

Université de Montréal

**Impact des antibiotiques céfprozil et céfoxitine sur le  
microbiote *Eggerthella lenta*, lié au métabolisme du  
cardiotonique digoxine**

par Jérémie Auger

Département de pharmacologie et physiologie  
Faculté de médecine

Mémoire présenté  
en vue de l'obtention du grade de maîtrise ès sciences (M. Sc.)  
en pharmacologie  
option pharmacogénomique

Décembre, 2018

© Jérémie Auger, 2018

*Ce mémoire intitulé*

**Impact des antibiotiques céfprozil et céfoxitine sur le microbiote *Eggerthella lenta*, lié au métabolisme de la digoxine**

*Présenté par*  
**Jérémie Auger**

*A été évalué par un jury composé des personnes suivantes*

**Guy Rousseau**  
Président-rapporteur

**Marie-Pierre Dubé**  
Directeur de recherche

**Bertrand Routy**  
Membre du jury

# 1 Résumé

La digoxine est un cardiotonique largement employé pour contrôler les symptômes de l'insuffisance cardiaque et de la fibrillation auriculaire. Il est connu depuis les années 1980 que le métabolite principal de la digoxine, la dihydrodigoxine, est produit exclusivement par le microbiome intestinal (métabolisme de premier passage) et plus précisément la bactérie *Eggerthella lenta*. Aux États-Unis, c'est 14% des participants à une étude qui excrétaient 40% et plus de la dose sous la forme de ce métabolite rapidement éliminable et ayant perdu son affinité pour sa cible. De plus, chaque année, la digoxine est le médicament qui engendre le plus d'hospitalisations pour effets secondaires toxiques. Les effets secondaires très problématiques de la digoxine sont souvent déclenchés par l'ajout d'antibiotiques (surtout les macrolides) à la prescription de digoxine. La théorie explorée ici explique les événements de toxicité chez les patients métabolisateurs. Ces derniers ont une dose quotidienne de maintien de digoxine plus élevée pour compenser l'action de la bactérie et, lorsque ces patients reçoivent un antibiotique pour une infection non reliée à leur condition cardiaque, l'arrêt du métabolisme par le microbiome engendre une augmentation de la biodisponibilité de la digoxine. Si la concentration plasmatique du médicament augmente trop, les effets secondaires peuvent aller jusqu'à causer la mort. Dans le présent projet, nous avons vérifié la sensibilité de *E. lenta* à deux antibiotiques de la famille des céphalosporines de seconde génération, *in vivo* et *in vitro*. Pour les 18 volontaires qui ont été exposés à 2x500mg de céfprozil durant une semaine, il y a une tendance à la baisse de l'abondance de la bactérie d'intérêt (par 58,3% par rapport au niveau initial), mais pas de significativité au niveau des tests statistiques. Pour les échantillons complets de microbiome fécal, mis en culture avec et sans antibiotiques, il y a une différence statistiquement significative avec une valeur-p de 0,0457, alors que la croissance de *E. lenta* a été impactée négativement par l'ajout de céfprozil au milieu de culture. Les résultats valident une prémisse importante pour la démonstration du rôle du microbiome dans la pharmacocinétique de la digoxine et la gestion clinique du médicament cardiotonique.

**Mots-clés :** Microbiome, *Eggerthella lenta*, *Eubacterium lentum*, Next-Generation Sequencing (NGS), Whole Genome Shotgun (WGS), Bioinformatique, Pharmacologie, Pharmacocinétique, Médecine personnalisée, Cardiologie, Insuffisance Cardiaque

## 2 Abstract

Digoxin is a widely used cardiotonic drug in the management of heart failure and atrial fibrillation. It has been known since the early 1980's that the main metabolite of digoxin, dihydrodigoxin, is synthesized by the gut microbiome during first pass metabolism and is exclusively produced by the bacteria *Eggerthella lenta*. In a clinical study done in the U.S.A., there were 14% of high metabolizers, for whom over 40% of the oral digoxin dose is transformed to the inactive metabolite and rapidly eliminated. Digoxin toxicity is the leading cause of hospitalization from medication's secondary effects. The toxicity events are often associated with the addition of an antibiotic (mostly from the macrolides class) to the patient's drugs regiments. The theory explored in this project could help explain the toxicity events in metabolizers. These patients have a higher daily digoxin maintenance dose to counteract the effects of the microbiome and are then prescribed antibiotics for an infection unrelated to their heart condition. The antibiotic alters *E. lenta* negatively, which cannot metabolize digoxin anymore and therefore augments the bioavailability of the cardiotonic. If the plasmatic concentration reaches dangerous levels (over 2ng/ml of plasma), the patients face adverse effects that include death. In the present project, we evaluated the susceptibility of *E. lenta* to two second generation cephalosporins, *in vivo* and *in vitro*. With the 18 healthy volunteers that were exposed to 2x500mg of cefprozil daily for 7 days, we observed a diminution of the abundance of the bacteria of interest by 58,3% from the initial levels. This change did not however produce statistically significant tests results. For the complete fecal microbiome that were cultivated *in vitro*, with or without cefprozil, the difference between the two conditions resulted in a statistically significant p-value of 0.0457, confirming the sensitivity of *E. lenta* to this cephalosporin. These results validate an important premise for the demonstration of the importance of the gut microbiome in the pharmacokinetics of digoxin and the clinical management of the drug to avoid toxicity events in clinical practice.

**Keywords** : Microbiome, *Eggerthella lenta*, *Eubacterium lentum*, Next-Generation Sequencing (NGS), Whole Genome Shotgun (WGS), Bioinformatics, Pharmacology, Pharmacokinetics, Personalised Medecine, Cardiology, Heart Failure

### 3 Table des matières

1	Résumé .....	i
2	Abstract.....	iii
3	Table des matières .....	iv
4	Liste des tableaux .....	vi
5	Liste des figures.....	vii
6	Liste des sigles.....	viii
7	Liste des abréviations .....	x
9	Remerciements .....	xi
10	Introduction .....	1
11	Revue de la littérature scientifique .....	6
11.1	Les microbiotes et les microbiomes .....	6
11.1.1	Le microbiome intestinal .....	6
11.1.2	Des bactéries jusqu'aux microbiomes.....	7
11.1.3	Les grands consortiums de recherche.....	11
11.1.4	La symbiose égoïste.....	19
11.2	Approches bio-informatique pour l'étude du microbiome.....	22
11.2.1	Types d'études métagénomiques – ADN ribosomique 16S .....	22
11.2.2	Types d'études métagénomiques – WGS .....	25
11.2.3	Bases de données publiques .....	28
11.3	Pharmacomicrobiomique.....	33
11.3.1	Le microbiome comme effecteur pharmacologique.....	33
11.3.2	Interactions avec les xénobiotiques (mécanismes directs et indirects).....	35
11.3.3	Interactions directes.....	39
11.3.4	Interactions indirectes – cas tu TMAO.....	40
11.4	La digoxine et le microbiome intestinal .....	42
11.4.1	La digoxine dans le traitement de l'insuffisance cardiaque (IC) et de la fibrillation auriculaire (FA) .....	42
11.4.2	L'utilisation contemporaine de la digoxine .....	45
11.4.3	L'interaction <i>Eggerthella lenta</i> - digoxine .....	48

11.4.4	Les études publiées depuis 35 ans .....	50
12	Objectifs.....	53
12.1	Dans le contexte de l'interaction pharmacomicrobienne proposée, confirmer que la bactérie <i>E. lenta</i> est susceptible aux antibiotiques des jeux de données (céfprozil et céfoxitine).....	53
13	Méthodologie.....	54
13.1	Source des données.....	54
13.1.1	Présentation du jeu de données – expérience avec volontaires ( <i>in vivo</i> ).....	54
13.1.2	Présentation du jeu de données – cultures <i>in vitro</i> .....	55
13.1.3	Aspect technique des données .....	56
13.1.4	Base de données de génomes bactériens .....	57
13.2	Analyses bioinformatiques d'abondance de <i>E. lenta</i> .....	58
13.2.1	<i>De novo</i> vs. closed-reference OTU picking.....	58
13.2.2	Comparaison des aligneurs D'ADN.....	60
13.2.3	Calcul des comptes normalisés.....	62
13.2.4	Comparaisons de groupes non-paramétriques et modèles statistiques.....	63
14	Résultats.....	65
14.1	Article .....	65
14.2	Résultats supplémentaires.....	82
15	Discussion.....	83
15.1	Susceptibilité de <i>E. lenta</i> au céfprozil dans l'intestin .....	83
15.2	Susceptibilité de <i>E. lenta</i> à la céfoxitine <i>in vitro</i> .....	84
15.3	Forces et faiblesses de l'étude .....	87
15.4	Perspectives futures .....	88
16	Conclusion .....	92
17	Bibliographie .....	i

## 4 Liste des tableaux

Tableau 1. Résumé des propriétés pharmacocinétiques de la digoxine chez l'humain.....	47
Tableau 2. Tableau sommaire du lieu de récupération de la digoxine administrée à des participants sains.....	49
Tableau 3. Tableau des principaux métabolites de la digoxine retrouvés dans l'urine après administration à des participants sains .....	49

## 5 Liste des figures

Figure 1. A. Échantillons intégrés au HMP, triés par provenance corporelle. B. Groupes majoritaires moyens selon les sites corporels de provenance des échantillons.....	13
Figure 2. Hétérogénéité de la composition du microbiome dans le tractus digestif murin. Tiré de (30).....	14
Figure 3. Schématisation de la technique de séquençage de Whole Genome Shotgun (WGS) et de l'assemblage des séquences (119) .....	26
Figure 4. Le rôle de Akkermansia Muciniphilia dans le maintien de l'intégrité de la barrière de mucus intestinale. Tiré de (151). .....	34
Figure 5. Principales classes d'interactions entre le microbiome et les xénobiotiques (et autres médicaments). Tiré de (163).....	37
Figure 6. Structures chimiques des cardioglycosides principaux de la digitale laineuse (digitalis lentana). <b>A.</b> structure de la digoxine <b>B.</b> Structure de la digitoxine.....	45
Figure 7. Digoxine et son métabolite principal, la dihydrodigoxine (209) .....	48
Figure 8. Niveau d'élimination de la digoxine sous forme de DRPs selon la provenance géographique des participants tiré de (12).....	51



## 6 Liste des sigles

16S ribosomal ARNr sequencing : Séquençage de la sous-unité 16S du gène ribosomal ADN :

Acide DésoxyriboNucléique

ADRB1: Beta-1 Adrenergic Receptor

ARN : Acide Ribonucléique

ASV : Amplicon Sequence Variants Concept semblable aux OTU

ATB : Antibiotique

ATP : Adénosine Tri-Phosphate

ATPase Na<sup>+</sup>/K<sup>+</sup> : Enzyme de transport d'ions sodium/potassium activée à l'ATP BAM : Binary

Alignment/Map

BWA : Burrows-Wheeler Aligner

CGR : CardioGlycoside Reductase

CRI : Centre de Recherche en Infectiologie (affilié à Université Laval)

DDi : Drug-Drug interaction

DIG : Digitalis Investigation Group

DMB : 3,3-dimethyl-1-butanol

DRP : Produits Réduits de Digoxine

ENA : European Nucleotide Archive

FAD : Flavine Adénine Dinucléotide

FMO3 : Flavin-containing monooxygenase 3

GLM : régression linéaire généralisé (Generalized Linear Model)

HMP : Human Microbiome Project

HPMC : Communautés Pan-Microbiennes Humaines

IGV : Integrative Genomics Viewer

IMG : Integrated Microbial Genomes and Microbiomes

JGI : Joint Genome Institute

KO : Knock-Out – Modèle cellulaire ou animal avec un gène supprimé du génome Log :

Logarithme – transformation mathématique

LXR : Liver X Receptors - Récepteurs des oxystérols

MetaHit : Metagenomics of the Human Intestinal Tract

MICI ou IBD : Maladies Inflammatoires Chroniques de l'Intestin

MTX : Methotrexate

NGS : Next-Generation Sequencing – Séquençage de nouvelle génération

NIH : National Institute of Health

OTU : Operational Taxonomic Unit – Concept approximativement équivalent à « espèce bactérienne »

P.O. : per os – Prise orale

pb : paires de bases (d'ADN ou d'ARN)

RDI : Raw Description and Inference plots

RNAseq : Séquençage de nouvelle génération d'ARN messagers

RPM : Reads Per Million – Métrique d'abondance relative d'amplicons bactériens 16S

SAM : Sequence Alignment/Map

TAAR5 : Trace Amine-Associated Receptor 5

THG : transfert horizontal de gènes – transmission génétique entre bactéries

TMA : Triméthylamine

TMAO : Triméthylamine N-oxyde

TMAU : triméthylaminuré

THG : transfert horizontal de gènes (ang. HGT)

VIH : Virus d'Immunodéficience Humain

WGS : Whole Genome (Shotgun) Sequencing

## **7 Liste des abréviations**

Lit. : Littéralement (dans le contexte de traductions)

## 9 Remerciements

Merci à Dre Marie-Pierre Dubé pour son support et ses contributions aux hypothèses de travail tout au long du projet. Merci à toute l'équipe du centre PGx Beaulieu-Saucier pour leur support pour les aspects plus technique du projet (et leurs qualités de collègues au quotidien).

## 10 Introduction

Avec les progrès exponentiels des technologies de séquençage de nouvelle génération (NGS en anglais), il est aujourd'hui possible de caractériser une communauté microbienne mixte et y quantifier les espèces bactériennes, fongiques, archéobactériennes, protistes et même les virus. Le principal avantage que ces technologies offrent aujourd'hui est celui de pouvoir identifier toutes ces espèces sans avoir recours à des méthodes de culture, principal outil de microbiologie classique. En effet, avant l'essor massif des technologies de séquençage et la baisse drastique des coûts de séquençage (prix par base séquencée), il fallait faire proliférer les microorganismes échantillonnés *in vitro* afin de les observer visuellement et/ou caractériser leurs capacités métaboliques. Ainsi, la microbiologie classique était fortement limitée, alors que certaines publications récentes estiment que 85% à 99% des espèces bactériennes ne peuvent pas être cultivées en laboratoire (1). En plus d'estimer la présence de toutes les bactéries présentes, même celles inconnues, les technologies NGS permettent de prendre une "photo" de ce qui est en train de se passer à un moment précis dans une communauté de microbiome.

La métagénomique est l'étude génomique des microbiomes, soit l'ensemble des microorganismes qui partagent un environnement donné. Elle s'appuie majoritairement sur 2 technologies, soit le séquençage du gène ribosomique codant pour la sous-unité 16S et la technique du Whole Genome Shotgun (WGS). La technique du séquençage 16S est utilisée depuis le début des années 1990 (2) et permet de détecter des bactéries sans les cultiver et même sans les connaître préalablement. En 2008, une équipe de recherche a séquencé des amplicons 16S qu'ils ont estimé appartenir à 215 nouvelles espèces bactériennes, dont 15 appartiennent à des nouveaux genres théoriques, dans les échantillons de microbiomes humains (3). Toutefois, la technique 16S comporte certaines limitations, par exemple le fait que certains organismes n'ont pas une région 16S homologue à celle des bactéries, rendant par exemple les virus totalement invisibles à cette technique. À l'inverse, la technique de séquençage WGS est basée sur le fractionnement aléatoire de tous les génomes échantillonnés et le séquençage de ces fragments génomiques. Au lieu d'amplifier seulement un gène comme l'ARNr, tous les éléments génomiques des bactéries sont amplifiés, puis séquencés. Le WGS permet de

reconstruire les séquences et même de reconstituer les génomes complets de nouveaux microorganismes. Nous y reviendrons dans la section revue de littérature.

Les premières campagnes de séquençage métagénomique ont démontré que la majorité des bactéries du microbiome intestinal humain sont très peu, voir totalement inconnues. Grâce à la culturomique, que certaines qualifient de renaissance de la culture en microbiologie, l'on répertorie maintenant 2 671 espèces bactériennes dont 23% ont été cultivées au moins à partir d'un échantillon humain (4). Beaucoup des techniques de culture ont été empruntées au domaine de l'écologie environnementale (5) et les techniques d'identification bactérienne à haut débit avec la spectrométrie de masse (MALDI-TOF) permettent d'identifier les nouvelles espèces rapidement (6,7). Malgré ces progrès, il reste clair qu'il y a encore beaucoup de chemin à faire pour commencer à bien comprendre l'énorme diversité des microorganismes avec qui nous cohabitons au quotidien.

Bref, le développement des nouvelles technologies de séquençage a permis l'émergence d'une nouvelle branche de la microbiomique qui est à l'intersection du monde bactérien et de la médecine personnalisée (pharmacogénomique) qui a été baptisée pharmacomicrobiomique. Cette discipline regroupe plusieurs domaines de la santé, alors que le microbiome peut transformer les médicaments ingérés (perte d'efficacité ou métabolites secondaires toxiques qui peuvent être prédits) ou encore, le microbiome est la cible des médicaments. Dans ces cas, c'est en modifiant le comportement / composition des communautés bactériennes que le médicament produit son effet et par exemple, selon plusieurs études récentes, c'est le cas avec la metformine (8–10). Dans le cas du présent projet, nous cherchons à valider un fondement important d'une théorie qui tient en compte l'importance d'une bactérie intestinale commune, *Eggerthella (E.) lenta*, dans la pharmacocinétique d'un médicament cardiotonique très largement utilisé, la digoxine.

La bactérie *E. lenta* a été identifiée comme responsable de la production du métabolite principal de la digoxine depuis le début des années 1980 (11), en réduisant le cardiotonique vers sa forme inactive, la dihydrodigoxine. Chez 14% des Américains participant à une étude, c'est 40% et plus de la dose orale de digoxine qui est transformée et marquée pour l'élimination avant même l'entrée en circulation du médicament (12) et les bactéries du groupe coriobactéries sont présentes chez 90% des adultes en santé dans le monde (13). Le présent projet cherche à pousser les hypothèses pharmacomicrobiomiques plus loin en s'interrogeant sur le rôle du microbiome intestinal dans les cas de toxicité à la digoxine observés en clinique suivant l'administration d'antibiotiques (surtout ceux de la classe des macrolides).

La digoxine est le médicament qui est responsable pour le plus d'hospitalisations chaque année pour ses effets secondaires (14) et quelques raisons expliquent ce constat, dont le fait que la digoxine a un index thérapeutique très étroit (récemment révisé à la baisse, de 0,5 ng/ml à 0,9 ng/ml de plasma (15)). De plus, les effets secondaires du cardiotonique incluent des nausées, des tremblements, des troubles de vision et des arythmies pouvant causer la mort. La littérature scientifique est d'accord sur l'importance de *E. lenta* dans la pharmacocinétique de la digoxine, surtout lors du métabolisme de premier passage. Dans le présent projet, nous cherchons à vérifier si cette bactérie peut avoir un rôle explicatif dans les cas de toxicité qui sont observés en clinique suivant l'ajout d'antibiotiques aux prescriptions des patients qui utilisent la digoxine quotidiennement pour contrôler leur insuffisance cardiaque ou leur fibrillation auriculaire.

