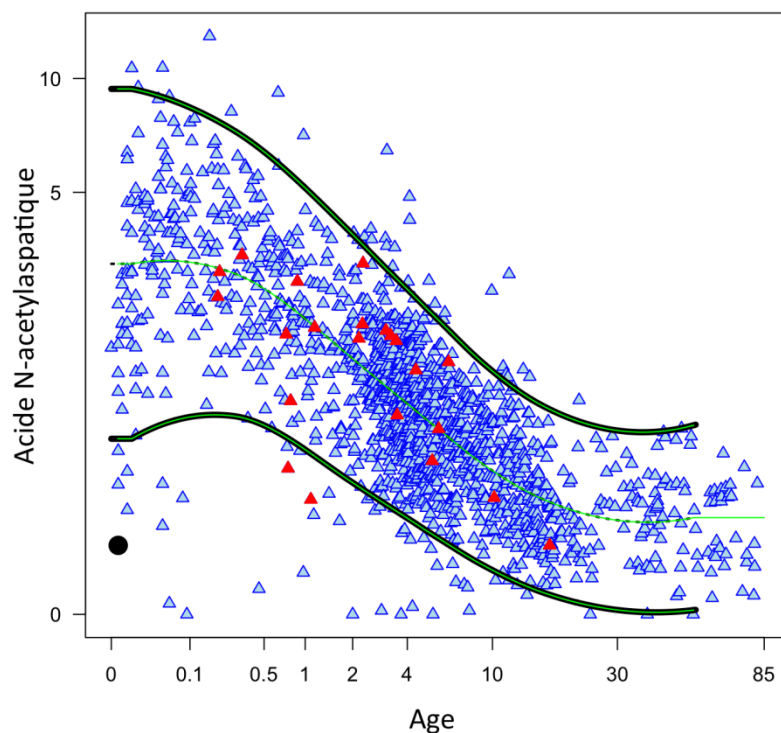


Figure 3. Concentrations de l'acide N-acetylaspartique urinaire, exprimées en  $\mu\text{mol}/\text{mmol}$  de créatinine. Chaque triangle bleu correspond à un individu contrôle, chaque triangle rouge à un patient présentant un désordre génomique non-bénin et le point noir au patient 2 présentant un remaniement génomique complexe ( $\text{der}(8)\text{t}(4;8)$  ( $\text{p}16.3;\text{p}23.3\text{p}23.1$ )). La courbe verte centrale correspond à la valeur médiane du métabolite et les courbes extérieures à  $+2\text{MAD}$  et  $-2\text{MAD}$ .

par GC-MS, ont été recherchés par les techniques adaptées en routine, sans mettre en évidence d'anomalies évocatrices d'une activité altérée de l' $\alpha$ -L-iduronidase chez ce patient.

Le second gène, *NAT8L*, codant pour une N-acetyltransferase a été considéré comme potentiellement causal dans le déficit en N-acetylaspartate du fait de son expression intense dans le cerveau, de son affinité pour l'aspartate comme substrat d'acétylation et de l'identification d'une délétion homozygote de 19 bp dans ce gène chez un enfant présentant un effondrement du N-acetylaspartate cérébrale à la spectro-IRM (33). Ce patient présentait un retard de développement psychomoteur, une ataxie, une épilepsie et une microcéphalie secondaire, à l'opposé de la megalencéphalie observée dans la maladie de Canavan causée par l'accumulation de N-acetylaspartate par déficit en aspartoacylase. De façon intéressante, notre patient présente une diminution importante du N-acetylaspartate urinaire (Figure 3) accompagnée d'une diminution plus modérée d'autres acides organiques (Table 1, Figure 1).

Ce résultat suggère que la baisse du N-acétylaspartate urinaire puisse être utilisée comme marqueur de délétion 4pter. Le gène *NAT8L* est situé à 110 kb du gène *NSD2*, codant pour une histone-lysine N-méthyltransférase, dont l'haploinsuffisance est jugée responsable du phénotype clinique présenté par le patient. Une baisse du N-acétylaspartate urinaire pourrait donc également orienter vers cette étiologie. La détection d'une baisse du N-acétylaspartate est difficile du fait de valeurs basses à la naissance. L'utilisation d'un étalon interne marqué aux isotopes stables et une méthode dédiée GC-tandem MS devrait augmenter la robustesse, la sensibilité et la spécificité de sa mesure. Le N-acétylaspartate est abondant dans le cerveau et détectable par spectro-IRM. Cette méthode pourrait également être utilisée comme outil de dépistage de baisse de N-acétylaspartate.

### 3.1.3. Autres gènes

Le patient 3 (Table 1) présentait un syndrome délétionnel Xp21, incluant le gène de la glycérol kinase à l'origine d'une augmentation importante du glycérol urinaire (+ 7DS). Le syndrome délétionnel associe hypoglycémie (liée au déficit en glycérol kinase impliquée dans la néoglucogenèse), myopathie de Duchenne (gène de la dystrophine) et hypoplasie congénitale des surrénales (gène *DR0A1/DAX1*) (8).

Le patient 4 présentait une délétion 16p11.2 alors que la patiente 5 présentait une duplication de ce même locus impliquant le gène *ALDOA* codant par l'aldolase A fructose bisphosphate dont le déficit entraîne une glycogénose de type XII. Ces deux patients ne présentaient pas d'anomalies métaboliques en lien avec ce gène.

Enfin, le patient 6 présentait une délétion 16q23.2q24.1 impliquant le gène *MLYCD* codant pour la malonyl-CoA décarboxylase dont le déficit est à l'origine de l'acidurie malonique caractérisée par une augmentation importante dans l'urine des acides malonique et méthylmalonique. Toutefois le patient ne présentait aucune de ces anomalies. Il est intéressant de noter que, chez certains patients hétérozygotes pour des mutations inactivatrices dans *MLYCD*, l'augmentation de la concentration d'acide malonique a parfois été observée transitoirement (34). Alors qu'il est peu probable que l'état déficitaire hétérozygote pour *MLYCD* soit suffisant pour exprimer des anomalies cliniquement significatives, un rôle éventuel dans le contexte d'autres anomalies coexistantes n'est pas exclu. Ainsi les résultats rapportés dans ces études suggèrent que l'acide malonique pourrait servir de marqueur des délétions hétérozygotes emportant le gène ; il est donc possible que notre technique de dosage

ne soit pas suffisamment précise. De ce fait, nous avons prévu une amélioration de la technique de dosage de ce marqueur (notamment par l'utilisation d'un étalon spécifique marqué aux isotopes stables) pour faciliter la détection des déficits hétérozygotes.

Chez 21 patients, les désordres génomiques affectaient des gènes qui ont des fonctions métaboliques démontrées *in vitro* ou présumées mais sans qu'ils aient encore été associés à des marqueurs d'intérêt clinique et/ou à des maladies métaboliques (Table Annexe 2).

### 3.2 Comparaisons des profils métaboliques entre patients présentant des désordres génomiques comparables ou opposés

Nous nous sommes ensuite intéressés aux désordres génomiques regroupant plusieurs individus. Nous avons cherché des altérations métaboliques similaires pour les patients présentant le même type d'anomalie à un locus donné et des altérations opposées pour des anomalies opposées (délétion vs duplication).

Les patients 9 et 12 présentaient une duplication 1q21.1 de 115 kb. L'analyse des acides organiques urinaires ne montrait pas de variations significatives consistantes entre ces deux patients. Les patients 5 et 8 portaient respectivement une délétion et une duplication du même locus 1q21.1. Les deux patients ne présentaient pas de variations significatives inverses. Le patient 5 présentait également une duplication 16p11.2 alors que le patient 4 présentait une délétion au locus 16p11.2. Nous n'avons pas non plus mis en évidence d'altérations opposées pour ces deux patients. Les patients 26, 28, 29 et 30 présentaient une délétion aux loci 15q11.2q13.2 ou 15q11.2q13.1 alors que le patient 27 présentait une duplication au locus 15q11.2q13.1. Le patient 29 n'a pas été pris en compte car celui-ci présentait un profil de cétonurie altérant un grand nombre de métabolites, secondaire à une durée de jeûne importante. Nous n'avons pas observé de variations inverses entre délétion et duplication.

Cette comparaison directe des variations des profils urinaires entre patients présentant des anomalies similaires ou opposées n'a pas mis en évidence d'altérations intéressantes. Ceci suggère que les valeurs déviantes présentées par ces patients ne sont pas en lien avec le déséquilibre génomique. Cette étude est cependant très limitée par le faible nombre de patients et devrait être reproduite sur des groupes plus importants.

### 3.3 Interprétation des profils métabolique en terme de retard de développement et vieillissement précoce

Nous avons interprété les profils métaboliques en terme de retard de développement ou de vieillissement précoce. Pour cela, nous avons calculé pour chaque métabolite l'âge théorique pour lequel la valeur médiane calculée dans l'ensemble de la population est la plus proche de la valeur observée (Figure 4, cf méthodes). Un âge théorique plus élevé s'interprétera comme une anticipation d'un niveau typique d'individus d'âges plus élevés. D'autre part, un âge théorique plus bas correspondra à un « retard », c'est à dire à la persistance d'un profil plus « jeune ». L'âge théorique d'un patient peut être calculé en faisant la *moyenne* des âges théoriques des métabolites composant son profil.

La Figure 5 représente les histogrammes des distributions des âges théoriques des patients porteurs de désordres génomiques et de la population contrôle. Les profils des patients semblent tendanciellement plus « âgés » (t-test, p-value = 0.003 ; Wilcoxon-test : p-value = 0.38) suggérant qu'une composante de vieillissement précoce puisse participer à la pathologie.