La théorie explorée repose sur les bases théoriques suivantes: Premièrement, les patients qui sont des métabolisateurs forts de la digoxine ont une dose de maintien quotidienne ajustée à la hausse pour atteindre la concentration plasmatique désirée. Ensuite, un patient sous thérapie à la digoxine contracte une infection non reliée à sa condition cardiaque et se voit prescrire des antibiotiques. En troisième lieu, l'antibiotique affecte la survie de la bactérie métabolisatrice *E. lenta* dans le microbiome intestinal. Cette perte d'abondance de la bactérie d'intérêt mène à une baisse drastique du métabolisme de premier passage de la digoxine, dont la biodisponibilité augmente et s'accumule dans le sang jusqu'à des niveaux problématiques.

Ainsi, le présent projet vise à démontrer la sensibilité de la bactérie *Eggerthella lenta* face à deux antibiotiques de la famille des céphalosporines de seconde génération, à savoir la céfoxitine et le céfprozil. Le projet de bioinformatique présenté s'intéresse à prouver cette sensibilité dans deux jeux de données de microbiomes intestinaux exposés à la céfoxitine *in vivo*, in situ dans l'intestin de volontaires sains. Le second jeu de données a été réalisé avec des aliquots de microbiome fécal complet mis en culture *in vitro* et exposés (ou non) au céfprozil.

Le volet réalisé avec les volontaires sains avait été mis en place pour évaluer l'effet du céfprozil sur le microbiome intestinal sain et voir comment ce dernier peut récupérer après l'arrêt de la prise de l'antibiotique. Après une collecte initiale de matière fécale au jour de l'enrôlement dans l'étude, les volontaires ont pris 2x500mg de céfprozil quotidiennement pour 7 jours, puis un autre échantillon de microbiome fécal a été recueilli. Un troisième échantillon a aussi été récupéré après une période de retour à la normale de 90 jours, mais ce n'est pas l'intérêt principal du présent projet. Les antibiotiques de la famille des céphalosporines ont été sélectionnés parce qu'ils sont réputés comme faibles perturbateurs du système digestif. Ce choix a été fait pour minimiser les effets secondaires pour les participants sains de l'étude clinique.

Les données de séquençage sont de type WGS (Whole Genome Shotgun) réalisés en chimie V3 sur la plateforme Illumina et les données brutes au format .fastq contiennent des séquences de 101 paires de bases, séquencés en paires complémentaires. Pour quantifier la bactérie d'intérêt, nous avons utilisé une technique d'attribution de groupes taxonomiques opérationnels (closed-reference OTU picking), c'est-à-dire que nous avons aligné toutes les séquences contre une base de données de génomes de microbiome intestinal humain, composée de 5 141 génomes rassemblés et curés manuellement. Les lectures de PCR de qualité suffisante (91pb/101pb) et qui possèdent un partenaire de séquençage de qualité aussi attribué au même génome ont été comptabilisés et ce nombre rapporté sur le nombre total de lectures de PCR provenant du même échantillon. Avec les 24 patients échantillonnés à trois points dans le temps (total de 72 échantillons au total) et une moyenne de 15 Gb de fichiers fastq (profondeur de séquençage relativement élevée), portant le total à 1 Tb de séquences bactériennes.



Pour ce qui est des tests de croissance de microbiomes fécaux complets mis en culture *in vitro*, nous avons opté pour moins de profondeur de séquençage, ce qui a allégé les résultats de séquençage et leur traitement. Les 24 échantillons collectés au jour initial de collecte et les 24 échantillons récoltés après une semaine d'exposition (total 48 échantillons) ont été cultivés dans 4 conditions, ce qui a fait un total de 192 échantillons (48 x 4). Avec une moyenne de 2,5 Gb par échantillon séquençé, cela porte le total des données brutes à près de 0.5 Tb. Les quatre conditions de culture sont un agencement des deux paramètres suivants: culture aérobie / anaérobique et culture avec / sans céfoxitine dans le milieu de culture. Ainsi, pour chaque échantillon mis en culture (provenant soit du jour initial de collecte ou du jour 7; soit cultivé aérobie / anaérobique, il existe un échantillon cultivé avec céfoxitine et son contrôle négatif cultivé sans antibiotique. L'abondance de *Eggerthella lenta* a été calculée de la même façon que pour les échantillons provenant des volontaires sains, décrite au paragraphe précédent. Nous avons finalement seulement considéré les échantillons provenant de la première collecte (donc non préalablement exposés au cefprozil) et cultivés en condition anaérobique (parce que cela réplique mieux le milieu intestinal) pour faire les tests statistiques.

# 11 Revue de la littérature scientifique

## 11.1 Les microbiotes et les microbiomes

### 11.1.1 Le microbiome intestinal

Le terme microbiome réfère à l'ensemble des gènes de la communauté de microorganismes présente dans un environnement, par exemple l'intestin humain; le microbiome c'est l'ensemble des génomes. Le terme microbiote réfère à l'ensemble des microorganismes (archaea, bactéries, microeucaryotes et virus) qui partagent un habitat, par exemple l'intestin d'un mammifère. Fonctionnellement, l'ensemble de ces microorganismes sera souvent référé à comme 'bactéries', alors que ce dernier groupe est en nombre supérieur par deux à trois ordres de magnitude par rapport aux autres cellules de microorganismes (16). Chez l'humain, les estimés récents dénombrent environ  $3.8 \times 10^{13}$  cellules bactériennes (17) soit environ 1.3 fois plus de bactéries que les cellules humaines (loin du ratio 10:1 souvent cité). Des chercheurs blaguent même sur le fait que ce ratio peut même charger en faveur de cellules eucaryotes... après un tour à la salle de bain. Il n'en reste pas moins que le microbiome est composé de plus de 3.3 millions de gènes uniques (18), ce qui éclipse les quelques 20 000 gènes humains par un facteur d'au moins 150 fois.

Les origines exactes des termes microbiome et microbiote ne sont pas claires et sont généralement attribuées, à tort, à une publication de 2001 de Joshua Lederberg, un microbiologiste récipiendaire d'un prix Nobel (19). Le terme microbiote était déjà en usage pour désigner les microorganismes d'une manière similaire à son usage actuel, alors que le terme microbiome désignait les organismes (pas seulement les bactéries) et leur environnement ayant des limites raisonnablement claires. Il a longtemps été pensé que les bactéries qui habitent le tube digestif (microbiome intestinal) produisent des vitamines et dégradent les molécules trop complexes pour l'hôte, sans plus. Il est clair aujourd'hui que le réseau d'interactions entre le microbiome et son hôte humain est d'autant plus complexe et plusieurs considèrent maintenant le microbiome comme "l'organe oublié" (20,21). Des chercheurs s'intéressent à son implication

avec le développement du système immunitaire et les maladies auto-immunes, le diabète de type II, la santé intestinale (ex. syndrome du côlon irritable), la santé mentale (ex. la dépression), la régulation du comportement et aussi, important à ce projet, les interactions avec les médicaments.

Tous les organes qui interagissent avec l'environnement, comme la peau, les oreilles, l'utérus, le nez, les poumons et l'intestin possèdent des groupes de bactéries qui y trouvent refuge et y vivent, majoritairement d'une manière commensale. Ces environnements ne sont pas nécessairement délimités de manière claire et les protocoles d'échantillonnage exploratoires recueillent des spécimens pour déterminer, par exemple, la diversité entre des sites comme l'avant-bras et l'aisselle. Ces études démontrent que chacun des sites, même si parfois mal défini dans l'espace, a des communautés bactériennes qui lui sont propres et il est clair que l'environnement où les bactéries sont les plus nombreuses et diversifiées est celui de l'intestin.

### **11.1.2 Des bactéries jusqu'aux microbiomes**

Il y a de cela 300 ans, un fils d'artisan de paniers hollandais allait révolutionner la biologie en y greffant une nouvelle branche, la microbiologie. Armé de divers microscopes de sa propre fabrication, Antonie van Leeuwenhoek est le premier à observer les protozoaires, qu'il nomme alors 'animalcules' en 1677. La découverte de la vie unicellulaire est tellement surprenante à l'époque qu'elle laisse même ses correspondants de la Société royale de Londres sceptiques au point de douter de leur existence. Toutefois, la rigueur de la démarche scientifique du hollandais confirmera la véracité et l'importance de ses découvertes. Ses microscopes à lentille simple, plus puissants que les microscopes composés de l'époque, lui ont permis de magnifier le contenu de gouttes d'eau jusqu'à ~500x et tiennent dans la paume de la main. En fait, cette figure de magnification de 500x (22) est approximative parce qu'estimée par rapport aux observations faites par le hollandais. Du total de 322 microscopes répertoriés à l'époque (total estimé à plus de 500 microscopes), il n'en reste plus qu'une dizaine réputés authentiques et le plus puissant produit une magnification de 200x (23). Parmi plusieurs groupes de microorganismes que Leeuwenhoek répertorie le premier, il observe aussi les bactéries, malgré

leur taille minimale. Il faudra attendre quelques 150 ans avant que les microscopes composés développés par un professeur de l'université Tulane en Louisiane ne permettent de faire des meilleures observations du monde microscopique.

Faisant preuve d'un esprit scientifique remarquable, Leeuwenhoek va observer toute sortes de structures naturelles avec une grande rigueur, comme les stries des muscles, les spermatozoïdes, les vacuoles des cellules, etc. Il observera même ses propres microbiomes oraux et fécaux et son intuition scientifique l'amènera même à comparer les microbiomes fécaux de personnes en santé avec ceux de patients atteints de maladies. Toutefois, avec la technologie limitée de l'époque, il est difficile de tirer des conclusions élaborées et reproductibles sur ces unicellulaires fraîchement découverts. Certains des contemporains de Leeuwenhoek spéculent que ces 'animalcules' pourraient causer des maladies, mais le Hollandais restera sceptique. Plusieurs médecines traditionnelles s'intéressent d'une manière approfondie à l'importance de la digestion dans la santé humaine et le microbiome fécal est très certainement le proxy le plus accessible pour l'observation de ce qui se passe dans le tube digestif. Bref, avec le savoir moderne qui est en train de se développer autour des unicellulaires omniprésents dans notre environnement et nos corps, il apparaît clair que Leeuwenhoek a eu des intuitions scientifiques des plus pertinentes.

Les microscopes développés à Tulane par le professeur chimiste J. Lawrence Smith vers 1850 (24) ont permis le développement de la bactériologie moderne, une sous-branche de la microbiologie, en popularisant des microscopes composés beaucoup plus simples d'utilisation que ceux de Leeuwenhoek. La popularisation de la microscopie va permettre à la communauté scientifique de mettre à la poubelle des théories mal fondées comme celle de la "génération spontanée". À cet effet, les travaux de Louis Pasteur en France permettent de confirmer la "théorie des germes" et de mettre en lumière l'importance de la vie microscopique dans le développement de maladies comme la rage. On lui crédite d'ailleurs le premier vaccin prophylactique pour prévenir ce virus. Vers le milieu des années 1880, son contemporain

germanique Robert Koch pousse encore plus loin l'établissement des microorganismes dans l'étiologie des maladies humaines avec ses 4 postulats désormais célèbres. Ils vont ainsi:

1. Le microorganisme doit être présent en abondance dans tous les organismes souffrant de la maladie, mais absent des organismes sains.
2. Ce microorganisme doit pouvoir être isolé et croître en milieu de culture pure.
3. Le microorganisme cultivé doit déclencher la même maladie chez un animal de laboratoire sensible.
4. Le microorganisme doit être à nouveau isolé du nouvel organisme hôte rendu malade puis identifié comme étant identique à l'agent infectieux original.

Ainsi, Koch isole les unicellulaires suspects dans les prélèvements des patients atteints de maladies, les observe au microscope, les isole, les cultive, puis les inocule à des animaux pour confirmer leur pouvoir pathogène. Cette approche méticuleuse lui a permis d'identifier hors de tout doute les bactéries causales des pathologies de l'anthrax, de la tuberculose et du choléra. Ses travaux sur la tuberculose lui valent d'ailleurs un prix nobel en 1905. Toutefois, pour chacune des étapes de cette méthode, il existe des agents biologiques pathogènes connus aujourd'hui qui vont éluder à la détection. Par exemple par faute d'être cultivable dans des conditions standard, comme c'est le cas avec tous les virus et la majorité des bactéries, beaucoup d'agents pathogènes pourraient passer inaperçus parce qu'ils ne se multiplieront pas *in vitro*.

Les progrès accomplis en pathologie humaine grâce aux travaux de Koch sont indéniables et leur rigueur a permis d'identifier les agents causaux de nombreuses autres maladies. Toutefois, la recherche actuelle sur le microbiome démontre que le portrait n'est pas aussi simple. En effet, un des grands fléaux de la médecine moderne ne correspond pas du tout au premier postulat de Koch, et la pathologie en question est celle de l'infection par *Clostridium Difficile*. Ce bacille est résistant aux antibiotiques et il cause des diarrhées potentiellement mortelles, surtout en tant qu'infection nosocomiale. La raison de sa non-conformité avec le premier postulat se trouve dans la nature quasi-ubiquitaire de *C. Difficile*, alors que la majorité

des humains sains sont déjà entrés en contact et seront encore en contact avec la bactérie au cours de leur vie et que dans bien des cas, ils seront colonisés à long terme. Une étude récente a étudié le rôle des enfants comme réservoirs potentiels de *C. Difficile* et la prévalence de la bactérie est de 45%, avec 13% de porteurs de la souche toxigénique (25) avant même de quitter l'hôpital à la naissance. Ainsi la présence de *C. Difficile* n'est pas synonyme d'une infection par *C. Difficile*. De plus, lorsque l'on consulte les archives de génomes bactériens complètement séquencés sur des archives publiques comme "Nucleotide" du NCBI, une très forte proportion des noms latins attribués aux espèces bactériennes réfèrent à la maladie qu'elles peuvent causer, à l'épidémie durant laquelle elle a été isolé ou le nom du chercheur qui a risqué sa vie en allant recueillir des échantillons de l'agent causal de l'épidémie.

Devant cette réalité, plusieurs chercheurs questionnent l'intérêt de concentrer la recherche uniquement sur le potentiel pathogène des unicellulaires et de plutôt s'intéresser à l'ensemble des microorganismes en présence. Trop souvent, la science a une attitude réductionniste qui pousse à classifier les microorganismes en pathogènes ou inintéressants. L'attitude est tout à fait justifiable considérant la panoplie de facteurs qui peuvent influencer l'aboutissement des recherches avec du matériel biologique. Toutefois, il apparaît aujourd'hui de plus en plus évident que nous ne pourrons pas comprendre entièrement les bactéries si l'on étudie que celles qui sont potentiellement pathogènes et c'est pour cette raison que beaucoup des fondements de la recherche sur le microbiome emprunte des termes, des concepts et des techniques qui ont d'abord été développés dans le cadre de travaux en écologie des communautés. Un autre concept qui fait émergence actuellement est celui de la dysbiose, une condition de déséquilibre des groupes de microorganismes en présence, pour expliquer le développement de certaines pathologies. Par exemple des études récentes sur maladies inflammatoires chroniques intestinales (MICI ou IBD en anglais) parlent du déséquilibre des groupes microbiens en présence pour expliquer le développement de la maladie, plutôt que de tenter d'isoler un seul taxon pathogène (26,27).

Heureusement, une multitude de progrès technologiques permettent aujourd'hui d'étudier les microorganismes sans avoir à les observer visuellement et sans avoir à les cultiver *in vitro*, les deux premières conditions des postulats de Koch. Par exemple, dès 1975 Anhalt et Fenselau ont publié une technique indépendante de culture pour l'identification de bactéries, reconnaissant les forts biais et les embûches techniques qui accompagnent inévitablement les techniques de culture *in vitro*. La technique nécessite de lyophiliser les bactéries et de les analyser avec la spectrométrie de masse, sans avoir à les cultiver *in vitro* ou à amplifier leur matériel génétique. Leur technique n'a pas connu la popularité qu'on aurait pu prévoir, probablement parce qu'elle nécessitait d'isoler une grosse quantité de l'échantillon à identifier et qu'elle ne permet pas d'identifier des bactéries préalablement inconnues. Heureusement, les progrès des techniques de biologie moléculaire, notamment le clonage par PCR (réaction de polymérase en chaîne), permettent depuis 20 ans de décoder directement le contenu des génomes des bactéries (28,29). L'avènement des techniques de séquençage à haut débit modernes permettent maintenant de déterminer virtuellement toutes les séquences de toutes les bactéries présentes dans un échantillon, toujours sans avoir à cultiver les microorganismes présents et évitant ainsi de biaiser les données obtenues.

### **11.1.3 Les grands consortiums de recherche**

Le terme microbiome réfère à l'ensemble des gènes de la communauté de microorganismes présente dans un environnement donné, par exemple l'intestin humain; le microbiome c'est l'ensemble des génomes. Le terme microbiote réfère aux microorganismes (archaea, bactéries, microeucaryotes et virus) qui partagent un habitat, par exemple l'intestin d'un mammifère. Fonctionnellement, ces microorganismes seront référés à comme 'bactéries', alors que ce dernier groupe est en nombre supérieur par deux à trois ordres de magnitude aux autres cellules de microorganismes (16). Chez l'humain, les estimés récents dénombrent environ  $3.8 \times 10^{13}$  cellules bactériennes (17) soit environ 1.3 fois plus de bactéries que les cellules humaines (loin du ratio 10:1 souvent cité). Toutefois, il vaut de mentionner que le microbiome est composé de plus de 3.3 millions de gènes uniques (18), ce qui éclipse les environ 20 000 gènes humains par un facteur d'au moins 100 fois.

En 2008 aux États-Unis, le Projet microbiote humain (Ou HMP pour Human Microbiome Project en anglais) a été lancé par le NIH (National Institute of Health) avec un financement de 115 millions \$ (US) et le projet a pris fin en 2012. Ayant comme premier objectif de séquencer et caractériser les génomes des microbiotes retrouvés dans les différents microbiomes d'humains sains, le projet a aussi fait voir le jour à plus de 190 publications dont certaines sur les microbiomes liés à des états pathologiques. Les participants de l'étude ont été prélevés à différents sites, comme la bouche, la plante des pieds, le postérieur, les aisselles, le vagin, l'intérieur du tractus intestinal et, comme illustré sur la Figure 1-a, beaucoup d'échantillons de selles. Les données générées par le consortium HMP sont disponibles au public, à leur état brut et dans diverses représentations graphiques et autres traitements bioinformatiques, comme l'assignation taxonomique. Ils ont ainsi généré plusieurs figures, comme la figure Figure 1-b, qui permettent de visualiser efficacement ce qui constitue un microbiome normal, par site de prélèvement. Leur interface web interactive nouvellement redessinée (<https://portal.hmpdacc.org/>) offre un accès attrayant vers une panoplie de types de données, comme les fastq bruts de plus de 6 363 échantillons de matière fécale, 1 224 provenant de la fosse nasale, 4 406 provenant de la muqueuse buccale, etc. pour un total de plus de 30 000 échantillons disponibles publiquement pour téléchargement. Ils ont aussi inclus une liste de 2 200 génomes bactériens de référence et plus d'une vingtaine d'outils informatiques ainsi que les explications détaillées pour permettre d'analyser les données dans le cadre d'études métagénomiques.



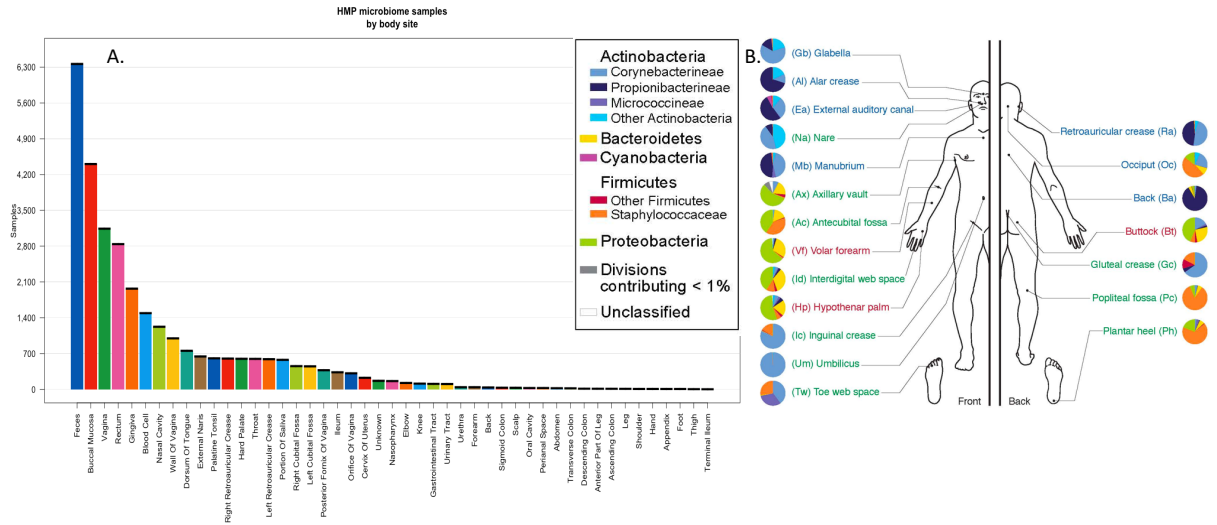


Figure 1. A. Échantillons intégrés au HMP, triés par provenance corporelle. B. Groupes majoritaires moyens selon les sites corporels de provenance des échantillons.

Ces données permettent de faire des observations intéressantes, par exemple que les mêmes espèces peuvent se retrouver à différents sites sur le corps, mais leurs proportions respectives qui varient dramatiquement. L'humain offre une grande diversité de niches écologiques, comparativement à l'environnement riche en nutriments et chaud de l'intestin, la peau de l'avant-bras ressemble à un désert pauvre en eau et en nutriments et exposé aux radiations UV. Aussi, les bactéries du genre *Lactobacillus* (largement utilisées en fermentation des aliments, la transformation des produits du lait et pour leur caractère dit probiotique) se retrouvent peu dans les selles, alors qu'incapable de s'établir dans le tractus digestif, mais est la population très largement majoritaire de la flore vaginale saine. Pour caractériser ces différences, une étude de 2016 s'est intéressé à l'hétérogénéité des groupes bactériens au long du tube digestif de souris de laboratoire et les groupes dominants changent de manière significative entre les différentes parties qui composent le tractus digestif, de la bouche au côlon (30). Deux choses ressortent comme très intéressantes par rapport au petit intestin, la première est que c'est le seul site de microbiome de mammifère où le groupe microbien dominant est celui des *mycoplasmataceae*, un groupe de bactéries (contrairement à ce que le suffixe du nom pourrait suggérer). De plus, ce groupe redevient un groupe très minoritaire du microbiome à partir de la transition vers le gros intestin et reste une portion presque négligeable dans les échantillons de selles des souris.

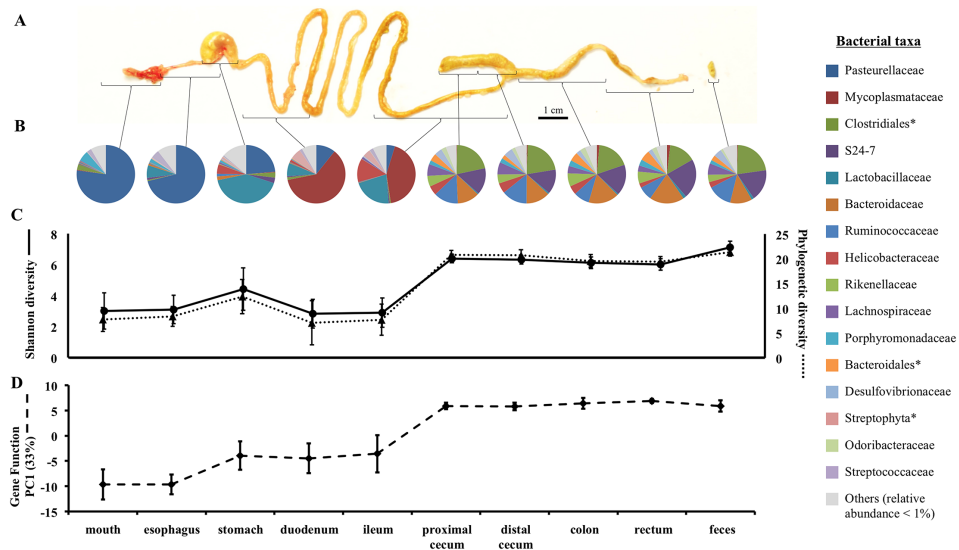


Figure 2. Hétérogénéité de la composition du microbiome dans le tractus digestif murin. Tiré de (30)

L'étude de l'hétérogénéité de la composition des microbiomes lors du transit intestinal signifie aussi que les échantillons de selles ne sont peut-être pas le meilleur proxy pour l'étude des phénomènes impliquant le microbiome intestinal, même si cette étude est basée sur des souris. Heureusement pour l'avancement de la science du microbiome, des méthodes de prélèvements beaucoup moins invasives que les cathéters ont été développées, comme les caméras sans fil avalables (31–33) et de ces capsules, un nouveau genre de capsules d'échantillonnage permet de recueillir le contenu de l'intestin au lieu désiré (34).

Avec des dates concomitantes et un budget similaire, la Commission Européenne et son septième programme-cadre ont permis de créer le MetaHIT, (pour Metagenomic of the Human Intestinal Tract) avec une enveloppe budgétaire de 11,4 millions d'euros. Le projet a été lancé officiellement en 2008 et a pris fin en 2012, à l'instar de son équivalent américain et comme son nom l'indique, le projet MetaHIT concentre ses recherches sur un seul microbiome, celui du tube digestif humain. Bien que plusieurs publications se recoupent avec celles du HMP, la mission que le consortium s'est donnée est aussi différente, avec comme objectif central d'établir des liens entre les gènes du microbiome, en état de santé et en état de pathologie. Cet objectif a comme visée principale de caractériser le rôle du microbiome dans l'étiologie de l'IBD et de

l'obésité, deux désordres en forte croissance en Europe. Toutefois, contrairement au HMP qui a lancé une nouvelle plateforme web à l'automne 2017, le fil de 'nouvelles' du MetaHIT est peuplé d'actualités datant de 2011 et 2012 (site non entretenu depuis).

En plus d'une mission articulée autrement à celle de son homologue américain, le consortium MetaHIT a aussi pris une approche différente à l'analyse des données générées. En effet, alors que le HMP s'est mandaté de catalogue les taxons microbiens présents (au niveau de l'espèce et au niveau des phylums), le groupe européen s'est surtout intéressé à l'aspect génétique des séquences retrouvées. Avec les méthodes de métagénomique en criblage aléatoire, il est possible de reconstituer virtuellement des longs fragments d'ADN et ensuite de soit: 1. les classer en groupes par similitude de séquence et de les attribuer à des espèces et les taxons connus; ou encore 2. s'intéresser à leur contenu génétique, c'est-à-dire de prédire les gènes selon les séquences fonctionnelles pour les attribuer à des groupes de fonctions métaboliques connues. Les deux approches ne sont pas mutuellement exclusives pour une étude métagénomique contemporaine et chaque approche offre un point de vue très révélateur sur les données, mais sont toutes deux très exigeantes du point de vue du traitement de données. Une chose qui ressort en commun entre plusieurs études et cela peu importe la méthode de traitement des données, soit que dans toutes les études comparant un microbiome sain à celui dans un état de dysbiose, la diversité est négativement affectée que ce soit au niveau des gènes ou au niveau des espèces (35–39). La majorité des études récentes évaluent la diversité du microbiome en termes d'indice alpha ou d'autre index de diversité écologique évaluant la diversité des espèces ou des familles de microorganismes présents.

Il existe plusieurs définitions du terme 'espèce' et celle la plus souvent citée est probablement celle de : "des animaux qui peuvent s'accoupler ensemble et produire une descendance viable et fertile proviennent de la même espèce". Toutefois, il est facile de voir qu'avec les microorganismes, cette définition est très certainement inappropriée, voire même inutile. En effet, non seulement les bactéries se reproduisent par clonage en se dupliquant en cellules filles identiques, les bactéries peuvent échanger sections de matériel génétique dans un

processus appelé le transfert horizontal de gènes (THG). L'exemple classique de THG est celui où les bactéries s'échangent des gènes de résistance aux antibiotiques avec une structure externe appelée pilus sexuel, en référence au fait que c'est la façon qu'ont les bactéries de trouver des partenaires pour faire évoluer leur génome. En fait, les bactéries ont même tendance à être réceptives à l'ADN étranger, une capacité appelée 'compétence' en clonage moléculaire. En fait même plus encore, les bactéries ont tendance à intégrer toutes sortes d'ADN libre qu'elles rencontrent dans l'environnement (40,41), à la recherche de gènes qui pourront augmenter leur succès évolutif et au risque d'incorporer des phages qui peuvent les parasiter. Et les bactéries ne sont pas les seules à accepter des séquences codantes sans nécessairement connaître l'aboutissement d'une telle intégration, comme nous verrons dans le prochain chapitre, plusieurs études font ressortir la possibilité que les microorganismes aient eu un rôle direct dans l'évolution des génomes des animaux et autres pluricellulaires.

Malgré cette impossibilité de définir sans équivoque ce qui constitue une espèce bactérienne, plusieurs souches et espèces bactériennes sont répertoriées et ont des noms connus, comme la désormais célèbre bactérie intestinale *Escherichia Coli*, très utilisée en recherche. En effet, il n'existe pas un seul génome officiel pour *E. Coli* et alors que la majorité des souches sont bénignes (une ou plusieurs souches présentes chez 19 de 20 volontaires sains (42)), certaines comme la *E. Coli* O157:H7 causent des fortes diarrhées à cause de sa toxine semblable à la shigatoxine et peut même engendrer un syndrome hémolytique et urémique permanent.

Bref, comme il n'existe pas un répertoire complet de toutes les espèces microbiennes et leurs déclinaisons, une technique statistique de regroupement a été développée pour faire de la taxonomie numérique en écologie environnementale. Concept proposé la première fois en 2005, l'OTU (ou Unité Taxonomique Opérationnelle en français) est un type de regroupement qui remplace fonctionnellement la notion d'espèce (43) et il s'agit généralement du métrique communément utilisé pour quantifier la diversité des organismes présents dans un écosystème. En écologie des macroorganismes classique, la quantité d'espèces animales présentes et la répartition des individus à l'intérieur des groupes (par exemple, dans un écosystème où une seule

espèce représente 90% des individus rencontrés, la diversité sera plus basse) permettent de calculer un indice de diversité numérique comme l'indice de Shannon ou l'index de Simpson, largement utilisés en écologie et dans l'étude des microbiomes. Pour ce qui est des OTU, la définition a quelque peu changé depuis la formulation initiale par Blaxter *et al.* en 2005 et il existe aujourd'hui plusieurs programmes d'analyse avec des algorithmes, comme uclust (44), cd-hit (45) et WGSQuikr (46). Le chercheur détermine un seuil de similarité de séquences génétiques, généralement 97% ou 95%, qui sera le seuil minimal pour regrouper des séquences en un groupe appelé OTU.

En 2011 dans une publication dans la revue Nature a fait beaucoup réagir, le renommé scientifique du microbiome Peer Bork a proposé une façon de réduire la complexité des données générées dans les études sur le microbiome en classifiant les participants aux études selon leur groupe bactérien majoritaire. Cette proposition est très certainement attirante, alors qu'il est difficile d'interpréter les résultats de séquençage de microbiome, minimalement pour la raison que chacun a un microbiome unique. L'interprétation des données de séquençage de génomes complet est tellement complexe qu'il est courant de voir les comparaisons des patients échantillonnés comparés par des graphiques 'd'analyse de composantes principales'. Ce type de représentation statistique est utilisé avec les données de hautes dimensions pour les simplifier, mais ces graphiques sont presque impossibles à interpréter de manière concrète, alors que les axes ne représentent pas des valeurs numériques matérielles simples. Bref, le fait d'attribuer un échantillon à un groupe permet par exemple de stratifier les analyses et augmente ainsi les chances de générer des résultats interprétables, malgré la complexité des données.

Le concept des trois entérotypes a séduit la communauté scientifique et même fait jaser les médias conventionnels. Les trois entérotypes proposés initialement sont caractérisés par le phylum dominant : *bacteroides*, *prevotella* et *ruminococcus*, selon les bactéries qui domine le microbiome fécal échantillonné. L'entérotipe bactéroïdes est dominant chez les occidentaux (47) et est associé à la consommation de gras saturés et de protéines animales, alors que l'entérotipe *prevotella* est prévalent ailleurs. Ce dernier est associé à des diètes riches en

carbohydrates et en sucres, alors qu'il est aussi plus observé chez les végétariens (48). Pour ce qui est de l'entérotype ruminococcus, il est plus rare et, par exemple, l'étude de Gorvitovskaia *et al.* parue en 2016 l'a même regroupé avec l'entérotype bactéroïde, en argumentant que la distinction entre les deux groupes n'est pas si clairement définie (47).

Depuis la proposition du concept en 2011, le terme entérotype s'est fait une place dans plusieurs publications scientifiques, mais ne fait pas l'unanimité. Les critiques s'articulent principalement autour de la robustesse des regroupements des patients en groupes, alors que plusieurs chercheurs estiment que la diversité des microbiomes intestinaux observés est un gradient des groupes principaux et des groupes associés (49,50). De plus, certains suggèrent que, par différents mécanismes associés au TGH, l'habilité des gènes à se propager dans l'environnement à partir de multiples sources vers les bactéries communes du tractus digestif compliquent et affaiblissent le paradigme entérotype-fonctions métaboliques (51).

Malgré le débat autour de l'utilité et même autour de l'existence des entérotypes, plusieurs travaux aussi récents que 2018 l'utilisent dans l'analyse de leurs données, que ce soit en écologie (52–54) ou en santé humaine et animale (55–61). Considérant la controverse autour du concept d'entérotypes, certains chercheurs discutent aussi du concept d'entérogradients et beaucoup se servent des ratios entre les deux groupes qui dominent généralement le microbiome fécal, soit le ratio de firmicutes sur les bactéroïdètes. Ainsi, le ratio F/B a été corrélé avec l'indice de masse corporelle chez les participants (62,63), ou avec l'activation immunitaire chez les souris (64) et que d'autres ont suivi l'évolution des entérotypes inférés grâce à un calcul dérivé du ratio F/B (65,66). Il vaut de noter que l'analyse des données de séquençage est tellement complexe qu'il est très courant de voir des publications où les patients sont comparés sur des représentations graphiques de type "analyse de composante principale", une représentation où les axes ne sont pas des variables quantifiables numériquement et sont difficile à interpréter. Certaines études s'intéressent à l'établissement d'un microbiome central ('core microbiome' en anglais) et de distinguer de ce qui est une composante stable du microbiome versus une composante transitoire du microbiome (67–69).

#### 11.1.4 La symbiose égoïste

Il n'existe pas de eucaryotes sans procaryotes. Que ce soit le plus simple des pluricellulaires, comme des éponges et les cnidaires, tous les animaux ont un microbiome (70–72) et même les colonies d'algues proto-pluricellulaires comme *volvox* ont des endosymbiontes bactériens (73) (i.e. microbiomes), une réalité que nos yeux macroscopiques ont tendance à nous faire oublier alors qu'aucun œil nu n'a la capacité de nous faire plonger seul dans ce monde de mystères. L'association est tellement vieille et inévitable que des hydres (cnidaires souvent appelées anémones; animaux diploblastiques très simples qui sont la phase de vie sessile des méduses) produisent leurs propres neuropeptides pour contrôler le développement des espèces bactériennes qui l'assistent dans la digestion de ses aliments (74). Les premiers ancêtres des protozoaires retrouvés dans des fossiles ont 3,47 milliards d'années (75) et auraient pu émerger aussi tôt que quelques centaines de millions d'années après la solidification du magma de la terre. Les eucaryotes attendront un autre 2 milliards d'années pour commencer à émerger et se diversifier et les eumétazoaires ('animaux vrais', spécialisation cellulaire complète comme chez les vertébrés) (76,77). Bref, malgré le fait que notre point de vue de conscience d'animal complexe peut nous amener à concevoir que nous vivons sur la terre des animaux, la biomasse des plantes et des animaux combinés est estimée comme équivalente à celle des procaryotes (78). Cela veut dire que, collectivement les procaryotes unicellulaires sont aussi massifs que tous les troupeaux d'éléphants, d'humains et de bovins combinés à la masse de tous les arbres et plantes des forêts boréales et tropicales. Bref, un argumentaire pourrait être entretenu que nous vivons encore sur la terre des procaryotes et cela même sans considérer tous les descendants encore relativement indépendants qui habitent les cellules eucaryotes, à savoir les mitochondries et les plastides photosynthétiques absolument essentiels à toute la vie pluricellulaire sur terre. Il n'existe pas d'estimation de la biomasse des plastides et mitochondries, mais il va sans dire que la formule « unicellulaire » a prouvé son efficacité, autant par son ubiquité que son nombre.

À la lumière de ces observations, certains auteurs ont proposé de percevoir les animaux complexes comme étant composés des holobiontes dont le génome collectif de l'eucaryote hôte et de ses communautés microbiennes constitue l'hologénome. Ce concept est rapidement devenu populaire dans les domaines proches de l'écologie et de la lutte biologique (79–83), alors que le concept d'holobionte proposé inclut les cellules de l'hôte métazoaire (animal), les bactéries et autres unicellulaires ainsi que les virus. La première proposition du concept en 2008 par Zilber-Rosenberg et Rosenberg s'appuie sur l'ubiquité du partenariat hôte-microorganisme, de la transmissibilité du partenariat et de l'influence du partenariat sur la survie des parties impliqués (84) et considère les microbes et l'hôte comme une seule unité intégrée.

Malgré la réalité des interactions décrites par le concept d'holobionte, ce dernier reçoit beaucoup de critiques. En effet, alors que certains défenseurs du concept veulent considérer les gènes de l'hôte et les gènes du microbiome sur un pied d'égalité, d'autres chercheurs soulignent les limitations d'une telle approche. Par exemple, des espèces proches pourraient avoir des microbiomes semblables sans les avoir hérités, mais plutôt en les ayant acquis de sources environnementales lors d'évènements non reliés. Cela va de pair avec l'absence de fidélité dans les relations symbiotiques hôte-microbiome, alors que chacun des parties impliqués interagissent avec l'autre seulement pour sa propre survie. De plus, il est très possible que les réactions de l'hôte soient régies par les mêmes mécanismes lorsqu'en présence de microorganismes qu'en présence de facteurs abiotiques (85–87). Bien que les membres des microbiotes rendent des services indéniables et souvent essentiels à leurs hôtes respectifs, la relation n'est pas basée sur l'altruisme ou la volonté de participer à un tout plus grand, mais bien comme le font toutes les formes de vie sur terre, les microbiotes utilisent les ressources d'une niche écologique pour leur survie.