La variance des différences entre âge théorique et âge observé calculée sur tous les métabolites de chaque patient permet d'évaluer l'hétérogénéité des variations des métabolites dans le sens de retard et d'anticipation, et dans quelle mesure des phénomènes de retard et d'anticipation peuvent coexister chez un patient. La Figure 5B montre la distribution des variances des patients et de la population contrôle. Pour les patients, la distribution est aussi large que pour les contrôles mais semble multimodale. Le test de multimodalité (dip.test) n'est significatif ni pour les patients (p = 0.45), ni pour les contrôles (p = 0.66). Cependant, les patients présentent tendanciellement une distribution plus platykurtique (anscombe-test, kurtosis = 1.67 [1.04-2.30]) que la population contrôle (anscombe-test, kurtosis = 2.29 [2.14-2.42]). Certains patients présentent une variance élevée suggérant une coexistence d'une anticipation de certaines voies de développement tardives et d'un retard d'autres voies plus précoces. À l'inverse, certains patients présentent une variance plus faible que la normale, suggérant l'existence d'une synchronisation des voies de développement qu'elles soient en phase avec leur âge, en avance ou en retard. Ces analyses sont exploratoires et les résultats, encore préliminaires, nécessitent sans doute une confirmation sur un nombre plus important de patients.

## Anticipation/Retard (années)

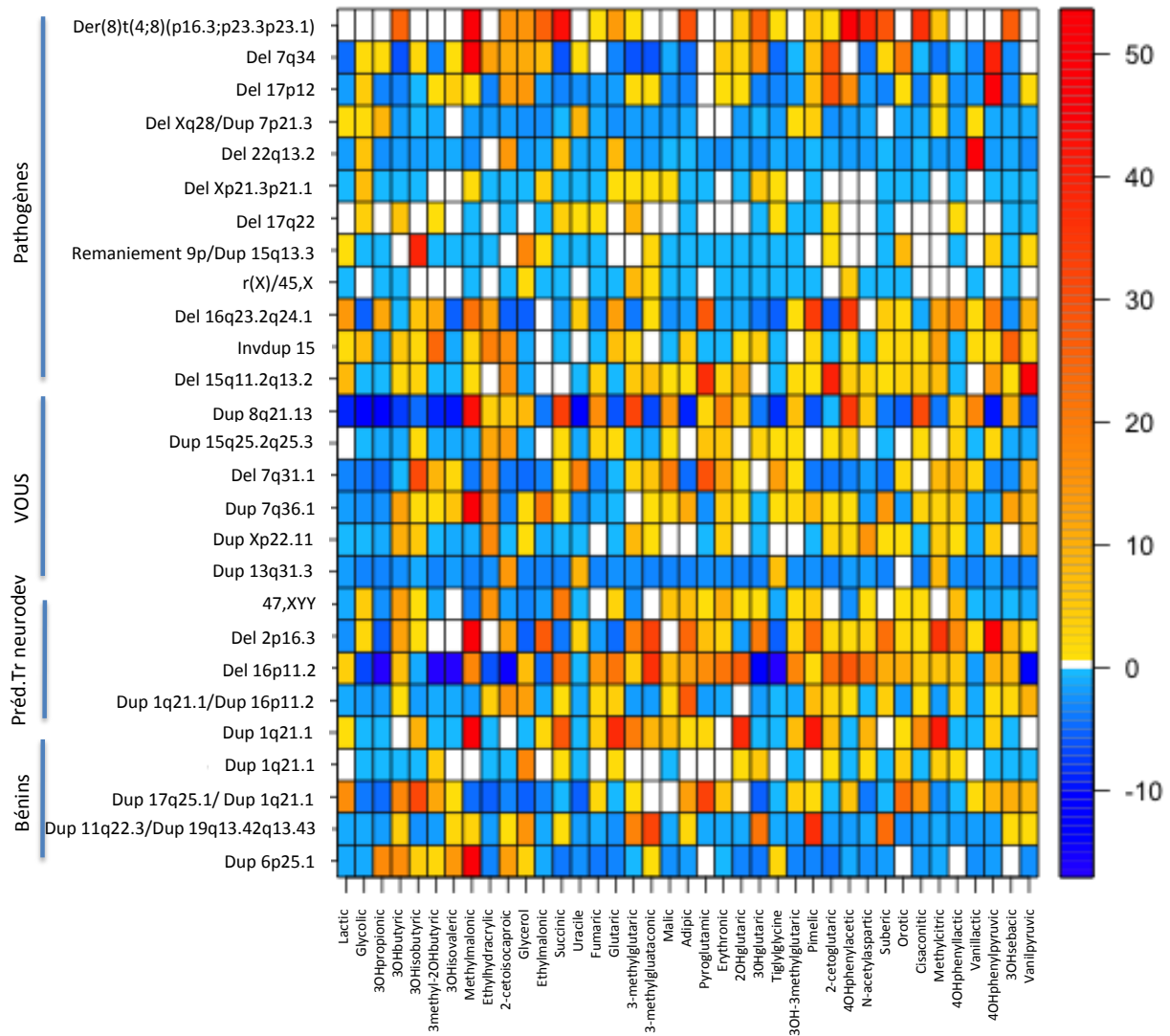


Figure 4. Heatmap des profils métaboliques interprétés en termes d'avance ou de retard de la cohorte de patients présentant des désordres génomiques. Chaque ligne correspondant à un patient, chaque colonne à un métabolite. Le code couleur correspond aux valeurs des différences d'âge entre âge réel et âge estimé à partir de la valeur du métabolite. Une valeur positive correspond à une avance, une valeur négative à un retard.



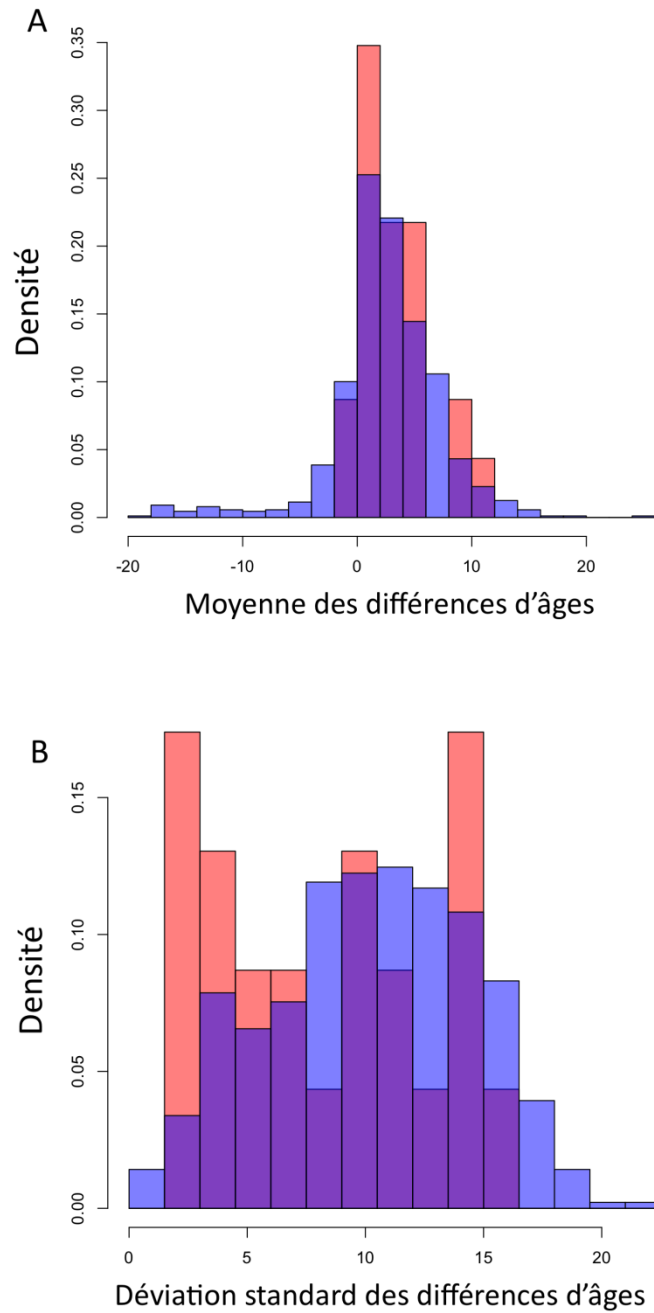


Figure 5. Histogrammes correspondant aux distributions A) des moyennes et B) déviations standards des différences entre âges réels et âges métaboliques. Distribution rouge : ensemble des patients présentant un désordre génomique non-bénin. Distribution bleue : ensemble des individus contrôles. Le violet correspond à la superposition des deux distributions.

### 3.4 Marqueurs métaboliques pour la détection de délétions affectant des gènes haploinsuffisants

Nous avons évalué la portée potentielle de l'utilisation de marqueurs métaboliques pour la détection de délétions affectant des gènes haploinsuffisants à l'origine de manifestations cliniques de manière comparable aux deux situations précédentes concernant *SHANK3* et *NSD2*. Pour cela, nous avons fusionné la liste des gènes haploinsuffisants présents dans la base OMIM (35) et la liste des gènes haploinsuffisants de la base de données ClinGen (clinicalgenome.org) conduisant à un total de 250 gènes haploinsuffisants. Nous avons ensuite calculé pour chacun de ces gènes, la distance avec le gène métabolique le plus proche, présent sur la liste de 221 gènes impliqués dans les maladies héréditaires du métabolisme central utilisée précédemment. Cette analyse montre que 6 gènes haploinsuffisants sont situés à moins de 100kb (cf Table Annexe 3) d'un gène métabolique et 27 à moins de 500kb. Il serait intéressant d'étudier spécifiquement le profil de métabolique de patients présentant des délétions affectant ces gènes afin d'augmenter le rendement de découverte de marqueurs métaboliques à intérêt diagnostique potentiel.