Sans se prononcer davantage de la réalité de l'holobionte comme unité de base de l'évolution, il reste clair que les hôtes eucaryotes interagissent avec les passagers de leur microbiomes et que ces interactions peuvent influencer le succès évolutif de tous les parties impliqués. Les traces de ces interactions complexes sont observables encore aujourd'hui dans le



code génétique humain. Par exemple, l'arrivée de bactéries pathogènes a récemment laissé des traces dans l'héritage génétique des européens. En effet, récemment plusieurs recherches sont faites sur la mutation CCR-5 delta32 parce que cette dernière protège les individus porteurs homozygotes contre l'infection au VIH (virus d'immunodéficience humaine). Chez les caucasiens originaires d'Europe, la mutation hétérozygotes est retrouvée chez 10% de la population, avec 1% des européens homozygotes au site de variation. Comparativement, la mutation apparue *de novo* il y a 700 ans environ est virtuellement absente chez les populations hors-Europe (88). Le fait que la mutation soit aussi répandue aujourd'hui est le fruit d'une pression de sélection sur les organismes hôtes (dans ce cas-ci les habitants de l'Europe autour des années 1300) qui a sélectionné pour les individus moins susceptibles à un pathogène, fort probablement le bacille de la peste noire, *Yersinia pestis*, qui a ravagé l'Europe autour de cette époque (épidémie majeure en 1347) (89,90).

À l'inverse, les organismes supérieurs travaillent aussi à moduler leurs partenaires microbiens. Par exemple, le lait maternel humain contient des oligosaccharides qui ne peuvent pas être dégradés en constituants simples par le système digestif du nourrisson. La présence de ces sucres complexes a longtemps laissé les chercheurs perplexes, alors que le corps de la mère doit les synthétiser à grand frais énergétiques et que cette énergie ne peut pas être récupérée par la progéniture. La raison de la présence de ces sucres complexes serait en fait reliée au microbiome, alors que les bactéries du type Lactobacilles et Bifidobactéries, deux probiotiques largement étudiées et reconnues pour leur effet bénéfique sur la santé et le système digestif. La présence et composition normale du cocktail d'oligosaccharides du lait maternel est nécessaire au bon établissement du microbiome dans les premiers moments de la vie en promouvant les bactéries mutualistes et en prévenant la croissance des pathogènes (91–94). C'est donc un exemple où l'organisme hôte s'est adapté au microbiome en intégrant des voies métaboliques pour entretenir un partenariat avec des groupes particuliers de microorganismes.

Il est fréquent de voir des conférences scientifiques ou des documents écrits sur le système immunitaire où les organes lymphatiques sont tous énumérés à l'exception des plaques

de Peyer. Ce qui est d'autant plus étonnant considérant le fait que c'est l'organe lymphoïde qui contient le plus de cellules lymphatiques (95), étant un centre où les cellules immunitaires se reproduisent en continu. Cela témoigne de l'importance de pouvoir réguler la prolifération bactérienne, sans quoi l'hôte est hautement vulnérable aux bactéries opportunistes et pathogènes. De plus, le système immunitaire est 'éduqué' par les bactéries du microbiome lors du développement, pour que les bactéries commensales soient tolérées, mais que les pathogènes soient surveillées de près (96–98). Cette éducation du système immunitaire est aussi importante dans l'étiologie des maladies auto-immunes (99,100).

## **11.2 Approches bio-informatique pour l'étude du microbiome**

### **11.2.1 Types d'études métagénomiques – ADN ribosomique 16S**

Le terme métagénomique inclut toutes les études de microbiome qui ciblent l'ensemble des microbiotes présente dans l'environnement étudié, que ce soit dans un échantillon de sol, un échantillon d'eau de lac, un frottis de peau, ou dans le cas présent un échantillon de matière fécale. Le terme fait référence au fait que l'on ne cherche pas à connaître un seul génome, mais un ensemble de génomes, soit ceux de toutes les bactéries, archaebactéries, champignons, protozoaire, et souvent aussi les virus. En fait les virus font toujours partie du microbiome, mais ils ne sont pas détectables par toutes les méthodes comme nous le verrons.

La méthode la plus répandue et qui a permis la première de générer des jeux de données de métagnomique est celle dite 16S, en référence à l'approche de clonage ciblé visant à amplifier le gène codant pour la sous-unité 16S du ribosome, un ARN fonctionnel. Cette région est ciblée parce qu'elle encode une composante essentielle à la synthèse protéique, donc essentielle à toute vie cellulaire et est par conséquent hautement conservée à travers différents groupes d'unicellulaires. Ainsi, avec des amorces d'ADN une PCR (Polymerase Chain Reaction, la méthode à la base des séquençages, servant à amplifier le matériel génétique) est réalisée pour

généraliser des amplicons qui seront séquencés par la méthode Sanger ou d'autres technologies plus récentes de séquençage. Il existe différents critères de similarité pour classer les espèces ensemble et souvent on considère qu'une homologie de séquence de 97% signifie que deux séquences appartiennent à la même espèce, ou plus précisément au même OTU (décrits plus tôt). Le nombre de copies attribué par les algorithmes informatiques à chaque espèce permet de faire une quantification relative de chaque groupe taxonomique par rapport au total, que ce soit au niveau de l'espèce, du genre ou de la famille, par rapport au total des espèces présentes. Une analyse standard consiste à produire le rapport des OTU majoritaires (top OTUs table), qui informe sur les espèces majoritaires et permet de calculer facilement la diversité alpha ou d'autres indices de diversité écologique.

Comme le séquençage 16S a joué un rôle prépondérant dans l'étude métagénomique et qu'il continue à le faire, il existe une panoplie de logiciels et de pipelines d'analyses pour traiter ces données. Beaucoup de ces logiciels sont sous licences libres et même disponibles sur des serveurs faciles d'utilisation comme MG-RAST (101), qui offre même la possibilité de faire le téléchargement et le traitement des données sur leurs serveurs. D'autres outils permettent d'inférer diverses propriétés du microbiome à partir des données de 16S, comme par exemple le logiciel PICRUST (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) qui permet d'établir un profil fonctionnel prédictif des communautés microbiennes. L'algorithme d'inférence des fonctions métaboliques a été développé grâce aux données du HMP, le grand consortium de recherche cité plus tôt. Le logiciel a été développé par Morgan Langille et son équipe à l'université canadienne de Dalhousie et fait partie d'une suite de logiciels globalement appelés 'BioBakery'. Développés par un réseau de collaborateurs, la suite de logiciel comprend le très populaire MetaPhlan (utilisé pour l'assignation taxonomique des WGS, terme défini dans la prochaine section) et l'anecdote veut que ce soit la conjointe d'un des chercheurs qui soit à l'origine du thème des pâtisseries (gâteaux, tartes et autres) pour les noms donnés aux différents outils informatiques regroupés dans le projet BioBakery (ex. ShortBRED, PPAANI, WAAFL, CCPE, MicroPITA, etc.).

La technique de séquençage 16S est tellement robuste et simple d'exécution qu'elle est même utilisée pour faire de la phylogénie et de l'identification d'animaux supérieurs. En effet, des chercheurs amplifient des parties de gènes mitochondriaux et, selon la très largement acceptée théorie de l'endosymbiose (102–105), les mitochondries dérivent de bactéries du type *rickettsiaceae* qui encore aujourd'hui ont un génome indépendant de celui de la cellule eucaryote qui l'héberge. Bien qu'incomplet, ce génome contient encore ses propres gènes ribosomaux et ces ARNr sont ceux ciblés, surtout dans les régions 12S et 16S, pour faire l'identification et la phylogénie de vertébrés et d'insectes (106–110).

Le choix des amorces peut aussi affecter le résultat du séquençage des gènes ribosomiques, alors qu'il a longtemps été nécessaire de considérer quelle section du gène l'on souhaitait amplifier. En effet, à l'intérieur du gène il y a à l'intérieur du gène ribosomal des régions génétiquement stables, entrecoupées de régions hypervariables. Ces dernières ont été étiquetées de V1 à V9 et plusieurs publications, dont certaines récentes, parlent de l'importance du choix des amorces (111) pour optimiser les régions variables à séquencer (ex. cibler la région V4-V6) pour choisir des amorces dites universelles (112) et optimiser la sensibilité de l'analyse. Le gène ARNr 16S est encodé dans une région d'un peu plus de 1 500 nucléotides et contrairement à ce que suggère la publication de 2016 tout juste citée, d'autres chercheurs présentent des amorces (Bac8f et UN1541r) qui amplifient l'entièreté de la région (113). Mais la majorité des technologies de séquençage de seconde génération ne produisent que des lectures de 200 pb +/- 50 et le séquençage de 1500 pb reste un défi de taille. Ces considérations techniques relatives à l'utilisation de la PCR et du traitement des données sont donc à la discrétion des expérimentateurs. Toutefois, il existe quelques limitations intrinsèques à la technique du séquençage des ARN ribosomaux 16S. Dans un premier temps, comme les virus détournent la machinerie de synthèse protéique des cellules qu'ils envahissent, ils ne possèdent pas d'ARNr et passent donc sous les radars de la détection par les méthodes d'amplicons 16S. Il existe des sources de biais pour lesquels il existe des outils informatiques pour les contrer, comme le fait que le nombre de copies du gène 16S varie selon les espèces, alors la corrélation nombre de copies 16S comparé au nombre de génomes de l'espèce n'est pas 1:1 et il est important d'ajuster les calculs d'abondance relative en conséquence (114–117). L'outil CopyRighter a été

développé à cet usage (117). Par contre, une autre source de biais qui n'est prise en compte est la distance entre les gènes 16S et l'origine de réplication des bactéries quantifiées, alors qu'un gène 16S à proximité du site d'initiation du dédoublement du génome bactérien sera logiquement sur-représenté par rapport à son vrai nombre de copies. L'établissement de bases de données complètes pour contrôler pour ces billets est, à toutes fins pratiques, impossibles avec les technologies actuelles.

### **11.2.2 Types d'études métagénomiques – WGS**

La technique du séquençage 16S a longtemps été la seule option pour faire de la métagénomique et a permis de caractériser un grand nombre d'écosystèmes de diverses origines. Depuis le développement à grande échelle des technologies de séquençage à haut débit, celles dites de seconde ou de nouvelle génération, une autre méthode a changé le monde de la métagénomique, c'est le Whole Genome Sequencing (ou Shotgun), WGS. Dans cette technique, l'ADN du microbiome est extrait et amplifié dans sa totalité, contrairement à l'approche ciblée de la technique 16S. Le terme Shotgun (lit. fusil à pompe, projectile chevrotine) réfère à la fragmentation aléatoire de tous les génomes bactériens présents et la génération d'amplicons de manière non ciblée. Ces amplicons amplifiés par PCR proviennent de tous les génomes présents, dans leur entièreté et sans région particulière de départ ou de fin. Une génération non-aléatoire des fragments (par exemple en utilisant des enzymes de restriction spécifiques) induirait un fort biais et n'est pas souhaitable. La fragmentation est donc réalisée par des méthodes de cisaillement mécaniques comme la sonication ou par force hydrodynamique (et parfois des enzymes de restriction non spécifiques). Il est courant de chercher à générer des amplicons de quelques centaines de paires de bases, par exemple une campagne canadienne de séquençage WGS de *Mycobacterium tuberculosis* a préparé sa librairie avec des fragments de taille moyenne de 800 bp (600 bp à 1000 bp) (118) dans une publication comparant les techniques de préparation de librairies.

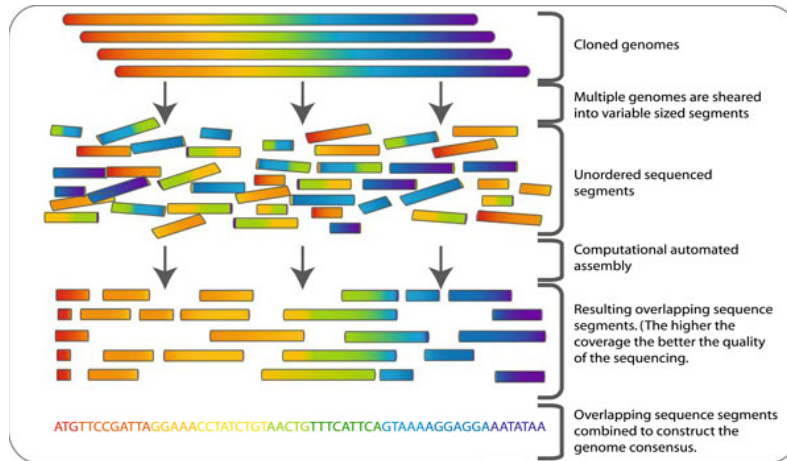


Figure 3. Schématisation de la technique de séquençage de Whole Genome Shotgun (WGS) et de l'assemblage des séquences (119)

L'avantage principal de la technique de WGS est qu'elle permet informatiquement de reconstituer des longs fragments. Ces longs fragments assemblés correspondent souvent à des génomes presque entiers et donc des voies métaboliques et gènes de résistance retrouvés dans chaque espèce. Les petits fragments générés dans la librairie sont séquencés par des techniques à haut débit sur des puces à ADN, à partir des extrémités des amplicons. La taille des fragments générés dans la librairie dépend de la technique et des paramètres choisis pour le cisaillement aléatoire des génomes bactériens. A majorité des plateformes de séquençage lisent chacun des deux bouts des fragments générés, c'est la technique en extrémités-couplées (en anglais 'paired-end') de séquençage, c'est-à-dire que les amplicons sont couplés à des adaptateurs (séquences d'ADN artificielles avec des séquences connues) qui permettent de les fixer sur la puce de séquençage. Les deux extrémités seront lues en étapes successives par l'automate de séquençage et les séquences nucléotidiques seront associées à des pointages de qualité, pour chaque base séquencée, dans un fichier brut de type fastq. Ces fichiers bruts sont des fichiers en texte brut à formatage minimal qui font généralement 2,5 GB par microbiome séquencé et qui peuvent être séquencés beaucoup plus en profondeur et, par exemple dans le jeu de données du présent projet, jusqu'à 15 GB en moyenne par échantillon.

Par rapport à la technique du séquençage ciblé du 16S, la technique de WGS est plus dispendieuse à réaliser à cause du coût des réactifs et des automates impliqués. De plus, la gestion des données nécessite des infrastructures informatiques avancées, comme des grappes de calcul et des 'super-ordinateurs', alors que chaque étape de traitement des données génère aussi des fichiers de grande taille qui atteignent rapidement plusieurs terabytes (TB) de fichiers numériques. Malgré cela, le WGS est une technique très intéressante qui gagne de plus en plus en popularité. Une des raisons est qu'en plus de donner de l'information encore plus précise sur la souche bactérienne observée, elle permet aussi d'établir un portrait fonctionnel du microbiome. En d'autres termes, en plus de savoir qui est retrouvé dans l'échantillon, le WGS permet de savoir ce qu'elles sont en train de faire. Par exemple, en plus de déterminer qu'une bactérie de l'espèce *Escherichia coli*, le WGS permet de déterminer son statut de résistance aux antibiotiques dans un seul test (120) dans un contexte clinique.

Les études WGS permettent donc de déterminer l'abondance relative des différentes espèces bactériennes retrouvées dans un échantillon et l'obtention des séquences relativement complètes permet de déterminer quels gènes elles contiennent et donc quelles voies métaboliques elles peuvent encoder. Les études de métatranscriptomique poussent l'analyse fonctionnelle plus loin en utilisant les techniques de RNAseq (séquençage d'ARN) pour déterminer quels gènes sont activement transcrits. La majorité des RNAseq, qui à l'instar des WGS, sélectionnent aléatoirement les séquences nucléotidiques à amplifier et séquencer, se retrouvent avec de très hautes proportions d'ARNr, qui sont des ARN fonctionnels essentiels à la synthèse protéique. En fait, c'est généralement 90% des ARN ainsi isolés qui sont des ARNr, ce qui rend l'analyse des transcrits peu abondants plus difficile. Des techniques d'enrichissement permettent de sélectionner pour les ARNm, messagers précurseurs des protéines ou encore pour les petits ARN interférant impliqués dans la régulation de l'expression génétique.

La métatranscriptomique permet d'établir le niveau d'activité des gènes présents et de dresser un profil fonctionnel plus complet que les WGS. Toutefois, tous les biologistes moléculaires vous le diront, l'ARN est une molécule beaucoup moins stable que l'ADN et donc

beaucoup plus susceptible aux biais induits par la préparation des échantillons. Les trousse de réactifs offertes par les compagnies de biotechnologies redoublent d'ingéniosité pour circonvenir à ce problème inhérent aux études d'ARN et des études démontrent la robustesse des méthodes de travail. Autrement, une étude a aussi comparé les résultats de WGS et de métatranscriptomique pour trouver que la majorité des génomes bactériens sont transcrits à des taux similaires et relativement fixes chez des communautés bactériennes bien adaptées à leur environnement (121). Cela veut aussi dire que l'abondance du gène retrouvé sous forme d'ADN est un bon indicateur du niveau d'activité des voies métaboliques qui y sont reliées, ou dans d'autres termes, il est possible de déterminer sous conditions normales, les voies métaboliques fonctionnelles actives en utilisant les données des campagnes de collecte de données WGS. La métatranscriptomique demeure un instrument puissant pour observer les réactions rapides des microbiotes à des conditions environnementales changeantes.

### **11.2.3 Bases de données publiques**

Généralement, après avoir assemblé des séquences génomiques à partir des amplicons de WGS ou après avoir regroupé les gènes 16S en OTU, l'étape suivante consiste en l'attribution de ces séquences à des groupes taxonomiques connus. Il existe plusieurs bases de données qui répertorient les bactéries connues avec les annotations fonctionnelles et les informations phylogénétiques. Il existe des bases de données spécialisées pour les génomes complets assemblés, pour les données de WGS brutes ainsi que pour les séquences d'ADN ribosomal 16S attribuées à des espèces bactériennes connues.

Pour les jeux de données de séquençage ciblé 16S, il existe de nombreux outils informatiques permettant de faire la classification des groupes taxonomiques de façon automatisée avec des interfaces graphiques pour usagers. Les interfaces graphiques permettent aux usagers peu expérimentés de lancer les analyses bioinformatiques par des clics de souris et plusieurs groupes de recherche offrent même des plateformes en ligne, qui offrent la possibilité de téléverser les données expérimentales vers leurs serveurs (122). Les analyses sont réalisées et les résultats interprétables pour les humains sont retournés, sous la forme d'arbres



phylogénétiques, profils d'abondance, analyses de composantes principales (PCA, PCoA, MDA...), analyses de diversité alpha et beta, etc. Les algorithmes implémentés dans les suites logicielles sont des sujets de recherche actuelle, alors qu'il existe une multitude de façon de comparer les séquences, comme les algorithmes 'classiques' de classification phylogénétique qui comparent les séquences génétiques par ces principes comme l'horloge moléculaire (123), la vraisemblance maximale (124–126), l'élagage et greffe de sous-arbre (127), etc. Ces techniques de classification de séquences ne sont toutefois pas implémentées spécifiquement pour la comparaison de séquences de 16S, alors que d'autres programmes comme '16S classifier' tient en compte les régions hypervariables décrites plus tôt (128), ou encore un Modèles de Markov Caché (anglais: HMM), qui se base sur les probabilités de transition d'état pour regrouper les séquences 16S (129).

Donc, les outils informatiques offerts peuvent surtout, soit regrouper les séquences en arborescences de voisins les plus probables (phylogénie moléculaire) et/ou les attribuer à des taxons microbiens connus. Pour ce qui est de la phylogénétique, la comparaison de séquences peut se faire relativement facilement et ne requiert que les résultats de séquençage et un algorithme de classification. Pour ce qui est de l'attribution taxonomique, toutes les suites de logiciel doivent utiliser une base de données pour attribuer les séquences à des espèces ou taxons plus larges. Beaucoup de ces outils le font sans effort pour l'utilisateur, avec une base de données intégrée dans la suite de logiciels et l'attribution est ainsi faite automatiquement. D'autres utilisateurs peuvent préférer de télécharger une base de données seule et de faire l'étape d'attribution taxonomique 'manuellement' comme étape dans leur flot de travail informatique. Ainsi, il existe plusieurs bases de données disponibles en ligne et entretenues par des groupes de recherche un peu partout dans le monde, comme la populaire Greengenes database (130), la Ribosomal Database Project (RDP ou la Base de Données Ribosomale) (131) et SILVA (132). Chacune de ces bases de données contiennent des séquences ribosomiques annotées selon leur microorganisme d'origine et sont offertes en conjonction avec des outils bioinformatiques pour les différentes étapes du processus d'analyse de données 16S.

Un autre outil intéressant de bioinformatique est la plateforme Qiime (pour Quantitative Insights Into Microbial Ecology) et est plus un lieu d'échange qu'un logiciel d'analyse à proprement parler. Cet outil implémente d'une manière standardisée plusieurs logiciels populaires (comme dada2, deblur, vsearch, MAFFT, Emperor - PCoA, PICRUST, etc.) et simplifie donc considérablement leur utilisation, surtout dans un contexte de pipeline, qui nécessite une succession (souvent très longue) de logiciels de diverses origines. Les visées des développeurs consistent en l'utilisation simplifiée de logiciels en sources libres dans une plateforme extensible (capable d'intégrer des nouveaux outils) par les efforts de la communauté d'utilisateurs et de développeurs. La plateforme assure aussi la traçabilité des procédures informatiques (les étapes exécutées) pour augmenter la reproductibilité des résultats et fournit les sources des logiciels utilisés (pour donner le crédit aux auteurs). La majorité des outils et pipelines sont axés vers le séquençage d'amplicon 16S et certains plugiciels (plugins) sont en développement pour les jeux de données WGS.

Les plugiciels de Qiime n'utilisent pas les OTU décrits plus tôt dans le chapitre, mais plutôt les ASV (Amplicon Sequence Variants or Exact Sequence Variant). Au lieu de faire des groupes basés sur une similitude de séquence de 95% ou 97% comme les OTU, les algorithmes d'ASV (exemple dada2 ou deblur) comptabilisent seulement des séquences identiques et enlèvent les séquences retrouvées en trop petits nombres dans une étape appelée le 'denoising' (généralement l'étape la plus exigeante informatiquement de tout le pipeline). Les séquences peu abondantes sont considérées comme des erreurs de séquençage et exclues des analyses. Plusieurs publications récentes rapportent une meilleure efficacité des ASV par rapport aux OTU (129–132). Dans des étapes ultérieures du pipeline principal, d'autres plugins de Qiime permettent d'attribuer les ASV à des groupes taxonomiques connus (souvent avec une précision jusqu'à l'espèce) et de générer une table d'abondance des groupes taxonomiques avec une grande précision (et les ASV attribués au même taxon sont regroupées à cette étape). Toutefois, les analyses de diversité (alpha et beta) et les analyses multivariées (PCoA) sont réalisées sur les tables d'abondance des ASV.

Pour ce qui est des jeux de données de séquençage aléatoire en profondeur (WGS), il existe deux directions générales de traitement des données. En effet, on peut soit assembler les plus longues séquences possibles pour reconstituer les versions les plus complètes des génomes bactériens séquencés, dans une approche dite d'assemblage *de novo* ou encore comparer les séquences retrouvées à des génomes bactériens complets et connus dans une approche dite d'assignation des OTUs en référence fermée (en anglais, closed-reference OTU picking).

La technique d'assemblage *de novo* possède l'avantage de permettre de découvrir des nouvelles espèces microbiennes et virales jusqu'à lors inconnues, mais se trouve à être très exigeante au niveau informatique. La technique requiert de comparer les séquences (généralement assez courtes, souvent une centaine de paires de bases environ) pour retrouver les enjambements les plus probables entre leurs extrémités. L'algorithme d'assemblage construit alors un réseau d'interactions des séquences brutes sous la forme d'un graphique De Bruijn et les séquences reconstituent ainsi des portions larges d'information génomique. Les fragments ainsi assemblés virtuellement à partir de leur régions contiguës sont appelés des contigs (en anglais et en français) et la qualité de l'assemblage est souvent jugée supérieure si les contigs sont d'une longueur importante. Un métrique très utilisé à cette fin est celui du N50, qui est une médiane pondérée (l'algorithme accorde légèrement plus de poids aux longs contigs), selon le calcul que 50% des contigs sont plus grands ou égaux à cette valeur. Plus le N50 est élevé, plus les contigs assemblés sont longs, ce qui est généralement un signe de qualité pour le jeu de données, ou minimalement d'une profondeur de séquençage suffisante pour avoir une couverture élevée. Toutefois, il est difficile de comparer les jeux de données entre eux parce que la taille de l'assemblage doit être égal pour que la comparaison tienne.

Donc, l'algorithme d'assemblage *de novo* reconstruit les plus longues séquences possibles à partir des chromosomes bactériens provenant de l'échantillon de microbiome. Avec une très grande profondeur de séquençage, l'on pourrait dire que les longs contigs sont des génomes bactériens, mais il reste toujours une incertitude et il est prudent de ne pas faire ce lien 1 pour 1. Le terme "profondeur de séquençage" réfère à la quantité de fragments qui, une fois

superposés à la référence, s'entrecoupent pour produire plusieurs lectures chevauchant la même région, confirmant les bases séquencées. Concrètement, une plus grande profondeur de séquençage est obtenue en générant une librairie de séquençage plus large et en séquençant plus de fragments pour le même échantillon.

Sachant qu'il est fréquent que les données générées par les campagnes de WGS contiennent plusieurs dizaines de millions de fragments par échantillon de microbiome, il n'est pas surprenant que l'assemblage *de novo* soit très demandant au niveau informatique. Les séquences d'ADN ainsi reconstituées peuvent ensuite être analysées pour y détecter les séquences codantes et ainsi prédire les fonctions métaboliques que les bactéries du microbiome échantillonné. Il existe plusieurs projets qui entretiennent des bases de données, comme le HMP (décrit plus tôt) qui offre plusieurs jeux de données de WGS, le IMG de JGI (pour, en anglais 'Integrated Microbial Genomes and Microbiomes' du projet 'Joint Genome Institute') et le HPMC (traduit, Communautés Pan-Microbiennes Humaines). Leur mission est de favoriser l'échange de jeux de données et l'annotation des génomes et gènes associés. Pour la génétique humaine, il existe une base de données qui est la référence 'en or' pour l'annotation fonctionnelle et l'établissement des voies métaboliques actives qui se nomme KEGG (en anglais Kyoto Encyclopedia of Genes and Genomes). Le projet IMG et le projet HPMC ont constitué et entretiennent des collections de génomes microbiens, mais il n'existe pas l'équivalent pour l'annotation fonctionnelle. L'annotation fonctionnelle des génomes bactériens donc est réalisée avec la base de données de KEGG, qui contient les gènes et voies métaboliques eucaryotes.

Indépendamment des difficultés relatives à l'étude de l'ARN dans les systèmes biologiques, l'application de ces stratégies à l'étude du microbiome amène des nouvelles considérations, comme le fait que les bactéries produisent des ARNm (ARN messagers, codant pour les protéines) qui ne contiennent pas de queue poly-A, contrairement à ceux des cellules eucaryotes. Cette différence oblige à adapter la méthodologie de travail en laboratoire, alors que beaucoup de plateformes de séquençage d'ARNm utilisent cette propriété des ARNm eucaryotes pour amorcer le séquençage. L'analyse des données de métatranscriptomique apporte aussi

plusieurs autres considérations, par exemple, il n'y a pas encore de flot de travail standardisé pour ce type d'expériences à cause des spécifications techniques de chaque protocole de séquençage de RNAseq (133). Ainsi, beaucoup des outils proposés ne sont pas des plateformes intégrées avec une interface graphique pour l'utilisateur, mais plutôt des suites de programmes dont il faut ajuster manuellement les étapes pour assurer un déroulement sans heurt du flot de travail. De plus, un outil d'analyse de données de métatranscriptomique doit non seulement inclure les sous-programmes nécessaires au traitement des données issues de différents protocoles, mais aussi être efficace au niveau de la gestion de la mémoire informatique et du temps d'exécution. Il est fréquent de voir des sections de code qui nécessitent plus de 24 heures à 48 heures d'exécution ininterrompue.

En 2016, deux logiciels ont été lancés pour permettre aux utilisateurs moins expérimentés en bioinformatique d'analyser leur données de métatranscriptomique, soit MetaTrans (133) et SAMSA (134), qui permettent de traduire les données brutes de séquençage en sommaires numériques et graphiques interprétables. Toutefois, ces suites de logiciel nécessitent des ordinateurs puissants et une configuration informatique particulière. D'autres projets comme COMAN (135) offre une interface web qui permet de téléverser les données d'expérimentations vers leur serveur dédié qui est capable de conduire les analyses grâce à une base de donnée et des algorithmes développés sur place. Les résultats des analyses peuvent ensuite, sans trop d'efforts, être récupérées par l'utilisateur.

## **11.3 Pharmacomicrobiomique**

### **11.3.1 Le microbiome comme effecteur pharmacologique**

Il est reconnu depuis un certain temps que les fibres alimentaires réduisent le taux de cholestérol sanguin et le risque de maladies cardiovasculaires (15,136–144), réduisent l'absorption des métaux lourds ingérés (145–147) ainsi que le risque de maladies infectieuses (148,149). Toutefois, au-delà des corrélations très claires, les mécanismes derrière les améliorations de santé observées en présence de fibres alimentaires restent nébuleux et plusieurs

théories ont été proposées au fil des années. Une étude récente met en lumière le mécanisme par lequel certaines bactéries protègent l'hôte contre les pathogènes en créant la couche de mucus dans un mécanisme dépendant des fibres alimentaires (150). Le maintien de l'intégrité de ce revêtement est garant de santé, agissant comme la première ligne de défense du tube digestif. Ce serait par cette interaction que la bactérie *Akkermansia muciniphila* est bénéfique pour la santé de l'hôte humain. Cette dernière est aussi clairement corrélée avec un état de santé supérieur, mais cette interaction ne semble pas passer par des métabolites ou autre voies classiques bien connues.

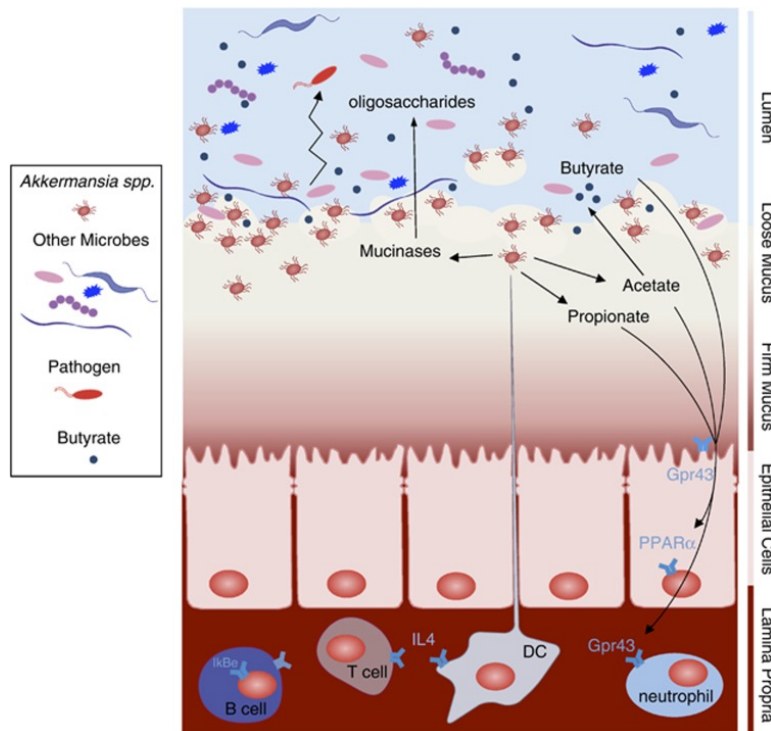


Figure 4. Le rôle de *Akkermansia Muciniphila* dans le maintien de l'intégrité de la barrière de mucus intestinale. Tiré de (151).

Le médicament antidiabétique metformine regagne en popularité depuis quelques années pour son action diminuant le glucose sanguin et ses effets secondaires presque inexistant permettant de l'utiliser en prévention du déclenchement du diabète. Il est aujourd'hui le médicament le plus couramment prescrit pour prévenir le diabète de type 2. Des études ont démontré les effets de la metformine sur les effecteurs classiques de la régulation du glucose,

comme l'AMPK, mais les liens ne sont pas établis hors de tout doute (152,153). D'autres chercheurs pensent plutôt que l'action du médicament s'opère dans l'intestin et non pas dans la circulation (10,154,155). Un microbiome modifié qui métabolise les aliments différemment pourrait en effet bien expliquer l'absorption réduite du glucose au niveau de l'intestin. Il faut toujours garder en tête de s'assurer que ces associations ne sont pas que corrélatives, mais bien causales. Dans tous les cas, plusieurs études démontrent une efficacité de traitement associée à une modification de la composition du microbiome (9,156,157) et d'autres suggèrent même de combiner la metformine à des modulateurs connus du microbiome pour augmenter l'efficacité de la thérapie hypoglycémiante (8). Les changements de microbiome sont généralement associés à une modulation à la hausse de la population de *Akkermansia Muciniphilia* (158,159), la bactérie associée à la couche de mucus citée plus tôt. Dans ces articles, on retrouve aussi quelques indices du rôle probable du microbiome dans le mécanisme d'action de la metformine, comme le fait que le médicament est moins efficace lorsqu'administré en intraveineuse que *per os*.

Dans le cas de pathologies cutanées atypiques comme l'eczéma et le psoriasis, le microbiome observé par métagénomique est différent de celui d'une peau normale. Cela permet aujourd'hui de mieux caractériser l'étiologie de ces maladies encore mal comprises et ainsi de permettre un meilleur diagnostic basé non seulement sur les symptômes, mais aussi sur les populations bactériennes qui peuvent être en cause. De plus, une meilleure compréhension du microbiome et de ses implications permettra sans doute de développer des nouvelles stratégies de traitement, comme d'utiliser des probiotiques (bactéries commensales actives), qui pourraient rétablir l'équilibre et la diversité associée à un microbiome sain. Dans l'avenue de la médecine personnalisée, une équipe a publié un article où ils poussent le pouvoir pronostic des tests métagénomiques pour assister dans le choix de médicament à prescrire (160).

### **11.3.2 Interactions avec les xénobiotiques (mécanismes directs et indirects)**

Les médicaments administrés *per os* (P.O. i.e. par voie orale) sont sujets à plusieurs étapes de transformation avant d'accéder à la circulation sanguine systémique, à savoir les

processus du métabolisme de premier passage. En pharmacologie, beaucoup de l'attention est portée sur les effets métaboliques des enzymes endogènes présentes surtout dans le foie (ex. CYP2D6, CYP2C19, UGT1A1, etc.), avec les polymorphismes génétiques des différentes enzymes déjà bien étudiés pour leurs effets sur la pharmacocinétique des médicaments. Les effets de métabolisme de premier passage des cytochromes associés à la paroi intestinale (surtout le CYP3A4) sont aussi reconnus. Depuis quelques années, de plus en plus de chercheurs s'intéressent à la branche naissante de la pharmacomicrobiomique, soit l'effet des enzymes produites par les microbiotes sur le devenir des médicaments ou la pharmacologie du microbiome.

En plus de l'effet du microbiome sur les médicaments lors du métabolisme de premier passage, plusieurs médicaments sont conjugués à la bile (qui permet l'émulsion des gras alimentaires à l'entrée du duodénum) et sont retournés dans le système digestif par la vésicule biliaire. Ainsi, même en utilisant la voie intraveineuse, il est possible que les médicaments soient exposés à l'activité métabolique complexe du microbiome intestinal (161,162). Aussi, dans le cas du métabolisme de premier passage, ce sont les produits enzymatiques du microbiome intestinal qui agissent sur les médicaments administrés *per os*, mais il n'est pas déraisonnable de concevoir que les médicaments en application topique sont aussi transformés par des bactéries du microbiome de la peau ou du vagin par exemple, ou encore qu'un traitement soit efficace en modifiant le microbiome du tissu ciblé. Bref, de plus en plus d'études démontrent l'importance de prendre en compte les bactéries et leurs voies métaboliques actives pour le développement moderne de la médecine personnalisée et des nouveaux médicaments.



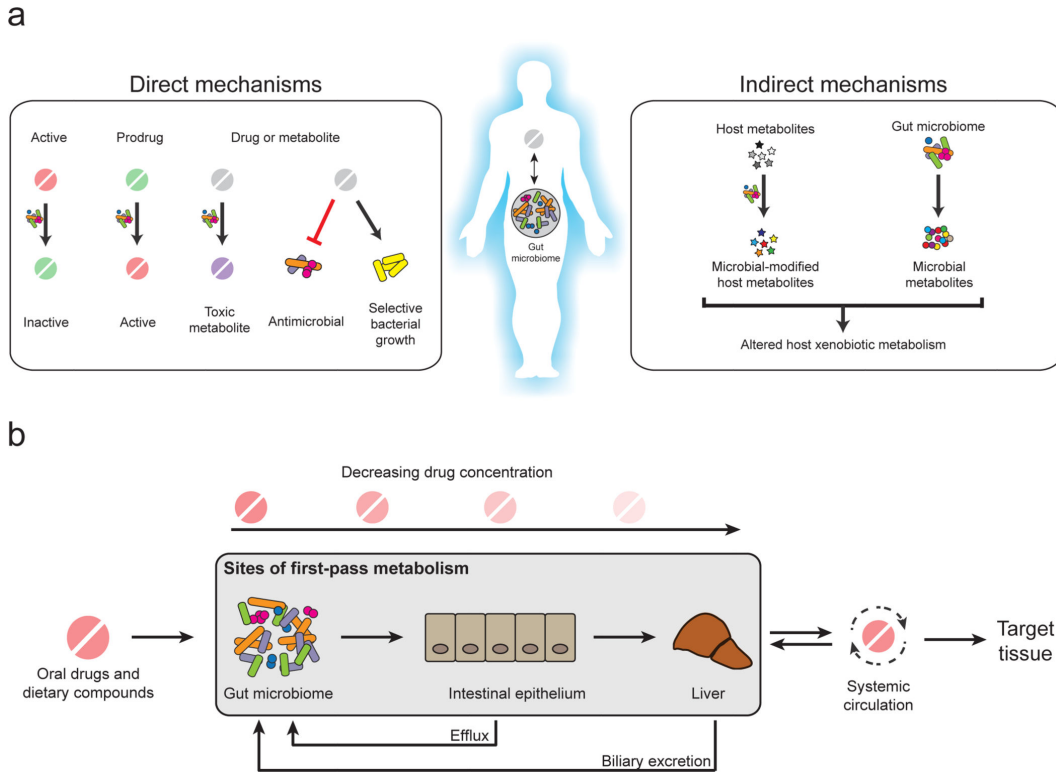


Figure 5. Principales classes d'interactions entre le microbiome et les xénobiotiques (et autres médicaments). Tiré de (163)

Les mécanismes par lesquels un microbiome modifie la pharmacologie de certains médicaments et autres xénobiotiques sont classifiés en deux grandes catégories, soit les interactions directes et les interactions indirectes, come illustré à la Figure 5. Les interactions directes (fig. 5-a.) sont celles où un médicament est métabolisé par les enzymes bactériennes de manière directe ou encore lorsque les populations bactériennes sont directement modulées par l'action du xénobiotique. La transformation des médicaments directement par les enzymes bactériennes peut avoir plusieurs effets, comme l'activation d'une prodrogue en son métabolite actif, l'inactivation d'un médicament actif ou encore la transformation d'un xénobiotique en un métabolite toxique. La quatrième classe d'interaction qualifiée de directe (fig. 5-a.) est celle par laquelle la substance exogène perturbe la croissance et la survie des microbiotes, induisant ainsi une pression de sélection sur les bactéries présentes. Un exemple de ce type d'interaction serait la sélection des espèces bactériennes résistantes aux antibiotiques par l'exposition fréquente du

microbiome à des antibiotiques, par exemple chez les animaux de ferme régulièrement supplémentés aux antibiotiques pour accélérer leur croissance.