### 3.5 Gènes métaboliques associés à des variations quantitatives

Nous nous sommes demandés quels gènes codant des enzymes ou transporteurs du métabolisme central présentent des variants génétiques connus pour s'exprimer par des variations modérées au niveau biochimique. Nous avons donc recherché dans les études GWAS les gènes impliqués dans les maladies héréditaires du métabolisme qui ont été associés à des variations quantitatives des concentrations des métabolites liés à leur fonction biochimique (13, 36). Nous avons ainsi identifié 15 gènes répondant à ces critères (cf Annexe Table 4). Ces gènes sont donc susceptibles d'induire des variations modérées du métabolisme lorsqu'ils sont impliqués dans des déséquilibres génomiques. Il serait intéressant d'étudier spécifiquement ces situations pour identifier de nouveaux marqueurs. Par ailleurs, puisque ces gènes sont associés à des maladies héréditaires du métabolisme central, il est possible que même des altérations modérées puissent avoir un effet sur un grand nombre d'autres fonctions (cf Discussion).

## Partie 4 Discussion

---

### 4.1 Signatures métaboliques des délétions de *TYMP* et *NAT8L*

Cette étude a mis en évidence des altérations des profils métaboliques urinaires de patients présentant une délétion 22q13.2qter ou un remaniement du chromosome 4, entraînant respectivement la perte d'une copie des gènes métaboliques *TYMP* ou *NAT8L*. Les perturbations observées, augmentation de l'uracile et de la thymine dans le premier cas et diminution de l'acide N-acétylaspartique dans le second, peuvent être mises en relation avec les fonctions connues des gènes *TYMP* et *NAT8L*. Ces résultats suggèrent que ces métabolites urinaires puissent être utilisés comme marqueurs des délétions impliquant *TYMP* et *NAT8L*. Ces deux gènes étant situés à proximité des gènes haploinsuffisants *SHANK3* et *NSD2* respectivement, et donc fréquemment impliqués dans les mêmes désordres génomiques, nous pouvons proposer l'augmentation de l'uracile et de la thymine urinaire comme marqueurs indirects de délétion de *SHANK3* et la diminution de l'acide N-acétylaspartique comme marqueur indirect de délétion de *NSD2*.

Le remaniement du chromosome 4 est un événement rare. L'intérêt pratique d'un marqueur métabolique comme la diminution de l'acide N-acétylaspartique reste de fait assez limité. En revanche, les délétions du chromosome 22 en qter sont des événements plus fréquents (37, 38) et d'importance clinique dans le cadre des troubles du spectre autistique. Il a été suggéré que l'haploinsuffisance de *SHANK3* soit une des causes monogéniques les plus fréquentes de troubles envahissants du développement et de déficiences intellectuelles expliquant au moins 0,5% des cas et reste probablement sous diagnostiquée (37). L'existence d'un marqueur métabolique présenterait donc un intérêt particulier, permettant de suggérer le diagnostic chez les patients n'ayant pas bénéficié d'analyse par CGH array, notamment hors CHU. Bien évidemment les délétions impliquant *SHANK3* seuls (microdélétions, ...) sans *TYMP* ou les mutations de *SHANK3* ne seront pas détectées par cette approche.

Marti et al. ont décrit une corrélation entre l'activité résiduelle de la thymidine phosphorylase et les concentrations plasmatiques de thymidine et de deoxyuridine chez des patients atteints

de MNGIE de type 1 par mutations homozygotes du gène *TYMP* (39–41). Les patients hétérozygotes pour les mutations *TYMP*, porteurs sains, ne présentaient pas d'élévation de la thymidine et de la deoxyuridine par rapport au groupe témoin. Cependant, il est possible que de faibles différences entre ces groupes ne soient pas détectées par manque de sensibilité de la technique utilisée (HPLC), ces deux métabolites étant retrouvés en faibles concentrations à l'état normal. Il serait intéressant de rechercher une augmentation de ces deux métabolites dans le plasma par GC-MS ou LC-MS, plus sensibles, chez des patients présentant une délétion 22qter affectant *TYMP* afin d'identifier d'éventuels biomarqueurs complémentaires.

La signature métabolique urinaire présentée par le patient porteur de la délétion 22q13.2qter montre, outre l'augmentation d'uracile et de thymine, de nombreux métabolites altérés, suggérant un état de décompensation métabolique loin de l'état de physiologique de base. Les anomalies concernant la thymine et l'uracile seraient-elles toujours présentes dans un état métabolique compensé ? Il serait intéressant d'obtenir plusieurs prélèvements de ce même patient afin d'évaluer la robustesse de cette signature. Cette remarque rejoint une problématique courante des maladies métaboliques où les altérations métaboliques signatures de la maladie ne sont parfois visibles que dans des conditions particulières de stress métabolique (état de décompensation, traitements et alimentation).

Cette étude pilote présente plusieurs limites. Il est tout d'abord impératif de valider la reproductibilité de ces résultats chez d'autres patients présentant des anomalies cytogénétiques comparables. Il serait intéressant de comparer le profil métabolique de patients présentant des délétions 22qter affectant *SHANK3* avec et sans *TYMP*. En admettant la robustesse de ces marqueurs, nous avons vu que leur spécificité n'est pas de 100%, comme le montre la Figure 2. Des individus présentent des taux d'uracile et de thymine à des niveaux comparables et peuvent être considérés comme des faux positifs. La caractérisation de ces individus permettrait de préciser si ces anomalies ont une origine évidente (diététique, autres gènes de prédisposition,...). De plus, cette étude ne permet pas d'évaluer la sensibilité de ces marqueurs comme nous ne disposons que d'un seul patient pour chaque condition.

## 4.2 La délétion de *TYMP* est-elle pathogène ?

Cette étude pilote pose la question de l'implication du gène *TYMP* dans la pathogénèse de la délétion 22q13.2qter. Plusieurs études ont montré que l'haploinsuffisance de *SHANK3* n'est

probablement pas seule en cause dans le phénotype présenté par les patients (42). Frye et al. ont par exemple suggéré que la délétion de plusieurs gènes codant pour des protéines mitochondriales situés dans la même région chromosomique puisse être à l'origine d'une partie des symptômes du tableau clinique donc via un mécanisme métabolique (43).

La thymidine phosphorylase, codée par *TYMP*, catalyse la première étape du catabolisme de la thymidine, transformant ce nucléoside en thymine et 2'-deoxyribose 1-phosphate (44). Elle joue donc un rôle important dans l'équilibre du pool de nucléosides et indirectement dans la réparation des dommages de l'ADN (45).

Les anomalies génétiques (mutation homozygote, double hétérozygote...) de *TYMP* entraînant une diminution sévère de l'activité de la thymidine phosphorylase (TP) sont à l'origine d'un tableau clinique d'encéphalomyopathie neurogastrointestinale (MNGIE de type 1) associant ptosis, ophtalmoplégie externe progressive, troubles de la motilité gastrointestinale, cachexie, neuropathie périphérique, myopathie et leucoencéphalopathie et apparaissant entre la seconde et cinquième décennie (28). Des formes moins sévères et d'apparition plus tardive, associées à des activités résiduelles de TP plus importantes, ont été décrites (41, 46). Les patients hétérozygotes ne présentent a priori pas de signes cliniques hormis peut-être dans une publication où il est rapporté la présence d'un ptosis chez trois sujets apparentés probablement hétérozygotes (47). La comparaison de ces tableaux cliniques avec ceux des patients présentant une délétion 22q ne plaide pas pour mécanisme primaire impliquant *TYMP*. En effet, le tableau clinique de MNGIE de type 1, forme la plus sévère, apparaît plus tardivement que celui présenté lors de délétions 22q avec des signes cliniques différents. De plus on s'attend à avoir une activité résiduelle de TP importante en cas de délétion (de l'ordre de 30-50%), qui ne semble pas induire de signes cliniques majeurs. Cependant, on ne peut formellement exclure que dans le contexte d'une délétion multigénique, une diminution de l'activité de la TP puisse devenir pathogène, par interaction avec les autres gènes délétés, affectant peut-être le cours de la maladie, avec un mécanisme différent de celui en œuvre dans la MNGIE, donc sur le modèle d'une maladie multigénique (48). Pour répondre à cette question, il serait intéressant de mener une étude de corrélation génotypes-phénotypes de cette région (49) en comparant spécifiquement le phénotype de patients présentant des remaniements 22qter comparables avec et sans implication de *TYMP*.

Nous avons envisagé jusqu'à maintenant le caractère pathogène de la délétion de *TYMP* sous l'angle de la maladie mendélienne associée, la MNGIE de type 1. On peut également

s'interroger de manière plus générale sur l'implication de *TYMP* dans les maladies complexes, même si cela n'est plus pertinent pour comprendre le phénotype présenté par notre patient. En effet, certaines maladies complexes sont associées à des variants de gènes impliqués dans les maladies mendéliennes (50). Par exemple, les individus hétérozygotes, porteurs sains de mutations du gène *GBA* codant pour l'enzyme lysosomale  $\alpha$ -glucocérébrosidase, responsables de la maladie de Gaucher, présentent un risque accru de développer une maladie de Parkinson (51, 52). De plus, dans des études GWAS, il a été mis en évidence un enrichissement de variants communs associés à des maladies complexes au niveau des loci des maladies mendéliennes, ce qui a même fait suggérer l'existence d'un « code » mendélien des maladies complexes (53). A l'inverse, certains variants sont probablement bénéfiques, à l'instar du gène *ASS1* pour lequel certains polymorphismes semblent avoir subi une sélection positive (54). Concernant *TYMP*, Guillem et al. ont identifié une association entre un polymorphisme de ce gène et la survie de patients après transplantation allogénique de cellules souches hématopoïétiques HLA-identiques (55) ce qui pourrait refléter le rôle de *TYMP* dans la réparation des dommages de l'ADN. La question de l'implication de variants et délétions de *TYMP* dans les maladies complexes est ouverte.

De manière plus générale, cette étude pose la question de la contribution possible des gènes impliqués dans les maladies héréditaires du métabolisme à la pathogénicité supposée des gènes haploinsuffisants situés à proximité d'un gène métabolique. Pour apporter des éléments de réponse de manière globale, il serait intéressant de rechercher une surreprésentation des gènes métaboliques à proximité des gènes haploinsuffisants. En pratique, on pourrait comparer la distance moyenne minimale entre gènes haploinsuffisants et gènes métaboliques les plus proches à la distance moyenne minimale de tous les autres gènes aux gènes métaboliques les plus proches. Un enrichissement des gènes métaboliques à proximité des gènes haploinsuffisants plaiderait pour une contribution de ces gènes à la pathogénicité supposée de leurs voisins et de manière indirecte pour leur propre haploinsuffisance.

### 4.3 Age, pathologie et métabolisme

L'étude des valeurs « normales » de notre cohorte de référence a mis en évidence une variation des concentrations urinaires des métabolites avec l'âge, nécessitant de réaliser une normalisation du jeu de données sur ce paramètre. Cette dépendance avec l'âge, déjà bien

connue (56), est maintenant l'objet de travaux dédiés qui cherchent à mettre en évidence une signature métabolique du vieillissement (57–59). Il serait donc intéressant de définir une signature métabolique de l'âge à partir de notre jeu de données et de tester sa robustesse en tentant de prédire l'âge vrai des individus à partir de leur âge métabolique. Nos résultats préliminaires montrent la faisabilité et l'intérêt de ce type d'étude métabolique dans l'exploration des mécanismes physiopathologies des maladies du développement. Dans l'ensemble, les patients ont une tendance à exprimer des marqueurs plus tardifs que leur âge ( $p = 0.003$ , cf ci-dessus; Figure 5A) suggérant que des mécanismes de vieillissement précoce aient un rôle important dans ces pathologies, comme cela a été suggéré depuis longtemps dans le syndrome de Down (60). De plus, les profils de certains patients sont hétérogènes, suggérant la coexistence entre l'accélération de certaines voies tardives (vieillesse précoce) et le ralentissement dans la mise en place d'autres voies métaboliques (retard développemental). En général, les marqueurs de retard développemental orientent vers le mécanisme déficitaire. Les marqueurs de vieillissement précoce peuvent orienter vers des mécanismes qui acquièrent, dans le contexte de la pathologie, une fonction compensatrice partiellement inadaptée. Il est intéressant de rapprocher ces résultats aux études suggérant que le déficit intellectuel, associé au syndrome de Down, pourrait dériver de mécanismes de type compensatoire, i.e., la surexpression anticipée de gènes de maintenance neuronale et une augmentation du métabolisme du glucose (61).

#### 4.4 Perspectives

Les perspectives pour ce projet sont de plusieurs ordres. Il est primordial d'étudier de nouveaux patients porteurs de délétions 22q13.3 impliquant *TYMP* ou de délétions affectant *NAT8L* afin de valider ces premiers résultats préliminaires. Un dosage ciblé de la thymine, de l'uracile et de l'acide N-acétylaspartique avec des étalons marqués de nature similaire permettrait d'améliorer la puissance de l'analyse.

De plus, il serait intéressant de reprendre l'analyse en étudiant non plus les concentrations des métabolites mais les rapports entre ces concentrations. Ceci permettrait d'augmenter le nombre de marqueurs potentiels. Cette approche est utilisée dans les études de GWAS et augmente la sensibilité des analyses (13).

L'analyse des profils métaboliques de patients présentant des délétions impliquant des gènes haploinsuffisants situés à proximité de gènes métaboliques (cf Annexe Table 3) permettrait d'augmenter le rendement de découverte de nouveaux marqueurs métaboliques à intérêt diagnostique potentiel. Il serait également intéressant de cibler les délétions affectant des gènes métaboliques susceptibles d'être associés à des variations quantitatives pour maximiser la probabilité d'observer des altérations des profils métaboliques (cf Annexe Table 4).

L'augmentation du nombre de patients présentant un même type de désordre génomique permettrait d'avoir des groupes suffisamment importants pour déduire des signatures métaboliques statistiquement robustes. Ceci permettrait de suggérer de nouvelles hypothèses physiopathologiques impliquant des gènes n'ayant pas de rôle connu dans le métabolisme.

Il serait également intéressant d'analyser les différences en terme d'anticipation/retard à partir des gènes présents/absents dans les déséquilibres chromosomiques d'intérêt.

Nous pourrions poursuivre l'étude avec l'utilisation de nouveaux équipements de spectrométrie de masse couplés à la chromatographie en phase liquide (Q-Exactive) offrant une couverture plus importante du métabolome (environ 3000 métabolites). On se rapprocherait alors véritablement d'une analyse croisant génome et métabolome dans leur ensemble. Dans ce cadre, notre étude suggère qu'il faille privilégier des techniques permettant de quantifier avec précision les métabolites appartenant à des voies connues pour être associées à des maladies héréditaires du métabolisme. Cela implique notamment l'usage d'étalons internes marqués aux isotopes stables correspondant à un nombre aussi large que possible de ces métabolites.

Enfin, une fois l'existence des altérations métaboliques démontrée, il serait intéressant d'évaluer leur caractère pathogène dans le contexte des déséquilibres génomiques. Il faudrait pour cela mener une étude de corrélation phénotype/« métabotype »/génotype permettant de mettre en évidence des modifications de phénotype (amélioration/aggravation) liées à la présence ou l'absence des altérations métaboliques. Le choix devrait se porter préférentiellement sur les gènes à la fois impliqués dans des maladies héréditaires du métabolisme central (du fait de leur sévérité habituelle) et ayant une expression biochimique spécifique à l'état homozygote, et connue ou suspectée à l'état hétérozygote (associations SNPs, désordres génomiques). Ceci permettrait de proposer des pistes thérapeutiques.

Ces questions peuvent être généralisées au-delà des désordres génomiques. Des mutations délétères hétérozygotes impliquant des gènes métaboliques et générant un phénotype métabolique modéré peuvent-elles modifier le phénotype associé à des mutations délétères



homozygotes dans d'autres gènes ? Dans ce cadre, il serait intéressant de génotyper une sélection ou la totalité des 250 gènes du métabolisme central (ce qui sera réalisable par une puce en cours de développement) chez des patients atteints de maladies génétiques associés à une grande hétérogénéité clinique. Il est tentant de suggérer que le degré de sévérité puisse dépendre, peut-être en grande partie, de l'existence de mutations hétérozygotes dans des gènes métaboliques.

## Partie 5 Conclusion

---

Cette étude pilote propose de nouveaux marqueurs métaboliques urinaires de délétions hétérozygotes de gènes métaboliques : augmentation de l'uracile et de la thymine dans les délétions de *TYMP* et diminution de l'acide N-acétylaspartique dans les délétions de *NAT8L*.

Ces marqueurs présenteraient un intérêt clinique pour la détection de délétions affectant les gènes haploinsuffisants *SHANK3* et *NSD2*, respectivement situés à proximité de *TYMP* et *NAT8L* et impliqués dans des cas de déficience intellectuelle pour le premier et dans un syndrome polymalformatif pour le second. Ces résultats préliminaires doivent être confirmés chez de nouveaux patients.

Cette étude illustre la possibilité de variations modérées du métabolisme central, secondaires à des causes génétiques et pose la question de leur implication dans le phénotype présenté par les patients. Leur étude nécessite le couplage entre des analyses génétiques à haute débit (CGH arrays, exomes, puces) et le développement de techniques robustes de quantification des concentrations absolues d'un grand nombre de métabolites, ciblant plus particulièrement les voies métaboliques connues pour être associées à des maladies héréditaires du métabolisme. Par ailleurs, il sera important d'augmenter le nombre de métabolites détectables en pratique hospitalière grâce aux nouvelles technologies de spectrométrie de masse à haute résolution, notamment dans le but de chercher d'éventuels métabolites spécifiques d'un grand nombre de maladies héréditaires affectant le métabolisme énergétique dont les marqueurs connus sont peu informatifs.

## Partie 6 BIBLIOGRAPHIE

---

1. G. Kastenmüller, J. Raffler, C. Gieger, K. Suhre, Genetics of human metabolism: an update. *Hum. Mol. Genet.* **24**, R93–R101 (2015).
2. V. Malan, S. Romana, Diagnosis of chromosomal abnormalities by array CGH in constitutional pathology: the end of the first-line karyotype. *Arch. Pédiatrie* **19**, 437–442 (2012).
3. B. Crespi, Diametric gene-dosage effects as windows into neurogenetic architecture. *Curr. Opin. Neurobiol.* **23**, 143–151 (2013).
4. R. Luo *et al.*, Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. *Am. J. Hum. Genet.* **91**, 38–55 (2012).
5. N. Brunetti-Pierrri *et al.*, Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat. Genet.* **40**, 1466–1471 (2008).
6. M. Laccaria *et al.*, A duplication CNV that conveys traits reciprocal to metabolic syndrome and protects against diet-induced obesity in mice and men. *PLoS Genet.* **8**, e1002713 (2012).
7. M. A. Deardorff *et al.*, Complex management of a patient with a contiguous Xp11.4 gene deletion involving ornithine transcarbamylase: a role for detailed molecular analysis in complex presentations of classical diseases. *Mol. Genet. Metab.* **94**, 498–502 (2008).
8. D. R. Sjarif, J. K. Ploos van Amstel, M. Duran, F. A. Beemer, B. T. Poll-The, Isolated and contiguous glycerol kinase gene disorders: a review. *J. Inherit. Metab. Dis.* **23**, 529–547 (2000).
9. L. Siggberg *et al.*, Familial 6p22.2 duplication associates with mild developmental delay and increased SSADH activity. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **156B**, 448–453 (2011).
10. M. A. Elkoushy, S. Andonian, Characterization of patients with heterozygous cystinuria. *Urology.* **80**, 795–799 (2012).
11. R. J. Morscher *et al.*, A single mutation in MCCC1 or MCCC2 as a potential cause of positive screening for 3-methylcrotonyl-CoA carboxylase deficiency. *Mol. Genet. Metab.* **105**, 602–606 (2012).
12. J. Raffler *et al.*, Genome-Wide Association Study with Targeted and Non-targeted NMR Metabolomics Identifies 15 Novel Loci of Urinary Human Metabolic Individuality. *PLoS Genet.* **11**, e1005487 (2015).