En aparté, l'utilisation d'antibiotiques dans le domaine agroalimentaire est préoccupante. En plus d'accumulation de gènes de résistance dans les bactéries qui peuplent les animaux (164), même l'épandage de fumier constitue une source de contamination environnementale considérable aux bactéries résistantes (165,166) et la contamination aux gènes de résistance s'observe à large échelle dans les estuaires (167). En fait, l'utilisation médicale à grande échelle des antibiotiques est une problématique environnementale (168,169) et certains pensent que les stations d'épuration des eaux usées pourraient servir de haut lieu d'accumulation et d'échanges de gènes de résistance (170)

Les antibiotiques sont sans contredits les médicaments qui ont le plus révolutionné la médecine moderne en sauvant un nombre incalculable de vies. Toutefois, même lorsqu'ils sont utilisés convenablement, les antibiotiques entraînent généralement des dysbioses du microbiome qui sont souvent associées à un transit perturbé (ex. diarrhées) dans 5% à 30+% des cas (171,172) et l'utilisation des antibiotiques est un facteur de risque pour la bactérie *C. difficile* (173). Dans tous les cas, l'utilisation d'antibiotiques réduit la diversité taxonomique des microbiomes (174,175) et entraînent l'accumulation de gènes de résistance (176) et des espèces qui les portent, comme les *Enterobacteriaceae* (177).

En pharmacomicrobiomique, les interactions indirectes (fig. 5-b.) sont celles par lesquelles les métabolites produits ou altérés par le microbiome influencent le métabolisme endogène de l'hôte humain. Par exemple avec le modèle murin, les souris dépourvues de microbiotes ont plus de 100 gènes métabolisateurs exprimés différemment dans leurs foies (surtout les CYP) par rapport à leur congénères élevées normalement, démontrant l'effet du microbiome sur le métabolisme de l'animal (178). Les animaux qui n'ont pas de microbiome normal voient leurs voies de métabolisation amplifiées et sont, par exemple, plus efficaces à métaboliser le phénobarbital (anesthésiant largement utilisé en laboratoire avec les souris). Il

existe plusieurs autres exemples de ce type d'interactions, comme celui du TMAO que nous verrons plus en détail dans une prochaine section.

### **11.3.3 Interactions directes**

Le premier type d'interaction directe est celui où une prodrogue (médicament sous sa forme inactive) qui est transformé en métabolite actif. C'est le cas avec la sulfasalazine, un médicament de première ligne dans le traitement de l'arthrite rhumatoïde aussi utilisé dans le traitement des colites ulcéreuses et de la Maladie de Crohn. La sulfasalazine est métabolisée en sulfapyridine par les bactéries du microbiome et c'est d'ailleurs la concentration plasmatique de ce métabolite actif qui est monitoré dans le sang des patients sous thérapie. C'est aussi un exemple du quatrième type d'interaction, alors que le médicament a été conçu artificiellement pour être dégradé dans l'intestin en acide 5-aminosalicylic (5-ASA), un anti-inflammatoire et en sulfapyridine cité plus tôt, un antibiotique. Cet exemple illustre de deux types d'interactions alors que c'est une interaction directe du microbiome qui hydrolyse une prodrogue en un antibiotique qui agit directement sur le microbiome.

Le deuxième type d'interaction est celui où un médicament est transformé d'une forme active vers un métabolite inactif. C'est le type d'interaction qui s'observe avec le méthotrexate (MTX), un médicament largement utilisé en chimiothérapie pour traiter plusieurs types de cancers et, avec des dosages significativement réduits, utilisé pour traiter certaines maladies auto-immunes. La diminution de la biodisponibilité du MTX n'est pas une préoccupation majeure, alors qu'à des doses élevées, le métabolisme microbien n'affecte pas une grosse partie de la dose administrée et que, lorsqu'utilisé pour le traitement des maladies auto-immunes, le médicament est bien toléré et son index thérapeutique est donc assez large. L'interaction est largement étudiée pour ses implications au niveau de la médecine de précision, alors que ce ne sont pas tous les patients qui réagissent aux traitements (179). La composition initiale du microbiome peut prédire l'efficacité du Mtx. Un autre exemple largement étudié de ce second type d'interaction est celui de la diminution de la biodisponibilité de la digoxine, où les bactéries peuvent métaboliser plus de 50% de la dose administrée P.O., ce qui est problématique parce

que l'index thérapeutique de cet anti-arithmétique est très faible (0.5 ng/ml à 2 ng/ml). Nous revendrons beaucoup plus en détail sur cette interaction pharmacomicrobiomique dans le chapitre principal.

Le troisième type d'interaction directe est celui de la transformation de xénobiotiques inactifs en métabolites toxiques (pour l'hôte) par le métabolisme enzymatique des bactéries. Pour cet exemple, proposons l'étude du composé organique mélamine, qui est très riche en azote qui avait été au cœur d'une controverse en Chine en 2008 alors que des dizaines de milliers d'enfants et de nourrissons se sont retrouvés à l'hôpital (dont quatre sont décédés). La mélamine a été utilisée pour falsifier les tests de contenu protéique du lait et elle est métabolisée par les bactéries comme *Klebsiella terrigena* pour former un métabolite toxique, l'aide cyanurique. Ce dernier se repolymérise dans les reins en se combinant à la mélamine et les composés y forment des pierres aux reins qui mène à l'insuffisance rénale grave pouvant mener à la mort (180).

### **11.3.4 Interactions indirectes – cas tu TMAO**

Un exemple d'interaction indirecte entre les xénobiotiques et le microbiome est celui du TMAO (Triméthylamine N-oxyde). Les bactéries métabolisent des nutriments introduits dans l'alimentation, comme la choline, et forment le TMA (Triméthylamine), qui agit comme une hormone endocrine une fois en circulation.

Il existe une corrélation claire entre les taux circulants de TMAO (Triméthylamine N-oxyde) et les risque de maladies cardiovasculaires, particulièrement l'athérosclérose. Le TMAO est produit par l'oxydation endogène du TMA (Triméthylamine), un métabolite dérivé de la dégradation des nutriments alimentaires comme la choline et la L-carnitine. Ce qui est intéressant par rapport au TMA, c'est qu'il est exclusivement produit par les bactéries du microbiome intestinal et aucunement par l'hôte animal. Dans plusieurs modèles animaux et études cliniques les niveaux de TMAO corrént aussi fortement avec des traits liés à l'obésité et le diabète de type 2.

La conversion du TMA en TMAO est attribuable à l'enzyme flavine monooxygénase 3 (FMO3), qui est par la même occasion un joueur majeur dans le métabolisme du cholestérol. Les souris transgéniques n'exprimant pas (KO-) FMO3 qui se sont fait nourrir des diètes riches en cholestérol sont caractérisées par une excrétion augmentée de cholestérol dans les fèces, accompagnée d'une accumulation du TMA sanguin, mais pas du TMAO. Toutefois, l'excrétion accrue du cholestérol est accompagnée d'une diminution des taux hépatiques de LXR (les récepteurs des oxystérols ou liver X receptors). La baisse de LXR résulte en l'inflammation du foie et le stress du réticulum endoplasmique, des effets secondaires très importants.

D'une manière intéressante, certains humains possèdent naturellement une mutation inactivant FMO3, mais ils sont affectés d'une condition singulière nommée triméthylaminurée (TMAU), aussi connu sous le nom de "fish odor syndrome" (syndrome de l'odeur de poisson). En effet, comme c'est le cas pour le modèle de souris KO-FMO3, le TMA s'accumule en circulation. Toutefois, cette dernière molécule est célèbre pour l'une de ses particularités: son odeur puissante de chair de poisson vieillissante. En effet, le TMAO est abondant chez les animaux aquatiques et les bactéries présentes sur les produits de la mer dégradent le TMAO (inodore) vers le TMA (181,182).

Comme mentionné plus tôt, il est de plus en plus accepté que le microbiome intestinal a une influence majeure sur l'hôte mammifère, que ce soit au niveau comportemental ou encore sur le niveau du métabolisme énergétique. Le TMA produit par les microbiotes pourrait agir à titre d'hormone endocrine lorsqu'elle rejoint la circulation par la veine portale hépatique. Cette hypothèse s'appuie initialement sur l'observation que les niveaux de TMA atteignent leur sommet deux heures suivant l'ingestion d'un repas. Le TMA est un agoniste puissant du récepteur encodé par le gène TAAR5 (trace amine-associated receptor 5). Cette protéine, ainsi que la majorité des membres de la famille TAAR, est fortement exprimée au niveau de l'épithélium nasal et plusieurs spéculent que le TMA pourrait agir comme phéromone chez les souris. Un lien intéressant peut être fait avec les études récentes où il a été démontré que la diète

influence les préférences sexuelles des mouches drosophiles pour les individus qui consomment la même diète (183,184). En plus de ce rôle de récepteur olfactif, la protéine TAAR5 est exprimée dans la majorité des tissus et pourrait servir de régulateur métabolique.

À l'instar de la protéine Taar5, le récepteur ADRB1 (beta-1 adrenergic receptor ou  $\beta$ 1 adrenoreceptor) est un récepteur couplé à la protéine G. ADRB1 est régulé à la hausse par l'injection de TMA ou de TMAO et une telle injection induit la phosphorylation d'une variété de kinases effectrices clés, majoritairement dans la cascade de signalisation du PKA.

Dans tous les cas, comme détaillé plus tôt, l'étape de conversion du TMA vers le TMAO ne constitue pas une cible pharmacologique intéressante à cause de la toxicité hépatique encourue et de la potentielle odeur corporelle. Toutefois, l'inhibition de la production du TMA par les bactéries du microbiome est sous investigation. L'analogue de la choline DMB (3,3-dimethyl-1-butanol) est apte à inhiber la conversion de la choline en TMA dans un mécanisme qui ne dépend pas des espèces bactériennes présentes. Dans les études menées sur des souris de laboratoire, il a démontré une capacité à réduire le poids corporel des animaux, alors qu'un analogue du DMB, le CCO8 a protégé les souris dépourvues de (KO-) ghreline contre l'obésité, dans un contexte de sur-alimentation.

## **11.4 La digoxine et le microbiome intestinal**

### **11.4.1 La digoxine dans le traitement de l'insuffisance cardiaque (IC) et de la fibrillation auriculaire (FA)**

La digoxine est un des plus anciens médicaments de la pharmacopée de l'Europe moderne, avec plus de 200 ans d'utilisation depuis sa description et son extraction à partir de la plante *digitalis purpurea* (grande digitale) par le britannique William Withering (185). Le terme digoxine réfère au principe actif, un cardioglycoside (à l'instar du Ouabain) qui est maintenant extrait à partir des feuilles de la *digitalis lantana* (digitale laineuse) et vendue majoritairement

sous le nom de lanoxin. La digoxine a été très largement prescrite et occupe même un pan important de l'imaginaire culturel, alors que le médicament est cité comme arme du crime dans plusieurs fictions. La digoxine a aussi été impliquée dans une série de meurtres à l'hôpital pour enfants malades de Toronto au début des années 1980, une affaire sordide qui reste irrésolue à l'heure actuelle. Bref, les effets secondaires des doses trop élevées de digoxine peuvent avoir des conséquences sérieuses sur la santé. D'une manière folklorique aussi, plusieurs études pointent aussi vers la possibilité que les traitements à la digoxine aient pu avoir un impact important sur certaines œuvres du célèbre peintre hollandais Vincent Van Gogh. Cette hypothèse est basée sur le fait que le peintre recevait de la digoxine comme traitement pour l'épilepsie, que deux de ses toiles présentent son médecin avec de la digitale dans les mains et que l'un des principaux effets secondaires de la digoxine est une vision floue tirant sur le jaune, possiblement un comme le fameux tableau 'nuit étoilée' (186,187).

La digoxine est efficace pour améliorer les symptômes de d'insuffisance cardiaque (IC) parce que c'est un inotrope positif. Cela signifie que le cardioglycoside augmente la force des battements cardiaques en augmentant la contractilité des fibres musculaires. La digoxine est aussi utilisée pour traiter la fibrillation auriculaire (FA), l'arythmie cardiaque la plus fréquente qui affecte les oreillettes cardiaques avec un rythme trop élevé. Pour cette raison, la digoxine est employée pour son effet chronotrope négatif, c'est-à-dire qu'elle ralentit la fréquence des pulsations cardiaques. La cible moléculaire de la digoxine est la pompe ATPase  $\text{Na}^+/\text{K}^+$ , dont elle inhibe l'action. Cela a pour conséquence de faire augmenter la concentration du sodium ( $\text{Na}^+$ ) intracellulaire dans les cardiomyocytes, ce qui provoque une augmentation subséquente du calcium intracellulaire, augmentant ainsi la contractilité du muscle cardiaque. La digoxine diminue la fréquence cardiaque (chronotrope négatif), empêche la conduction du stimulus électrique (dromotrope négatif) et promeut l'excitabilité myocardique (bathmotrope positif) (188). Ces améliorations cliniques claires de la digoxine en ont fait un médicament très populaire pour le traitement de l'IC et de la FA. Toutefois certains rappelleront que cette mise en marché a été faite avant la mise en place des essais double randomisés contemporains.

Étant donné que l'action de la digoxine passe par un effet sur les ions potassium et calcium, certaines conditions de déséquilibres électrolytiques rendent la prescription de la digoxine peu recommandée. En effet, pour les patients en situation d'hypokaliémie et les patients en situation d'hypercalcémie, les cardioglycosides ne sont pas recommandés, car leur action sur les cardiomyocytes est altérée (189). Aussi, la toxicité à la digoxine est souvent accompagnée d'une augmentation des ions potassium, menant à l'hyperkaliémie (190,191), mais contrairement à ce que les directives médicales courantes peuvent suggérer, ces cas d'hyperkaliémie ne devraient pas être traités avec l'apport d'ions calcium (192).

La digoxine n'est pas le seul cardioglycoside présent dans la digitale laineuse et l'autre principe actif le plus utilisé en cardiologie est la digitoxine. Cette dernière a été isolée beaucoup plus tard que la digoxine, vers 1875. Leur action pharmacodynamique (action moléculaire) est très similaire et les molécules sont chimiquement très semblables aussi (voir Figure 6). Toutefois, la digitoxine est plus stable et la pharmacocinétique (ADME - Absorption Distribution Métabolisme et Excrétion) est très différente, alors que la digoxine est très majoritairement excrétée inchangée par les reins comparativement à la digitoxine qui est éliminée par l'action du foie, ce qui rend cette dernière plus indiquée pour l'utilisation avec les patients qui ont une fonction rénale basse ou erratique. Possiblement à cause de son nom qui ressemble à un nom qu'on aurait pu donner à une toxine de l'ère digitale, la digitoxine n'a jamais été aussi populaire que la digoxine dans l'utilisation en cardiologie, malgré ses propriétés pharmacocinétiques légèrement supérieures. Une étude américaine datant de 2000 sur les taux d'hospitalisation de patients de 60 ans et plus traités soit avec l'un ou l'autre des cardioglycosides a démontré un risque trois fois moins élevé de toxicité pour les patients utilisant la digitoxine (193).



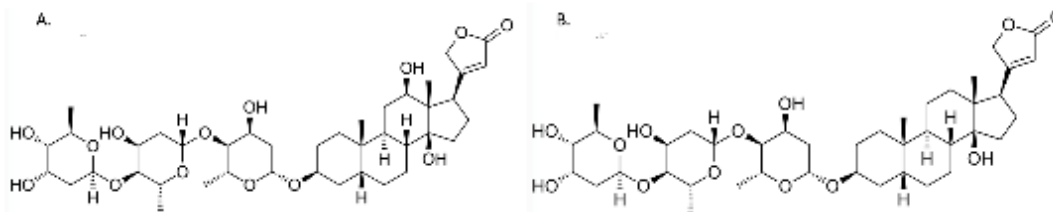


Figure 6. Structures chimiques des cardioglycosides principaux de la digitale laineuse (*digitalis lanata*). A. structure de la digoxine B. Structure de la digitoxine.

### 11.4.2 L'utilisation contemporaine de la digoxine

L'intoxication à la digoxine est la réaction adverse à un médicament la plus fréquemment rencontrée en clinique (14). Les effets délétères les plus courants de l'intoxication à la *digitalis* sont les nausées, vomissements, l'anorexie, des symptômes neurologiques et des arythmies pouvant causer la mort (194). Dans une étude de 2016 sur l'utilisation de la digoxine en pédiatrie, les nausées affectaient 36,4% et la tachycardie 29,5% des enfants traités (195). Comme mentionné plus tôt, la digoxine a une très longue histoire d'utilisation médicale, mais ce n'est que dans les années 1980 qu'elle est vraiment entrée en usage pour traiter l'insuffisance cardiaque et améliorer le rythme sinusal, suite à la publication de trois études cliniques qui ont démontré une amélioration de la tolérance à l'exercice et une amélioration des symptômes de l'IC congestive (196).

La digoxine a continué à être très largement prescrite et utilisée jusqu'à la parution en 1997 d'une étude randomisée en double aveugle où la digoxine a été comparée à un placebo. L'étude du Digitalis Investigation Group (DIG) a démontré que la digoxine ne réduisait pas la mortalité dans la cohorte de 6800 patients par rapport à un placebo (197), ce qui a amené plusieurs à questionner l'intérêt de ce médicament qui est encore classé dans la liste des médicaments essentiels par l'agence mondiale de la santé. En fait, l'étude a démontré que la digoxine réduit le taux d'hospitalisation, mais ne réduit pas la mortalité globale pour les usagers. Un article en italien paru en 2017 y va même d'un titre plus tranché se traduisant comme-ci: "La digoxine, un médicament à envoyer aux poubelles?" (198). Bref, il est clair qu'un médecin qui

voudrait prescrire le cardioglycoside doit redoubler d'attention pour prévenir les effets secondaires indésirables.