13. S.-Y. Shin *et al.*, An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **46**, 543–550 (2014).
14. H. Kacser, J. A. Burns, The molecular basis of dominance. *Genetics.* **97**, 639–666 (1981).
15. A. Cornish-Bowden, Dominance is not inevitable. *J. Theor. Biol.* **125**, 333–338 (1987).
16. G. Seidner *et al.*, GLUT-1 deficiency syndrome caused by haploinsufficiency of the blood-brain barrier hexose carrier. *Nat. Genet.* **18**, 188–191 (1998).
17. P. M. van Hasselt *et al.*, Monocarboxylate transporter 1 deficiency and ketone utilization. *N. Engl. J. Med.* **371**, 1900–1907 (2014).
18. M. Coutelier *et al.*, Alteration of ornithine metabolism leads to dominant and recessive hereditary spastic paraplegia. *Brain J. Neurol.* **138**, 2191–2205 (2015).
19. G. Ho *et al.*, Maternal riboflavin deficiency, resulting in transient neonatal-onset glutaric aciduria Type 2, is caused by a microdeletion in the riboflavin transporter gene GPR172B. *Hum. Mutat.* **32**, E1976–1984 (2011).
20. A. Dervent *et al.*, Photosensitive absence epilepsy with myoclonias and heterozygosity for succinic semialdehyde dehydrogenase (SSADH) deficiency. *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.* **115**, 1417–1422 (2004).
21. Y. H. Hung *et al.*, Neurological Dysfunction in Early Maturity of a Model for Niemann-Pick C1 Carrier Status. *Neurother. J. Am. Soc. Exp. Neurother.* (2016), doi:10.1007/s13311-016-0427-5.
22. E. Napoli *et al.*, Mitochondrial Citrate Transporter-dependent Metabolic Signature in the 22q11.2 Deletion Syndrome. *J. Biol. Chem.* **290**, 23240–23253 (2015).
23. A. Verloes *et al.*, Diagnostic investigations for an unexplained developmental disability. *Arch. Pédiatrie* **19**, 194–207 (2012).
24. H. M. Engbers *et al.*, Yield of additional metabolic studies in neurodevelopmental disorders. *Ann. Neurol.* **64**, 212–217 (2008).
25. C. D. M. van Karnebeek, M. Shevell, J. Zschocke, J. B. Moeschler, S. Stockler, The metabolic evaluation of the child with an intellectual developmental disorder: diagnostic algorithm for identification of treatable causes and new digital resource. *Mol. Genet. Metab.* **111**, 428–438 (2014).
26. D. T. Miller *et al.*, Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* **86**, 749–764 (2010).
27. R. Possemato *et al.*, Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature.* **476**, 346–350 (2011).
28. I. Nishino, A. Spinazzola, M. Hirano, Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. *Science.* **283**, 689–692 (1999).

29. L. D. Fairbanks, A. M. Marinaki, E. A. Carrey, S. R. Hammans, J. A. Duley, Deoxyuridine accumulation in urine in thymidine phosphorylase deficiency (MNGIE). *J. Inherit. Metab. Dis.* **25**, 603–604 (2002).
30. C. M. Durand *et al.*, Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat. Genet.* **39**, 25–27 (2007).
31. A. Guilmatre, G. Huguet, R. Delorme, T. Bourgeron, The emerging role of SHANK genes in neuropsychiatric disorders. *Dev. Neurobiol.* **74**, 113–122 (2014).
32. H. S. Scott *et al.*, Identification of mutations in the alpha-L-iduronidase gene (IDUA) that cause Hurler and Scheie syndromes. *Am. J. Hum. Genet.* **53**, 973–986 (1993).
33. E. Wiame *et al.*, Molecular identification of aspartate N-acetyltransferase and its mutation in hypocacetylaspartia. *Biochem. J.* **425**, 127–136 (2010).
34. F. Baertling *et al.*, Malonic aciduria: long-term follow-up of new patients detected by newborn screening. *Eur. J. Pediatr.* **173**, 1719–1722 (2014).
35. S. Petrovski, Q. Wang, E. L. Heinzen, A. S. Allen, D. B. Goldstein, Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet.* **9**, e1003709 (2013).
36. A. M. Molloy *et al.*, A Common Polymorphism in HIBCH Influences Methylmalonic Acid Concentrations in Blood Independently of Cobalamin. *Am. J. Hum. Genet.* **98**, 869–882 (2016).
37. C. Betancur, J. D. Buxbaum, SHANK3 haploinsufficiency: a “common” but underdiagnosed highly penetrant monogenic cause of autism spectrum disorders. *Mol. Autism.* **4**, 17 (2013).
38. L. Soorya *et al.*, Prospective investigation of autism and genotype-phenotype correlations in 22q13 deletion syndrome and SHANK3 deficiency. *Mol. Autism.* **4**, 18 (2013).
39. R. Martí, Y. Nishigaki, M. Hirano, Elevated plasma deoxyuridine in patients with thymidine phosphorylase deficiency. *Biochem. Biophys. Res. Commun.* **303**, 14–18 (2003).
40. R. Martí *et al.*, Definitive diagnosis of mitochondrial neurogastrointestinal encephalomyopathy by biochemical assays. *Clin. Chem.* **50**, 120–124 (2004).
41. R. Martí *et al.*, Late-onset MNGIE due to partial loss of thymidine phosphorylase activity. *Ann. Neurol.* **58**, 649–652 (2005).
42. V. Disciglio *et al.*, Interstitial 22q13 deletions not involving SHANK3 gene: a new contiguous gene syndrome. *Am. J. Med. Genet. A.* **164A**, 1666–1676 (2014).
43. R. E. Frye *et al.*, Mitochondrial Dysfunction may explain symptom variation in Phelan-McDermid Syndrome. *Sci. Rep.* **6**, 19544 (2016).

44. N. S. Brown, R. Bicknell, Thymidine phosphorylase, 2-deoxy-D-ribose and angiogenesis. *Biochem. J.* **334** ( Pt 1), 1–8 (1998).
45. T. Matsuura *et al.*, Thymidine phosphorylase expression is associated with both increase of intratumoral microvessels and decrease of apoptosis in human colorectal carcinomas. *Cancer Res.* **59**, 5037–5040 (1999).
46. R. Massa *et al.*, Late-onset MNGIE without peripheral neuropathy due to incomplete loss of thymidine phosphorylase activity. *Neuromuscul. Disord.* **19**, 837–840 (2009).
47. A. Nalini, N. Gayathri, Mitochondrial neurogastrointestinal encephalopathy in an Indian family with possible manifesting carriers of heterozygous TYMP mutation. *J. Neurol. Sci.* **309**, 131–135 (2011).
48. N. Katsanis *et al.*, Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science.* **293**, 2256–2259 (2001).
49. S. M. Sarasua *et al.*, 22q13.2q13.32 genomic regions associated with severity of speech delay, developmental delay, and physical features in Phelan-McDermid syndrome. *Genet. Med.* **16**, 318–328 (2014).
50. J. R. Lupski, J. W. Belmont, E. Boerwinkle, R. A. Gibbs, Clan genomics and the complex architecture of human disease. *Cell.* **147**, 32–43 (2011).
51. M. F. Coutinho, S. Alves, From rare to common and back again: 60years of lysosomal dysfunction. *Mol. Genet. Metab.* **117**, 53–65 (2016).
52. E. Sidransky *et al.*, Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N. Engl. J. Med.* **361**, 1651–1661 (2009).
53. D. R. Blair *et al.*, A nondegenerate code of deleterious variants in Mendelian loci contributes to complex disease risk. *Cell.* **155**, 70–80 (2013).
54. L. B. Barreiro, G. Laval, H. Quach, E. Patin, L. Quintana-Murci, Natural selection has driven population differentiation in modern humans. *Nat. Genet.* **40**, 340–345 (2008).
55. V. Guillem *et al.*, A polymorphism in the TYMP gene is associated with the outcome of HLA-identical sibling allogeneic stem cell transplantation. *Am. J. Hematol.* **88**, 883–889 (2013).
56. O. Boulat, M. Gradwohl, V. Matos, J.-P. Guignard, C. Bachmann, Organic acids in the second morning urine in a healthy Swiss paediatric population. *Clin. Chem. Lab. Med.* **41**, 1642–1658 (2003).
57. S. Collino *et al.*, Metabolic signatures of extreme longevity in northern Italian centenarians reveal a complex remodeling of lipids, amino acids, and gut microbiota metabolism. *PLoS One.* **8**, e56564 (2013).
58. H. Gu *et al.*, <sup>1</sup>H NMR metabolomics study of age profiling in children. *NMR Biomed.* **22**, 826–833 (2009).

59. Z. Yu *et al.*, Human serum metabolic profiles are age dependent. *Aging Cell*. **11**, 960–967 (2012).
60. O. Ogawa, G. Perry, M. A. Smith, The “Down’s” side of mitochondria. *Dev. Cell*. **2**, 255–256 (2002).
61. E. Head, I. T. Lott, D. Patterson, E. Doran, R. J. Haier, Possible compensatory events in adult Down syndrome brain prior to the development of Alzheimer disease neuropathology: targets for nonpharmacological intervention. *J. Alzheimers Dis*. **11**, 61–76 (2007).