À l'inverse des articles tout juste cités, une étude de 2013 a stratifié les résultats de la cohorte de DIG et ils ont démontré que chez les patients atteints d'IC avec un haut risque (tel que défini selon les critères de l'association de cardiologie de New-York, NYHA) voyaient une amélioration non seulement du risque d'hospitalisation sur le suivi de deux ans, mais aussi une amélioration au niveau du risque de décès (199). Aujourd'hui au Canada, les pharmaciens et médecins n'utilisent plus la digoxine comme médicament de première ligne à cause de l'étroitesse de son index thérapeutique, mais l'utilisent encore pour traiter les patients qui restent symptomatiques malgré l'utilisation de bloqueurs-beta ( $\beta$ -blocker), avec un diurétique de l'anse rénale et un inhibiteur de l'Enzyme de Conversion de l'Angiotensine (ACEi). Avec les études démontrant le potentiel néfaste de la digoxine, certaines études ont été faites pour voir l'effet de l'arrêt de la médication et plusieurs ont conclu qu'il était nécessaire de poursuivre la médication afin de traiter l'IC et la FA (200–202). Il vaut de noter que la majorité de ces études ont été réalisées avant l'utilisation large des bêta-bloquants et des ACEi, mais que les médecins préfèrent généralement continuer à prescrire la digoxine aux patients qui l'utilisent déjà régulièrement pour ne pas risquer la décompensation du cœur. Avec ces études, la société américaine de l'IC recommande maintenant des doses maximales de digoxine dans un intervalle beaucoup plus étroit, c'est-à-dire qu'au lieu de l'ancienne recommandation de concentration dans le sérum de 0.5-2.0 ng/ml de plasma et maintenant un intervalle de 0.5-0.9 ng/ml (203–205). En fait, même dans l'étude DIG (Digitalis Investigation Group) citée plus tôt comme la première source d'inquiétude par rapport à l'utilisation de la digoxine, un sous-groupe avec les patients dont la concentration de digoxine dans le sérum à l'intérieur de la valeur inférieure de cet intervalle avaient une amélioration nette au niveau de la survie et des risques d'hospitalisation. Ainsi, la digoxine semble bien avoir des effets positifs sur la santé des patients atteints d'IC et de FA, mais ces effets peuvent être contrebalancés par les effets toxiques aux concentrations plus élevées. Dans le cas des cardioglycosides, il est même fréquent que le terme 'index thérapeutique' soit remplacé par le 'ratio niveau thérapeutique – niveau toxique', qui est très bas pour la digoxine.

Avec son index thérapeutique très étroit, la digoxine est reconnue pour pouvoir augmenter à des niveaux toxiques suite à l'interaction avec de nombreux médicaments, comme les Anti-inflammatoire non stéroïdiens, les ACEi, les antagonistes du récepteur II de l'angiotensine, la ciclosporine, ainsi que les antibiotiques macrolides et certains médicaments pour le cœur (spécialement l'amiodarone) (206). Les interactions avec les médicaments cardiaques comme l'amiodarone peuvent s'expliquer par les effets additifs des deux médicaments agissant sur le rythme cardiaque. D'autres interactions délétères, comme celle avec les macrolides, sont généralement expliquées par l'inhibition compétitive entre les deux molécules, soit l'antibiotique et la digoxine. La P-glycoprotéine 1 (P-gp) a été découverte en 1976 sur les cellules ovariennes mutantes de hamsters chinois (207) et son activité de transporteur transmembranaire a été associée au transport d'une multitude de médicaments, que ce soit au niveau de l'absorption intestinale, de l'exclusion cellulaire ou surtout, de l'excrétion au niveau rénale. Ainsi, l'interaction entre les macrolides et la digoxine est engendrée par le fait que les deux types de médicaments sont excrétés dans les reins pas la P-gp. Selon cette explication, la digoxine n'est plus éliminée aussi rapidement qu'à l'équilibre, elle s'accumule et entraîne de la toxicité.

Tableau 1. Résumé des propriétés pharmacocinétiques de la digoxine chez l'humain.

<b>Paramètres pharmacocinétiques de la digoxine</b>	<b>Valeur</b>
Biodisponibilité orale	~70%
Clairance (CL; élimination)	6 L/h
Volume de distribution (Vd)	400 L
Demi-vie (t ½)	40 h

### 11.4.3 L'interaction *Eggerthella lenta* - digoxine

Il est connu depuis plus de 35 ans que la bactérie *Eggerthella lenta* a la capacité de métaboliser la digoxine (208). En 1982 lors de cette découverte initiale, ce microbiote était encore connu sous le nom de *Eubactérium lentum*. Le métabolite de transformation principal de la digoxine est la dihydrodigoxine, un produit de la réduction du cycle lactone de la partie stéroïde du cardioglycoside et il est produit par le microbiote *E. lenta*, comme illustré à la Figure 7. Les produits de digoxine réduite (PDR) qui ont une affinité réduite pour leur cible et qui sont plus solubles et sont éliminés plus rapidement.

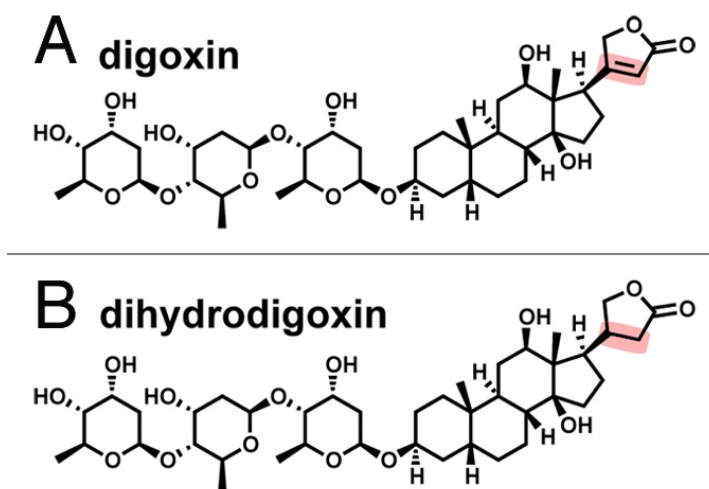


Figure 7. Digoxine et son métabolite principal, la dihydrodigoxine (209)

Dans une étude réalisée sur la pharmacocinétique de la digoxine réalisée avec trois sujets sains, dont les résultats sont présentés aux Tableau 2 et au Tableau 3, pour les résultats de l'endroit de récupération de la digoxine radio-marquée et de ses principaux métabolites. Un point intéressant à remarquer dans le Tableau 2 est le fait que, comme mentionné plus tôt, même lorsque la digoxine est administrée directement par intraveineuse (IV), une partie importante de la dose est retournée dans le tube digestif, en association avec la bile. En effet, on retrouve 17,1% de la dose dans les selles des participants, ce qui indique la circulation entéro-hépatique de la digoxine. Il y a ensuite une réabsorption intestinale du médicament, alors que l'on retrouve

le métabolite dihydrodigoxine dans l'urine, comme présenté au Tableau 3. Cela signifie que la digoxine administrée en IV est sujette au métabolisme de premier passage, bien qu'à des niveaux plus bas que lorsqu'administrée par voie orale (P.O. *per os*).

Tableau 2. Tableau sommaire du lieu de récupération de la digoxine administrée à des participants sains

	Administration intraveineuse		Administration voie orale	
	Pourcentage de la dose (%)	Variabilité	Pourcentage de la dose (%)	Variabilité
<b>Urine</b>	81.3	+/- 2.0	65.7	+/- 1.98
<b>Selles</b>	17.1	+/- 2.8	31.6	+/- 7.6

Tableau 3. Tableau des principaux métabolites de la digoxine retrouvés dans l'urine après administration à des participants sains

	Administration intraveineuse		Administration voie orale	
	Pourcentage de la dose (%)	Variabilité (%)	Pourcentage de la dose (%)	Variabilité (%)
<b>Digoxine</b>	75.6	+/- 3.0	51.5	+/- 11.4
<b>Dihydrodigoxine</b>	2.8	+/- 1.6	4.5	+/- 3.9
<b>Digoxigénine bisdigitoxoside</b>	1.6	+/- 0.1	1.9	+/- 0.1
<b>Métabolites additionnels</b>	1.5	+/- 0.3	6.8	+/- 4.4

L'autre métabolite en importance est le digoxigénine bisdigitoxide et il est généré par l'hydrolyse d'un groupement glucose sur les trois qui sont attachés au cardioglycoside natif. Ce type de métabolisme est réalisé pas les enzymes endogènes humaines (210). Le processus peut être répété pour former digoxigénine monodigitoxide et la digoxigénine simple, mais ces métabolites sont générés en très petite quantité et sont éliminés sans s'accumuler. À chacune des étapes de dégradation des groupements glycosides, il peut y avoir formation de métabolites conjugués, mais leur concentration est encore plus négligeable. La dihydrodigoxine peut elle aussi être métabolisée par les enzymes endogènes, mais encore ici, les proportions sont

négligeables. Les métabolites bisdigitoxide et monodigitoxide sont considérés comme approximativement aussi cardioactifs que la digoxine (211,212), mais le sous-groupe stéroïde seul et ses conjugués polaires sont beaucoup moins actifs. Dans tous les cas, ces métabolites issus du métabolisme humain sont rapidement éliminés (213).

#### **11.4.4 Les études publiées depuis 35 ans**

La réduction de la digoxine a été rapporté pour la première fois en 1981 par l'équipe du Dr. Lindenbaum (214). Avec son équipe, ils ont rapporté que chez 10% des volontaires sains prenant de la digoxine P.O., la digoxine est fortement métabolisée en DRP (Produits Réduits de Digoxine, surtout la dihydrodigoxine) et que cette tendance pouvait être partiellement ou entièrement arrêtée par l'administration d'érythromycine ou de tétracycline. L'année suivante, une autre publication paraissait où ils ont démontré *in vivo* et *in vitro* que c'était la bactérie *Eggerthella lenta* (alors appelée *Eubacterium Lentum*) qui était responsable de cette activité de réduction de la digoxine vers les DRPs (dihydrodigoxine et ses dérivés) (208). Sachant que la bactérie est normalement présente chez une large proportion des participants aux études, ils ont posé la question de déterminer ce qui pouvait faire que les taux de métabolisme étaient différents entre les individus. Ils avaient remarqué que, lors que les microbiotes étaient cultivés *in vitro*, le taux de métabolisme était inversement proportionnel à la concentration en arginine additionnée au milieu de culture (11), c'est-à-dire que plus la concentration de l'acide aminé était augmenté, plus les bactéries se répliquent et moins elles métabolisent la digoxine.

Dans les années qui ont suivies, des études ont été réalisées pour mieux comprendre cette interaction médicament-microbiote, mais avec les techniques de biologie moléculaire de l'époque une telle étude était difficile. Une étude intéressante est celle où les taux d'inactivation de la digoxine ont été comparés par provenance géographique des participants, avec une cohorte américaine et une cohorte d'Inde. De leur 95 participants, 67 étaient Américains, 23 d'Inde urbaine et 5 d'Inde rurale. Comme présenté à la Figure 8, les Américains ont un taux de production de DRPs beaucoup plus élevé, allant jusqu'à plus de 70% de la dose pour un participant. Les Indiens vivant en région rurale ont un taux beaucoup plus bas de métabolisme

et les Indiens urbains ont des taux intermédiaires entre les deux groupes (12). Les auteurs ont établi la corrélation entre la consommation de viande et le taux de transformation de la digoxine élevé, ce qui va presque en contradiction avec leur observation précédente de la corrélation négative entre le taux d'acide aminé arginine et le taux de réduction de la digoxine. De plus, ces recherches n'avaient pas permis d'établir une corrélation entre la présence de *E. lenta* et le niveau de réduction de la digoxine (11,12).

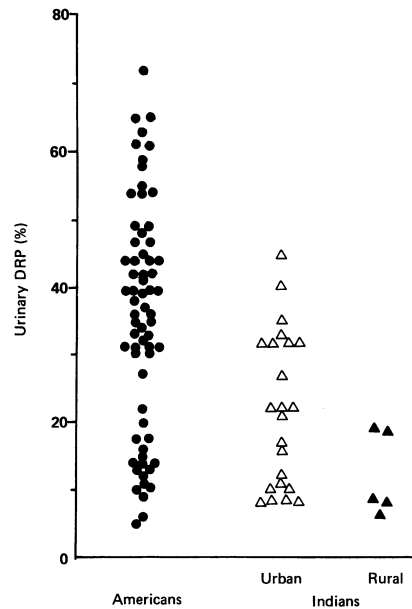


Fig. 1 Per cent digoxin reduction products (DRP) in the urine of 67 American metabolisers resident in New York City, 23 urban Indian metabolisers living in Vellore and five Indian metabolisers resident in rural villages 24 hours after an oral 0.25 mg dose of digoxin.

Figure 8. Niveau d'élimination de la digoxine sous forme de DRPs selon la provenance géographique des participants tiré de (12)

La raison suspectée pour le manque de corrélation entre la présence/absence de *E. lenta* et le niveau de métabolisme de la digoxine vient du fait que ce ne sont pas toutes les souches de la bactérie qui métabolisent la digoxine. En effet, avec l'avancement des technologies de séquençage, une équipe a récemment jeté un nouveau regard sur l'inactivation de la digoxine par *E. lenta*. En 2013, ils ont découvert deux gènes cytochromes dans la souche type, la souche DSM2243, mais qui était absent de deux autres souches qu'ils ont observées, les souches FAA 1-3-56 et FAA 1-1-60 (215). Ils ont réussi à démontrer que les souches étaient génétiquement différentes et que la souche DSM2243 possédait les flavocytochromes qu'ils ont nommés

l'opéron CGR (CardioGlycoside Reductase), composé des gènes CGR1 et CGR2. Ils ont stratifié avec succès la cohorte en établissant un ratio de l'opéron CGR sur la quantification 16S de toutes les souches de *E. lenta* présentes pour prédire le taux d'inactivation de la digoxine. Avec des qPCR (quantitative Polymerase Chain Reaction, une méthode de quantification des ARN messagers) ils ont démontré que les gènes de cet opéron sont activés lorsque la bactérie est exposée à la digoxine et que cette activité transcriptionnelle est inhibée par les hautes concentrations en arginine *in vitro*.

L'année suivante, ils ont publié une étude où ils décrivaient dans les détails moléculaires la structure des produits protéiques des gènes qu'ils avaient identifiés, CGR1 et CGR2. Durant le déroulement du présent projet de recherche, le gène CGR1 a été discontinué sur les bases de données du NCBI. Le gène CGR2 est celui qui conduit l'activité catalytique est une fumarate réductase liant le FAD (Flavine Adénine Dinucléotide, une coenzyme d'oxydo-réduction) et son action sur la digoxine est décrite comme 'par promiscuité de substrat' (209). En effet, le substrat naturel de cette enzyme est le fumarate, comme son nom l'indique, et l'anneau lactone du groupement stéroïde de la digoxine a des similarités structurelles avec ce dernier.

Les études qui ont suivi ont continué dans cette volonté de tenter d'inhiber l'activité métabolique que *Eggerthella lenta* exerce sur la digoxine, par exemple en proposant un plan de repas riche en arginine aux patients sous thérapie à la digoxine (216). Un autre point de vue pourrait être adopté où l'objectif est de ne pas perturber la bactérie pour garder la biodisponibilité de la digoxine stable. Généralement, les patients qui reçoivent de la digoxine sur une base quotidienne la reçoivent à long terme. Ainsi, lorsqu'ils doivent prendre des antibiotiques pour des infections non reliées à leur condition cardiaque, il est important que leur concentration plasmatique de digoxine reste stable pour éviter la toxicité. Cette considération pourrait s'appliquer à tous les patients, à savoir d'utiliser un antibiotique qui ne perturbe pas trop la bactérie métabolisatrice *E. lenta*, et tout particulièrement si les patients concernés sont des forts producteurs de DRPs. Ainsi, une approche de médecine personnalisée pour patients utilisant la digoxine pourrait proposer un séquençage du microbiome fécal pour quantifier les *E. lenta*



métabolisatrices de digoxine ou encore un dosage des DRPs urinaires. Les patients démontrant une forte activité de métabolisme par le microbiome pourraient ainsi avoir la recommandation d'antibiotiques les moins susceptibles de perturber *E. lenta* et conserver la biodisponibilité de la digoxine stable.

## 12 Objectifs

1. Caractériser et quantifier la bactérie *E. lenta* dans un jeu de données de métagénomique intestinale à l'aide d'approches bio-informatiques.
2. Évaluer le changement de concentration de la bactérie *E. lenta* chez des volontaires sains après 7 jours d'exposition à l'antibiotique céfprozil et à 90 jours suivant l'exposition.
3. Évaluer le changement de concentration de la bactérie *E. lenta* en culture en fonction de l'exposition à l'antibiotique céfoxitine.

**12.1 Dans le contexte de l'interaction pharmacomicrobiomique proposée, confirmer que la bactérie *E. lenta* est susceptible aux antibiotiques des jeux de données (céfprozil et céfoxitine).**

# 13 Méthodologie

## 13.1 Source des données

### 13.1.1 Présentation du jeu de données – expérience avec volontaires (*in vivo*)

Le présent projet a été réalisé avec la collaboration du Centre de Recherche en Infectiologie (CRI) affilié à l'Université Laval, à Québec. Sous la supervision de Dr. Jacques Corbeil, un projet de recherche a été mené avec comme méthodologie l'exposition de bactéries du microbiome à des antibiotiques de la famille des céphalosporines de seconde génération, à savoir le céfprozil (expérience *in vivo*; participants humains) et la céfoxitine (expérience *in vitro*). En quelques sortes, le but de leur projet de recherche était d'obtenir un éventail large de microbiomes fécaux et de cultures *in vitro* de ces microbiomes pour développer et confirmer le bon fonctionnement d'un outil informatique de pointe, un assembleur de séquences *de novo* (nous reviendrons sur ce type de logiciels au prochain chapitre) et sa suite de logiciels appelés Ray Meta (disponible à <http://denovoassembler.sourceforge.net/>). Avec une cohorte de 24 volontaires sains, ils ont administré du céfprozil à 18 des participants quotidiennement pour une durée de 7 jours consécutifs. Trois échantillons de selles ont été récupérés par les participants pour analyse de séquençage WGS (Whole Genome Sequencing) au laboratoire, le premier au jour initial (J0), le second après la semaine de traitement (J7) et le troisième après une période de trois mois de repos (J90). Avec ces données, ils ont publié un article dans lequel ils décrivent comment l'état initial du microbiome détermine sa restructuration après l'exposition à l'antibiotique, vers un état semblable à celui de la composition initiale du microbiome (217). L'exposition à l'antibiotique a été fait à raison de 500mg de céfprozil, deux fois par jour avec des comprimés P.O. que les participants ont pris eux-mêmes.

Le jeu de données associé au projet est disponible en ligne dans sa forme brute sur les serveurs du ENA (European Nucleotide Archive) sous le nom de projet PRJEB8094 (description et liens de téléchargement au <https://www.ebi.ac.uk/ena/data/view/PRJEB8094>). La description technique des données sera couverte plus bas. En plus de la conclusion relativement prévisible

citée plus tôt, quelques autres points intéressants sont ressortis suivant l'administration d'antibiotiques à des volontaires sains. La plus généralisée est l'augmentation de l'espèce *Lachnoclostridium bolteae* chez 16 des 18 participants exposés à l'antibiotique. Il y a aussi eu un sous-groupe de participants, semblables par leur microbiomes à diversité initiale basse et d'entérotipe bactéroïde, chez qui il y a eu une augmentation marquée du pathogène opportuniste *Enterobacter cloacae* suivant la prise du céfprozil.