## Partie 7 Annexes

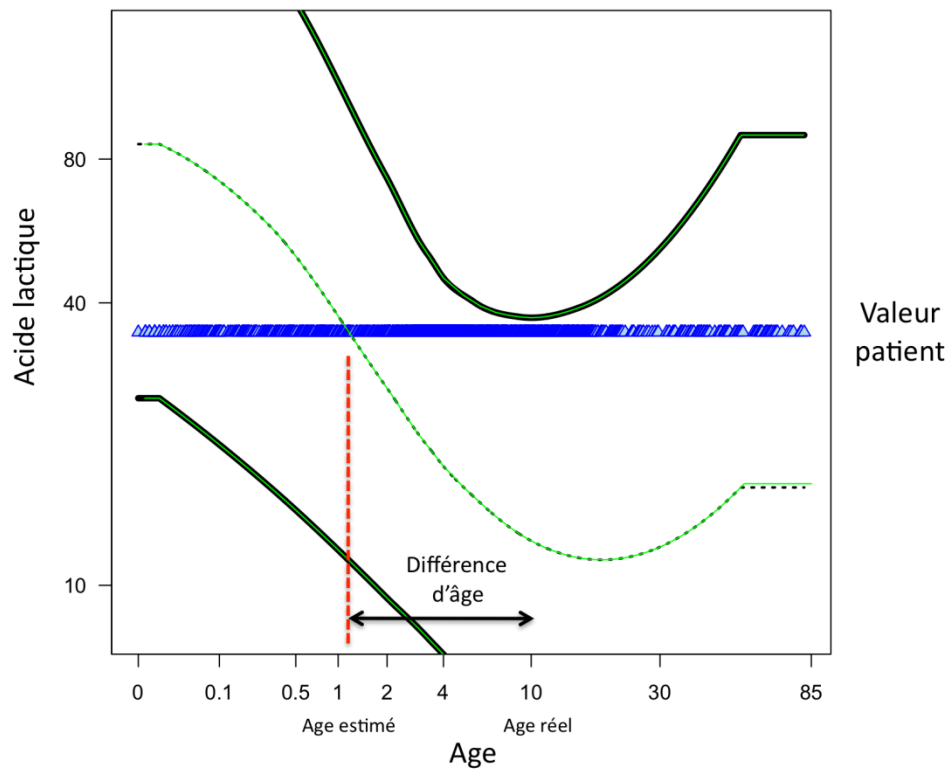


Figure S1. Calcul de la différence d'âge avance/anticipation. Dans cet exemple, le patient a 10 ans (âge réel) et présente une concentration d'acide lactique urinaire de 35  $\mu\text{mol}/\text{mmol}$  (matérialisée par la courbe bleue). Cette valeur correspond à la valeur médiane pour un âge de 1,1 an (intersection courbe bleue et courbe verte-pointillé noir centrale). L'âge métabolique optimal estimé pour la valeur de concentration du patient est donc de 1,1 an. La différence d'âge est de -8,9 ans correspondant à un retard.



Table Annexe 1. Liste des désordres génomiques présentés par les patients de la cohorte.

Patients	Sexe	Age	Diagnostic	Chromosomes	Del/Dup	Borne Inf	Borne Sup	Taille	Réurrence	Conséquences
1	F	2.3	del 22q13.2qter	22	del	43006204	51193680	8.2 Mb	Non recurrent	Pathogène
2	M	0	der(8)t(4;8)(p16.3;p23.3p23.1)	4	del	1	9628003	9.6 Mb	Recurrent	Pathogène
				8	dup	1	7753583	7.8 Mb		
3	M	0.7	del Xp21.3p21.1	X	del	28919849	35296972	6.3 Mb	Non recurrent	Pathogène
4	M	17	de 16p11.2	16	del	29673954	30198600	525 kb	Recurrent	Prédisposition tr. neuro. dev.
5	M	2.2	del 1q21.1 / dup 16p11.2	1	del	145076147	147786706	2.7 Mb	Recurrent	Prédisposition tr. neuro. dev.
				16	dup	29674848	30198600	524 kb	Recurrent	
6	M	5.7	del 16q23.2q24.1	16	del	81228359	85111857	3.9 kb	Non recurrent	Pathogène
7	F	5.3	del 2p16.3	2	del	51122091	51314430	192 kb	Non recurrent	Prédisposition tr. neuro. dev.
8	M	0.8	dup 1q21.1	1	dup	145747214	147786706	2 Mb	Recurrent	Prédisposition tr. neuro. dev.
9	F	0.5	dup 1q21.1	1	dup	145632334	145747269	115 kb	Non récurrent	Bénin
10	M	3.9	dup 6p25.1	6	dup	6243694	6760582	516 kb	Non recurrent	Bénin
11	F	2.7	dup 11q22.3 / dup 19q13.42q13.43	11	dup	106737140	107260943	524 kb	Non recurrent	Bénin
				19	dup	56238771	56353892	115 kb		Bénin
12	M	5.1	dup 1q21.1 / dup 17q25.1	1	dup	145632334	145747269	114 kb	Non récurrent	Bénin
				17	dup	71963482	72671896	708 kb		
13	M	3.5	arr (X)x1,(Y)x2							
14	M	4.4	del 7q31.1	7	del	110767384	111202026	434 kb	Non recurrent	VOUS
15	M	3.3	dup 7q36.1	7	dup	152351807	153360454	1 Mb	Non recurrent	VOUS
16	F	10.2	dup 8q21.13	8	dup	82484070	84280878	1.8 Mb	Non recurrent	VOUS
17	M	3.1	dup 13q31.3	13	dup	94844214	95363143	518 kb	Non recurrent	VOUS
18	F	1.2	dup 15q25.2q25.3	15	dup	85124542	86487178	1.4 Mb	Recurrent	VOUS
19	F	1.1	dup Xp22.11	X	dup	23926318	24495376	569 kb	Non recurrent	VOUS
20	M	3.5	del 17p12	17	del	14111772	15442066	1.3 Mb	Recurrent	Pathogène
21	F	0.2	del 17q22	17	del	55761943	58025806	2.2 Mb	Recurrent	Pathogène
22	F	0.3	r(X),45X : del Xp22.33p21.1 del Xq21.31q28	X	del	61091	34052781	34 Mb	Non recurrent	Pathogène
				X	del	88988259	155190083	62 Mb		
23	F	2.3	del Xq28 / dup 7p21.3	X	del	154120738	154490709	370 kb	Non recurrent	Pathogène
				7	dup	9436795	10676383	1.2 Mb		

24	F	6.4	del 7q34	7	del	138810389	139966852	1.1 Mb	Non recurrent	Pathogène
25	M	0.2	del 9p24.3p24.1 /	9	del	1	749364	749 kb	Non recurrent	Pathogène
			trip 9p22.3 /	9	trip	7608025	10786134	3.2 Mb		
			dup 9p22.3 /	9	dup	10882873	12196201	1.3 Mb		
			15q13.3	15	dup	32076948	32510863	434 kb		Bénin
26	M	0.8	del 15q11.2q13.2	15	del	23699701	30322138	6.6 Mb	Recurrent	Pathogène
27	M	0.9	gain 15q11.1q13.1 (invdup15)	15	trip	20102541	28525460	8.4 Mb	Recurrent	Pathogène
28	M	0.0	del 15q11.2q13.1	15	del	23699701	28525460	4.8 Mb	Recurrent	Pathogène
29	M	5.8	del 15q11.2q13.2	15	del				Recurrent	Pathogène
30	M	0.0	del 15q11.2q13.1	15	del	23739358	28119374	4.4 Mb	Recurrent	Pathogène

del : délétion ; dup : duplication ; tr. neuro. dev. : troubles neuro-développementaux ; VOUS : variant de signification inconnue ; Borne Inf/Borne sup : localisation chromosomique (GRCh37)

Table Annexe 2. Patients présentant des désordres génomiques impliquant des gènes impliqués dans le métabolisme selon Possemato et al.

Patients	Age	Remaniement	Chr	Borne.Inf	Borne.Sup	Symbole	Nom Complet	Classification
1	2.3	22q13.2qter	22 del	43006204	51193680	A4GALT	alpha 1,4-galactosyltransferase (globotriaosylceramide synthase)	Glycan
1	2.3	22q13.2qter	22 del	43006204	51193680	ALG12	asparagine-linked glycosylation 12 homolog (S. cerevisiae, alpha-1,6-mannosyltransferase)	Glycan
1	2.3	22q13.2qter	22 del	43006204	51193680	ARSA	Arylsulfatase A	Multipurpose
1	2.3	22q13.2qter	22 del	43006204	51193680	ATP5L2	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit G2	Proton Transport
1	2.3	22q13.2qter	22 del	43006204	51193680	CERK	ceramide kinase	Sphingolipid
1	2.3	22q13.2qter	22 del	43006204	51193680	CPT1B	carnitine palmitoyltransferase 1B (muscle)	Fatty Acid
1	2.3	22q13.2qter	22 del	43006204	51193680	CYB5R3	cytochrome b5 reductase 3	Redox
1	2.3	22q13.2qter	22 del	43006204	51193680	MCAT	malonyl CoA:ACP transferase (mitochondrial)	Fatty Acid
1	2.3	22q13.2qter	22 del	43006204	51193680	MIOX	myo-inositol oxygenase	Inositol Phosphate
1	2.3	22q13.2qter	22 del	43006204	51193680	PNPLA3	patatin-like phospholipase domain containing 3	Fatty Acid
1	2.3	22q13.2qter	22 del	43006204	51193680	PNPLA5	patatin-like phospholipase domain containing 5	Fatty Acid
1	2.3	22q13.2qter	22 del	43006204	51193680	SULT4A1	sulfotransferase family 4A, member 1	Sulfate
1	2.3	22q13.2qter	22 del	43006204	51193680	TYMP	Thymidine phosphorylase	Pyrimidine
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	ACOX3	acyl-Coenzyme A oxidase 3, pristanoyl	Fatty Acid
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	ATP5I	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit E	Proton Transport
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	DGKQ	diacylglycerol kinase, theta 110kDa	Membrane Lipid
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	IDUA	iduronidase, alpha-L-	Glycan Degradation
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	MAN2B2	mannosidase, alpha, class 2B, member 2	Glycan Degradation
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	MFSD10	Major Facilitator Superfamily Domain Containing 10	Other Transport
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	MFSD7	Major Facilitator Superfamily Domain Containing 7	Other Transport
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	NAT8L	N-acetyltransferase 8-like	Multipurpose
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	PDE6B	phosphodiesterase 6B, cGMP-specific, rod, beta (congenital stationary night blindness 3, autosomal dominant)	Signalling
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	PIGG	phosphatidylinositol glycan anchor biosynthesis, class G	Glycan Anchor
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	4 del	1	9628003	SLC26A1	solute carrier family 26 (sulfate transporter), member 1	Small Molecule Transport
2	0.0	der(8)t(4;8)(p16.3;p23.3p23.1)	8 dup	1	7753583	AGPAT5	1-acylglycerol-3-phosphate O-acyltransferase 5 (lysophosphatidic acid acyltransferase, epsilon)	Membrane Lipid
4	17.0	del_16p11.2	16 del	29673954	30198600	ALDOA	aldolase A, fructose-bisphosphate	Glycolysis
4	17.0	del_16p11.2	16 del	29673954	30198600	CDIPT	CDP-diacylglycerol--inositol 3-phosphatidyltransferase (phosphatidylinositol synthase)	Inositol Phosphate
4	17.0	del_16p11.2	16 del	29673954	30198600	GDPD3	glycerophosphodiester phosphodiesterase domain containing 3	Other
4	17.0	del_16p11.2	16 del	29673954	30198600	KCTD13	potassium channel tetramerisation domain containing 13	Ion Transport
4	17.0	del_16p11.2	16 del	29673954	30198600	QPRT	quinolinate phosphoribosyl(nicotinate-nucleotide pyrophosphorylase )	NAD
5	2.2	del_1q21.1/_du p_16p11.2	16 dup	29674848	30198600	ALDOA	aldolase A, fructose-bisphosphate	Glycolysis
5	2.2	del_1q21.1/_du p_16p11.2	16 dup	29674848	30198600	CDIPT	CDP-diacylglycerol--inositol 3-phosphatidyltransferase (phosphatidylinositol synthase)	Inositol Phosphate
5	2.2	del_1q21.1/_du p_16p11.2	16 dup	29674848	30198600	GDPD3	glycerophosphodiester phosphodiesterase domain containing 3	Other

5	2.2	del_1q21.1/_/du p_16p11.2	16	dup	29674848	30198600	KCTD13	potassium channel tetramerisation domain containing 13	Ion Transport
5	2.2	del_1q21.1/_/du p_16p11.2	16	dup	29674848	30198600	QPRT	quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating))	NAD
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	ADAD2	adenosine deaminase domain containing 2	Purine
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	ATP2C2	ATPase, Ca <sup>++</sup> transporting, type 2C, member 2	Ion Transport
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	BCMO1	beta carotene 15, 15 monooxygenase 1	Vitamin A
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	HSD17B2	hydroxysteroid (17-beta) dehydrogenase 2	Steroid
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	HSDL1	hydroxysteroid dehydrogenase like 1	Steroid
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	KCNG4	potassium voltage-gated channel, subfamily G, member 4	Ion Transport
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	MLYCD	malonyl-CoA decarboxylase	Fatty Acid
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	PLCG2	Phospholipase C, gamma 2 (phosphatidylinositol-specific)	Membrane Lipid
6	5.7	perte_16q23.2q2 4.1	16	del	81228359	85111857	SDR42E1	short chain dehydrogenase/reductase family 42E, member 1	Fatty Acid
7	0.8	gain_1q21.1	1	dup	145747214	147786706	FMO5	flavin containing monooxygenase 5	Detox
8	2.2	del_1q21.1/_/du p_16p11.2	1	del	145076147	147786706	FMO5	flavin containing monooxygenase 5	Detox
8	2.2	del_1q21.1/_/du p_16p11.2	1	del	145076147	147786706	NUDT17	nudix (nucleoside diphosphate linked moiety X)-type motif 17	Nucleotide
9	6.4	del_7q34	7	del	138810389	139966852	TBXAS1	thromboxane A synthase 1 (platelet, cytochrome P450, family 5, subfamily A)	Hormone
10	10.2	dup_8q21.13	8	dup	82484070	84280878	IMPA1	inositol(myo)-1(or 4)-monophosphatase 1	Inositol Phosphate Small Molecule Transport
10	10.2	dup_8q21.13	8	dup	82484070	84280878	SLC10A5	solute carrier family 10 (sodium/bile acid cotransporter family), member 5	
11	0.2	del_9p24.3p24.1 /_/trip 9p22.3/_/dup_9 p22.3/_/15q13.3	15	dup	32076948	32510863	CHRNA7	Cholinergic receptor, nicotinic, alpha 7	Ion Transport
12	2.7	dup_11q22.3/_/ dup_19q13.42q1 3.43	11	dup	106737140	107260943	GUCY1A2	guanylate cyclase 1, soluble, alpha 2	Signalling
13	3.1	dup_13q31.3	13	dup	94844214	95363143	DCT	dopachrome tautomerase (dopachrome delta-isomerase, tyrosine-related protein 2)	Pigment
13	3.1	dup_13q31.3	13	dup	94844214	95363143	TGDS	TDP-glucose 4,6-dehydratase	Nucleotide Sugar
14	0.0	del_15q11.2q13.1	15	del	23699701	28525460	ATP10A	ATPase, class V, type 10A	Small Molecule Transport
14	0.0	del_15q11.2q13.1	15	del	23699701	28525460	GABRA5	Gamma-aminobutyric acid (GABA) A receptor, alpha 5	Ion Transport
14	0.0	del_15q11.2q13.1	15	del	23699701	28525460	GABRB3	Gamma-aminobutyric acid (GABA) A receptor, beta 3	Ion Transport
14	0.0	del_15q11.2q13.1	15	del	23699701	28525460	GABRG3	Gamma-aminobutyric acid (GABA) A receptor, gamma 3	Ion Transport
14	0.0	del_15q11.2q13.1	15	del	23699701	28525460	OCA2	Oculocutaneous albinism II (pink-eye dilution homolog, mouse)	Ion Transport
15	0.8	perte_15q11.2q1 3.2	15	del	23699701	30322138	ATP10A	ATPase, class V, type 10A	Small Molecule Transport
15	0.8	perte_15q11.2q1 3.2	15	del	23699701	30322138	GABRA5	Gamma-aminobutyric acid (GABA) A receptor, alpha 5	Ion Transport
15	0.8	perte_15q11.2q1 3.2	15	del	23699701	30322138	GABRB3	Gamma-aminobutyric acid (GABA) A receptor, beta 3	Ion Transport
15	0.8	perte_15q11.2q1 3.2	15	del	23699701	30322138	GABRG3	Gamma-aminobutyric acid (GABA) A receptor, gamma 3	Ion Transport
15	0.8	perte_15q11.2q1 3.2	15	del	23699701	30322138	OCA2	Oculocutaneous albinism II (pink-eye dilution homolog, mouse)	Ion Transport
16	0.9	gain_15q11.1q1 3.1	15	dup	20102541	28525460	ATP10A	ATPase, class V, type 10A	Small Molecule Transport
16	0.9	gain_15q11.1q1 3.1	15	dup	20102541	28525460	GABRA5	Gamma-aminobutyric acid (GABA) A receptor, alpha 5	Ion Transport

16	0.9	gain_15q11.1q13.1	15	dup	20102541	28525460	GABRB3	Gamma-aminobutyric acid (GABA) A receptor, beta 3	Ion Transport
16	0.9	gain_15q11.1q13.1	15	dup	20102541	28525460	GABRG3	Gamma-aminobutyric acid (GABA) A receptor, gamma 3	Ion Transport
16	0.9	gain_15q11.1q13.1	15	dup	20102541	28525460	NIPA1	non imprinted in Prader-Willi/Angelman syndrome 1	Ion Transport
16	0.9	gain_15q11.1q13.1	15	dup	20102541	28525460	OCA2	Oculocutaneous albinism II (pink-eye dilution homolog, mouse)	Ion Transport
17	1.2	dup_15q25.2q25.3	15	dup	85124542	86487178	PDE8A	phosphodiesterase 8A	Signalling
17	1.2	dup_15q25.2q25.3	15	dup	85124542	86487178	SLC28A1	solute carrier family 28 (sodium-coupled nucleoside transporter), member 1	Small Molecule Transport
18	0.0	del_15q11.2q13.1	15	del	23739358	28119374	ATP10A	ATPase, class V, type 10A	Small Molecule Transport
18	0.0	del_15q11.2q13.1	15	del	23739358	28119374	GABRA5	Gamma-aminobutyric acid (GABA) A receptor, alpha 5	Ion Transport
18	0.0	del_15q11.2q13.1	15	del	23739358	28119374	GABRB3	Gamma-aminobutyric acid (GABA) A receptor, beta 3	Ion Transport
18	0.0	del_15q11.2q13.1	15	del	23739358	28119374	GABRG3	Gamma-aminobutyric acid (GABA) A receptor, gamma 3	Ion Transport
18	0.0	del_15q11.2q13.1	15	del	23739358	28119374	OCA2	Oculocutaneous albinism II (pink-eye dilution homolog, mouse)	Ion Transport
19	3.5	del_17p12	17	del	14111772	15442066	HS3ST3B1	Heparan sulfate (glucosamine) 3-O-sulfotransferase 3B1	Heparin Sulfate
20	0.2	del_17q22	17	del	55761943	58025806	EPX	eosinophil peroxidase	Other
20	0.2	del_17q22	17	del	55761943	58025806	GDPD1	glycerophosphodiester phosphodiesterase domain containing 1	Other
20	0.2	del_17q22	17	del	55761943	58025806	LPO	lactoperoxidase	Redox
20	0.2	del_17q22	17	del	55761943	58025806	MPO	myeloperoxidase	Other
21	5.1	dup17q25.1/_dup1q21.1	17	dup	71963482	72671896	TTYH2	Tweety homolog 2 (Drosophila)	Ion Transport

Table Annexe 3. Liste des gènes haploinsuffisants situés à moins de 500kb d'un gène métabolique.

Gènes haploinsuffisants : Symboles	Gènes haploinsuffisants : Noms	Gènes métaboliques : Symboles	Gènes métaboliques : Noms	Marqueurs putatifs	Bandes
<i>KCNH2</i>	Potassium channel, voltage-gated, sufamily H, Member 2	<i>NOS3</i>	nitric oxide synthase 3 (endothelial cell)	citrulline	7q36.1
<i>SDHD</i>	Succinate dehydrogenase complex, subunit D	<i>DLAT</i>	Pyruvate dehydrogenase E2	acide lactique	11q23
<i>RPS26</i>	Ribosomal protein S26	<i>SUOX</i>	sulfite oxidase	sulfocystéine	12q13
<i>SHANK3</i>	SH3 and multiple ankyrin repeat domain 3	<i>TYMP</i>	Thymidine phosphorylase	uracile, thymine	22q13.3
<i>ATM</i>	ataxia-telangiectasia mutated gen	<i>ACAT1</i>	acetyl-CoA acetyltransferase 1	acide 2-methylacetoacétique, tiglylglycine	11q22-q23
<i>NFIX</i>	Nuclear Factor I/X	<i>GCDH</i>	glutaryl-CoA dehydrogenase	acide glutarique, acide 3-hydroxyglutarique	19q13.3
<i>ZIC2</i>	Zinc finger protein of cerebellum 2	<i>PCCA</i>	propionyl CoA carboxylase, alpha polypeptide	acides 3-hydroxypropionique et methylcitrique	13q32
<i>ANKRD11</i>	Ankyrin repeat domain-containing protein 11	<i>ACSF3</i>	Combined malonic and methylmalonic aciduria	acides malonique et methylmalonique	16q24.3
<i>KAT6B</i>	Lysine acetyltransferase 6B	<i>ADK</i>	Adenosine kinase	méthionine	10q22.2
<i>COL6A1</i>	Collagen, Type VI, Alpha-1	<i>FTCD</i>	formiminotransferase cyclodeaminase	acide formiminoglutamique	21q22.3
<i>BRCA1</i>	Breast cancer 1 gene	<i>G6PC</i>	glucose 6 phosphatase, catalytic I	acide lactique, acide urique	17q21
<i>BCMO1</i>	Beta-carotene 15, 15-prime-monoxygenase 1	<i>GCSH</i>	glycine cleavage system protein H (aminomethyl carrier)	glycine	16q23.2
<i>STK11</i>	Serine/threonine protein kinase 1	<i>GAMT</i>	guanidinoacetate N-methyltransferase	acide guanidinoacétique	19p13.3
<i>BUB1B</i>	Budding uninhibited by enzimidazoles 1	<i>IVD</i>	isovaleryl-CoA dehydrogenase	isovalerylglycine	15q15
<i>TRPV4</i>	Transient receptor potential cation channel, subfamily V, member 4	<i>MMAB</i>	methylmalonic aciduria (cobalamin deficiency) cblB type	acide methylamlonique	12q24.1
<i>ABL1</i>	Abelson murine leukemia viral oncogene homolog 1	<i>ASS1</i>	argininosuccinate synthase 1	citrulline	9q34.1
<i>ATL1</i>	Atlantin GTPase 1	<i>L2HGDH</i>	L-2-hydroxyglutarate dehydrogenase	acide L-2-hydroxyglutarique	14q22.1
<i>RAD51</i>	RAD51	<i>IVD</i>	isovaleryl-CoA dehydrogenase	isovalerylglycine	15q15.1
<i>RB1</i>	RB1 gene	<i>SUCLA2</i>	succinate-CoA ligase, ADP-forming, beta subunit	acide methylamlonique	13q14.2
<i>GRN</i>	Granulin precursor	<i>NAGS</i>	N-acetylglutamate synthase	glutamine, citrulline, ammoniémie	17q21.32
<i>DICER1</i>	Dicer, Drosophila, homolog of , 1	<i>GLRX5</i>	Glutaredoxin / iron-sulfur (Fe-S) clusters	glycine	14q32.13
<i>PLIN1</i>	Perilipin 1	<i>IDH2</i>	isocitrate dehydrogenase 2 (NADP+), mitochondrial	acide D-2-hydroxyglutarique	15q26
<i>LEMD3</i>	LEM Domain-containing protein 3	<i>GNS</i>	N-acetylglucosamine-6-sulfatase	heparan sulfate	12q14
<i>DYRK1A</i>	Dual-specificity tyrosine phosphorylation-regulated kinase 1A	<i>HLCS</i>	holocarboxylase synthetase	acides 3-hydroxypropionique, lactique, 3-hydroxyisovalérique	21q22.13
<i>RPS19</i>	Ribosomal protein S19	<i>BCKDHA</i>	branched chain keto acid dehydrogenase E1, alpha polypeptide	leucine	19q13.2
<i>TP53</i>	Tumor protein p53	<i>ACADVL</i>	acyl-CoA dehydrogenase, very long chain	C14:1-carnitine	17p13.1
<i>GDF5</i>	Growth/Differentiation factor 5	<i>GSS</i>	glutathione synthetase	5-oxoproline	20q11.2

Table Annexe 4. Liste des gènes métaboliques et de leurs marqueurs métaboliques, associés de manière quantitative dans des études GWAS (13, 36).

Noms	Symboles	Bandes	Marqueurs
3-hydroxyisobutyryl-CoA hydrolase	<i>HIBCH</i>	2q32.2	acide méthylmalonique
Acyl-CoA déshydrogénase, C-4 to C-12 straight chain	<i>ACADM</i>	1p31	acetylcarnitine/hexanoylcarnitine
Acyl-CoA déshydrogénase, C-2 to C-3 short chain	<i>ACADS</i>	12q24.31	butyrylcarnitine
Cystathionine-beta-synthase	<i>CBS</i>	21q22.3	betaine
Carnitine palmitoyltransferase 2	<i>CPT2</i>	1p32	glutaroylcarnitine
Dopa décarboxylase (aromatic L-amino acid decarboxylase)	<i>DDC</i>	7p12.2	indoleacetate/ N2,N2-dimethylguanosine
Electron transfer flavoprotein déshydrogénase	<i>ETFDH</i>	4q32-q35	octanoylcarnitine/ X-13435
Glutaryl-CoA déshydrogénase	<i>GCDH</i>	19p13.2	glutaroylcarnitine
Isovaleryl-CoA déshydrogénase	<i>IVD</i>	15q14-q15	propionylcarnitine/ isovalerylcarnitine
5-oxoprolinase (ATP-hydrolysing)	<i>OPLAH</i>	8q24.3	5-oxoproline
Phosphoglycerate déshydrogénase	<i>PHGDH</i>	1p12	serine
Phosphoserine phosphatase	<i>PSPH</i>	7p11.2	serine
Proline déshydrogénase 1	<i>PRODH</i>	22q11.21	valine/proline
Solute carrier family 22 member 5	<i>SLC22A5</i>	5q23.3	valine/ isovalerylcarnitine
Thymidine phosphorylase	<i>TYMP</i>	22q13.33	uridine

Titre en français : Métabolomique par chromatographie en phase gazeuse couplée à la spectrométrie de masse et désordres génomiques détectés par CGH array : biomarqueurs ou mécanismes ?

Résumé (français) : L'objectif de cette étude pilote est d'étudier l'impact des désordres génomiques au niveau métabolique afin de mettre en évidence des marqueurs diagnostiques et de proposer des pistes de mécanismes physiopathologiques. Nous avons analysés les profils d'acides organiques urinaires mesurés par chromatographie en phase gazeuse couplée à la spectrométrie de masse de 30 patients porteurs de désordres génomiques détectés par CGH array. Nous avons observé une augmentation des concentrations urinaires de l'uracile et de la thymine chez une patiente présentant une délétion 22q13.2qter en rapport avec la délétion du gène *TYMP* retrouvé dans ce désordre génomique. Une diminution de l'acide N-acétylaspartique urinaire est retrouvée chez un patient porteur d'un désordre chromosomique complexe emportant une copie du gène *NAT8L* expliquant le phénotype métabolique. De manière intéressante, *TYMP* et *NAT8L* sont situés à proximité des gènes *SHANK3* et *NSD2*. Les signatures métaboliques observées pourraient donc être utilisées comme marqueurs indirects des délétions de *SHANK3* et *NSD2*. Ces résultats sont à valider avec un nombre plus grand de patients. Cette étude illustre la possibilité de variations modérées du métabolisme intermédiaire, secondaires à des causes génétiques et pose la question de leurs implications dans le phénotype présenté par les patients.

Mots clés (français) : CNV, marqueur métabolomique, GC-MS, *TYMP*, *SHANK3*, *NAT8L*, *NSD2*

Titre en anglais : Metabolomics by gas chromatography coupled to mass spectrometry in patients diagnosed with genomic disorders detected by CGH array : biomarkers or mechanisms ?

Abstract (english) : The aim of this pilot study is to assess the effects of genomic disorders at the metabolic level to derive metabolic markers and suggest pathological mechanisms. We analysed urinary organic acids by mass spectrometry coupled to gas chromatography in 30 patients carrying a genomic disorder. We observed in a patient carrying a 22q13.2qter deletion an increase in uracil and thymine due to the deletion of the *TYMP* gene found in this disorder. We found a decrease in N-acetylaspartic acid in a patient carrying a complex genomic disorder deleting one copy of the gene *N8TL*, explaining the biochemical phenotype. Noteworthy, *TYMP* and *NAT8L* are close to the haploinsufficient genes *SHANK3* and *NSD2*. These metabolic signatures could thus be considered as indirect markers of *SHANK3* and *NSD2* deletions. These results need to be validated in a larger cohort. This study highlights the possibility of moderate variations of intermediary metabolism due to genetic causes and asks the question of their involvement in the phenotypes of patients.

Keywords (english) : CNV, metabolomic markers, GC-MS, *TYMP*, *SHANK3*, *NAT8L*, *NSD2*

**Université Paris Descartes,  
Faculté de Médecine Paris Descartes  
15, rue de l'École de Médecine  
75270 Paris cedex 06**