### **13.1.2 Présentation du jeu de données – cultures *in vitro***

Dans un second volet d'expérimentations, un deuxième jeu de données (aussi d'une taille considérable) a été généré en cultivant les échantillons de matière fécale *in vitro* dans 4 différentes conditions. Les données relatives à cette phase du projet n'ont pas été soumises pour publication sous forme d'article, mais le design expérimental est très intéressant pour le présent projet. Les échantillons de microbiome fécal complet des 24 patients qui ont été collectés au jour initial J0 et après l'exposition au céfprozil au J7 ont été mis dans des milieux de culture complets. Les conditions de culture varient sur ces deux variables, soit la condition d'oxygène (avec oxygène ou sans oxygène) ainsi que l'ajout ou non de l'antibiotique céfoxitine (une autre céphalosporine de seconde génération. Ainsi les 48 échantillons (24 participants X 2 collectes) ont été mis en culture dans quatre conditions différentes, les paires de combinaisons de condition d'oxygène et de présence d'antibiotique, pour un total de 196 échantillons en culture. Les bactéries ont été cultivées *in vitro* pour une durée de 7 jours avant d'être traitées pour le séquençage en Whole Genome Sequencing (même technologie que pour les microbiomes fécaux dont le traitement a été décrit à la section précédente).

Nos collaborateurs du CRI ont initialement cultivé les échantillons de microbiome sous différentes conditions pour faire diverger les communautés bactériennes afin de pouvoir analyser un éventail plus large de composition microbiennes. Ces résultats de séquençage ont servi à entraîner et à confirmer la qualité des assemblages *de novo* du logiciel d'analyse informatique Ray Meta. Comme nous le verrons plus en détail dans les prochaines sections, la bactérie *Eggerthella lenta* a proliféré de façon significative dans toutes les conditions de culture

et a été plus fortement amplifiée dans certaines conditions comparativement à d'autres. Considérant les différentes pressions de sélection artificielles auxquelles les microbiomes sont soumis dans cette phase d'expérimentations, ce volet a été nommé expériences de 'sélectomique' lors de la génération du jeu de données.

### **13.1.3 Aspect technique des données**

Les échantillons ont été traités selon le protocole développé par le HMP (Human Microbiome Project) pour l'uniformisation des techniques de laboratoire et décrit dans la publication de Aagaard *et al.* en 2014 (218). L'ADN a été isolé puis séquencé sur des puces de séquençage de nouvelle génération produites par la compagnie Illumina. Les séquençages sur les échantillons de l'expérience avec les volontaires ont généré en moyenne 15 Gb par patient en format texte brut fastq (avec les 72 échantillons, c'est plus d'un terabyte). Pour les échantillons provenant des tests de culture en sélectomique, c'est une moyenne de 2,5 Gb par échantillon pour plus de 196 conditions de culture, pour un total d'environ 0.5 terabyte.

Comme décrit dans la publication de Raymond *et al.* en 2016, les échantillons ont été récupérés par les 24 participants selon les explications et avec les troussees fournies et acheminées au laboratoire en deux heures ou moins. Le transport a été fait dans un environnement anoxique grâce au dispositif GasPak EZ Anaerobe sachet (Becton, Dickinson and Company, Sparks, MD, USA). Au laboratoire, les échantillons ont été caractérisés selon l'échelle de Bristol, aliquotés dans plusieurs tubes et conservés dans un congélateur à -80 °C. L'extraction de l'ADN bactérien a été réalisé avec la trousse MO BIO PowerMax Soil DNA Extraction Kit (MO BIO Laboratories, Carlsbad, CA, USA), selon les recommandations du protocole de HMP, avec les étapes relativement standard de suspension dans un tampon de lyse avec des micro-billes, incubation et capture d'ADN et élution, toujours selon les directives de la trousse PowerMax (219). La qualité de l'extraction du matériel génétique du microbiome a été confirmée par trois techniques, soit l'appareil de spectrophotométrie NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA, USA), des gels d'agarose et un QuantiFluor ADNdb (ADN double brin) (Promega, Madison, WI, USA).

Les bibliothèques de séquençage ont été préparées en utilisant la trousse Nextera Sample Preparation (Illumina, San Diego, CA, USA) en utilisant 50 ng d'ADN purifié. La qualité de la préparation des bibliothèques de séquençage a été confirmée avec l'appareil Agilent Bioanalyzer 2100 high sensitivity DNA chips (Agilent Technologies, Santa Clara, CA, USA). Le séquençage a ensuite été mené avec un séquenceur HiSeq 1000 (Illumina) en utilisant une chimie v3 et des lectures en paires de 101 paires de bases (pb). Spécifiquement, cela signifie que les inserts (fragments d'ADN) d'une taille médiane de 600 pb ont été séquencés à chaque extrémité vers l'intérieur sur 101 pb. Pour certains de ces inserts très courts, les régions séquencées se chevauchent au centre, mais pour la majorité des inserts, une région centrale reste non séquencée. Ce point est important parce que dans le présent projet d'analyse, seuls les fragments lus en paires (du même insert) ont été sélectionnés afin d'augmenter la qualité des fragments analysés.

### **13.1.4 Base de données de génomes bactériens**

Pour le présent projet, il est nécessaire d'utiliser une base de données qui inclut le plus de génomes bactériens représentatifs des espèces bactériennes qui peuplent normalement le microbiome intestinal humain. Cette base de données servira à attribuer tous les fragments d'ADN séquencés à l'une ou l'autre espèce bactérienne pour quantifier quelle part des données appartient à quelle bactérie. Le plus d'espèces bactériennes présentes, la plus fine sera la résolution de cette étape d'alignement de séquences, alors que les séquences apparentées pourront être attribuées au bon génome. La base de données utilisée pour ce projet provient aussi de nos collaborateurs de l'Université Laval, alors qu'ils avaient regroupé 5 141 génomes bactériens dans une base de données pour le développement de leur aligneur, Ray Meta.

Malgré les nombreux projets publics de séquençage, il reste difficile de trouver une bonne base de données de génomes bactériens assemblés, identifiés convenablement et rassemblés en nombre assez important qu'ils représentent la diversité des espèces bactériennes

retrouvées dans l'un ou l'autre des microbiomes étudiés. Il existe de nombreux outils (cités plus tôt) pour l'identification et l'annotation des jeux de données de séquençage de la sous-unité 16S du ribosome, une technique très largement utilisée depuis plusieurs années. Les techniques de WGS gagnent en popularité récemment consécutivement avec le prix de plus en plus abordable des technologies de séquençage de nouvelle génération à haut débit. Il existe un bon nombre d'outils pour l'assurance qualité et l'inférence de fonctions de gènes sur des jeux de données de WGS. Il existe aussi plusieurs jeux de données de WGS de microbiome fécal humain, par exemple les données brutes du HMP qui sont disponibles en ligne pour téléchargement. Toutefois il n'y a pas de base de données de génomes bactériens assemblés et complets pour travailler localement.

## **13.2 Analyses bioinformatiques d'abondance de *E. lenta***

### **13.2.1 *De novo* vs. closed-reference OTU picking**

Le principe de la technique de WGS repose sur la fragmentation aléatoire des génomes bactériens et permet donc de séquencer des portions d'ADN génomique. Toutefois, ces fragments ne représentent pas des génomes complets et leurs positions ne sont pas connues. Il existe donc deux façons de traiter des données, soit l'assemblage *de novo* et la sélection d'unités taxonomiques opérationnelles (eng. OTU) en référence fermée (closed-reference OTU picking). La première technique constitue un de problèmes classiques de bioinformatique, c'est-à-dire l'assemblage de séquences par homologie des extrémités complémentaires. Sachant la direction du séquençage, les algorithmes programmés pour cette tâche vont donc examiner la fin d'un brin de séquençage et trouver le début d'un autre brin qui a le plus de similarité avec lui. Les brins séquencés sont ainsi reliés pour compiler des fragments d'ADN les plus longs possibles dans ce qui s'appelle un graphe de *de bruijn*, soit l'ensemble de ces liens. Les fragments ainsi reconstitués représentent les génomes bactériens, partiels ou complets, des bactéries présentes dans l'échantillon analysé. Le principal avantage de cette technique est que, sans ajout externe, les séquences contiguës (abrégé contigs) représentent des nouvelles séquences, c'est-à-dire que des contigs peuvent être des génomes de bactéries jusqu'à lors inconnues. Cela est possible parce

qu'il n'y a pas de génome ou de séquences connues *a priori* par le logiciel. Les contigs ainsi générés peuvent ensuite être comparés à des génomes bactériens connus pour les identifier et ceux qui restent inconnus sont (généralement) de génomes de nouvelles bactéries. Cette technique d'assemblage *de novo* est particulièrement intéressante lors de l'analyse d'écosystèmes peu connus, comme c'est le cas avec les échantillons d'extrêmophiles ou encore avec les échantillons du Earth Microbiome Project (EMP - [www.earthmicrobiome.org/](http://www.earthmicrobiome.org/) ), qui vise à identifier des nouveaux microorganismes dans des niches écologiques isolées.

La seconde technique de traitement des données de séquençage WGS est celle dite de référence fermée, c'est-à-dire que les lectures de séquençage sont comparées à un ensemble fermé de séquences de référence de génomes bactériens. Pour une raison évidente, cette technique ne permet pas de découvrir des séquences génomiques de bactéries inconnues, mais permet de quantifier efficacement les espèces présentes. Cette technique s'apparente aux techniques plus connues de génomique animale, pour lesquelles l'alignement des lectures de PCR est fait contre un génome eucaryote connu. La tâche est plus complexe avec le microbiome, alors qu'en génétique humaine, le gabarit d'alignement est déjà connu et assez bien caractérisé dans son ensemble, alors que le génome humain a été séquencé plusieurs fois dans son intégralité. Pour l'alignement de séquences bactériennes en référence fermée, le gabarit sur lequel les lectures de PCR sont alignés doit représenter un métagénome, qui provient donc d'une collection de génomes et qui ne sera jamais aussi défini que le génome humain. C'est pourquoi il est important d'avoir une base de données aussi étendue que possible pour augmenter la possibilité d'avoir le génome duquel provient le fragment de PCR et augmenter la résolution fine de cette attribution entre des espèces proches. Avec quelques autres avantages, cette technique nécessite aussi beaucoup moins de ressources computationnelles que la technique d'alignement *de novo*. Ainsi, la technique d'alignement en référence fermée permet de quantifier les espèces connues clairement et efficacement.

Comme nous venons de le voir, le whole genome shotgun (WGS) est une technique qui continuera à gagner en popularité, surtout avec la diminution des coûts associés aux techniques

de nouvelle génération. Parmi les deux techniques d'analyse de résultats de séquençage de ces lectures de PCR courtes, la technique d'assemblage *de novo* est toute désignée pour permettre la découverte de séquences génomiques jusqu'à lors inconnues. À l'inverse, la technique d'alignement à une référence fermée permet de quantifier rapidement les espèces déjà répertoriées dans un échantillon. C'est pour cette raison que dans la présente étude a, nous utilisons la seconde technique pour quantifier le plus précisément possible la bactérie déjà connue *Eggerthella lenta* (son génome est entièrement séquencé depuis plusieurs années) en alignant les lectures de PCR sur notre base de données complète de 5 141 génomes bactériens.

### 13.2.2 Comparaison des aligneurs D'ADN

Il existe plusieurs ressources de sources différentes quant aux ensembles logiciels pour attribuer les lectures de séquençage à des génomes de référence. Plusieurs outils et pipelines ont été comparés afin de s'assurer de la performance du présent pipeline. Parmi ceux-ci, nous avons comparé les résultats de quantification de *E. lenta* provenant de Ray Meta (l'outil d'alignement *de novo* développé à l'université Laval), l'aligneur BWA (Burrows-Wheeler Aligner) ainsi que deux métriques générés par l'outil RSEM (RNA-Seq by Expectation Maximization). Les résultats préliminaires obtenus avec chacune des techniques sont présentés dans les résultats supplémentaires.

Le premier outil, Ray Meta, est un aligneur *de novo*, c'est-à-dire que tous les fragments courts de PCR sont comparés et regroupés en séquences plus longues, sans génomes de référence pour influencer l'assemblage. Cette technique est donc une façon non biaisée d'étudier les génomes présents dans l'échantillon, même s'il s'agit d'un génome jusqu'à lors non répertorié. Les génomes ainsi reconstitués ont été redécoupés en k-mers (sous-séquences d'ADN de 21 ou 31 pb) *in silico* et chacun de ces fragments a été attribué à une espèce ou un genre. Lorsqu'une ambiguïté était observée, le k-mer analysé était attribué à un ordre taxonomique plus élevé (exemple, attribué au genre au lieu d'à l'espèce) et une table d'abondance des espèces a ainsi été générée (table disponible dans le matériel supplémentaire de (217) Raymond *et al.* 2016). Les



valeurs associées à *E. lenta* ont été extraites de la table d'abondance et comparées avec les valeurs générées avec les autres techniques.

L'aligneur BWA (Burrows-Wheeler Aligner) est un aligneur de lectures de PCR courtes contre une référence large comme le génome humain. Développé par le Wellcome Trust Sanger Institute, il comprend deux algorithmes adaptés pour les résultats de séquençage Illumina et après comparaison de BWA-backtrack et BWA-MEM, nous avons sélectionné le second (suivant les résultats préliminaires et les recommandations des développeurs). Le format de sortie de BWA est le format SAM (Sequence Alignment/Map), très largement utilisé en génomique et manipulable avec le logiciel ouvert SAMtools.

Nous avons aussi testé le pipeline de RSEM (RNA-Seq by Expectation Maximization), un pipeline d'alignement de lectures courtes. Cette suite logiciel a été sélectionnée parce que c'est un pipeline récent qui implémente Bowtie2, un autre aligneur d'ADN très populaire en bioinformatique. Il existe un parallèle entre les abondances des ARN messagers et la composition des microbiomes. Comme mentionné plus tôt, les gènes observés en études WGS et les ARNm bactériens observés sont hautement corrélés. Les profils d'abondances des copies de gènes bactériens sont donc corrélés avec les profils d'abondance des ARNm bactériens. Dans les études de métagénomique, les espèces bactériennes varient, mais les analyses fonctionnelles (quels gènes sont présents) révèlent des profils généralement conservés. Et un peu comme avec les cellules, les gènes les plus rencontrés sont les gènes "d'entretien cellulaire", comme les gènes de métabolisme et de synthèse protéique. Bref, comme les technologies de séquençage et les profils d'abondance des gènes bactériens séquencés peuvent avoir des similarités avec les ARNm, ce pipeline offre un code adapté pour faire une quantification des séquences courtes et les attribuer à une banque de séquences de référence. En plus des similitudes entre les deux types de données, le pipeline RSEM génère des résultats déjà normalisés et ne demande donc aucun traitement supplémentaire des données.

Bref, il n'existe pas, au meilleur de notre connaissance, un outil informatique optimisé pour quantifier une seule espèce bactérienne à partir de données de métagénomique WGS. En effet, la majorité des études déjà publiées sur le microbiome ont été réalisées avec des résultats de séquençage ciblé sur l'ARN mitochondrial de la sous-unité 16S. De plus, comme les coûts associés au séquençage WGS sont significativement plus élevés que ceux pour une campagne de séquençage 16S, la technologie WGS est surtout utilisée pour faire des études fonctionnelles des microbiomes, pour étudier les gènes présents et les clusters de gènes biosynthétiques (ensembles de gènes dont les extrants ont une importance économique ou de recherche). Lorsque l'on cherche à simplement identifier et quantifier les phyla présents, c'est majoritairement la technique 16S qui est privilégiée à cause de son faible coût pour le séquençage et le traitement des données. En effet, en plus d'être plus onéreux au niveau du laboratoire (réactifs, temps de traitement, etc.) les études WGS génèrent des données de taille considérable qui demandent des infrastructures informatiques puissantes en conséquence. Les WGS exigent presque de facto d'utiliser des super-ordinateurs, par exemple, simplement pour le premier jeu de données, c'est plus de 1Tb de séquences brutes (qui ont été doublées lors de la concaténation et autre étapes) et des dizaines et des dizaines de Gb de fichiers d'analyses; loin d'être envisageable sur un ordinateur de bureau.

### **13.2.3 Calcul des comptes normalisés**

Une fois les lectures courtes alignées avec BWA, la liste des lectures attribuées à *Eggerthella lenta* a été extraite pour chaque échantillon analysé. Les listes ont été inspectées sommairement visuellement, avec leurs informations. Plusieurs échantillons ont aussi été visualisés avec IGV (Integrative Genomics Viewer) du Broad Institute (209), permettant de voir directement les lectures de PCR alignées contre le génome de référence de *E. lenta*. Cette étape n'a toutefois pas servi à ajouter d'information, simplement visualiser les alignements.

Les lectures bien alignées ont aussi été sélectionnées selon les deux critères suivants, qui sont résumés par l'indicateur "SAM flag", un indicateur qui résume la qualité de l'alignement et d'autres informations. Toutes les lectures de PCR se voient individuellement attribuer un tel

indicateur qui utilise un système d'addition de bits. La valeur de l'indicateur est la somme de tous les bits informatiques et, comme deux bits ont tous des valeurs différentes, le chiffre total peut à tout moment être redécomposé en l'ensemble des indicateurs associés à une lecture de PCR. Il y a 12 caractéristiques avec deux états possibles et un SAM Flag et il le site <https://broadinstitute.github.io/picard/explain-flags.html> permet de d'équivaloir les valeurs des indicateurs avec les listes de caractéristiques des lectures de PCR. De plus, le fait d'utiliser l'indicateur '2' fait que le '1' est implicite. Le '1' spécifie que la lecture de PCR doit provenir d'une paire de lectures et le '2' que chacune des lectures de la paire ait été attribuée à la même séquence de référence. Le fait que les deux lectures aient été indépendamment attribuées à la même séquence de référence augmente par beaucoup la probabilité que le génome de référence choisi soit le bon. Le traitement des données demande de faire une transformation des fichier d'alignement en format ".sam" vers le format compact ".bam" (binary Alignment/Map) et lors de cette étape, le filtre '2' a été ajouté dans la commande "samtools view -f 2 -S" dans le but d'augmenter la fiabilité des alignements choisis.

### **13.2.4 Comparaisons de groupes non-paramétriques et modèles statistiques**

Les résultats de quantification de E. lenta n'ont pas produit des distributions normales, mais c'était attendu. En effet, le simple étendu des valeurs possibles (de 5 RPM à 4740 RPM dans les données des volontaires sains) empêche la formation de distributions simples, possiblement que c'est un reflet de la nature exponentielle de la croissance des bactéries. Durant une phase d'exploration des données, nous avons aussi tenté de modéliser ces distributions et le modèle statistique le plus ressemblant est celui de la bêta-régression. Les distributions observées sont généralement bimodales, avec un pic mineur d'abondance dans les valeurs plus élevées.

Bref, il apparaît clair que les données obtenues ne correspondent pas aux exigences des modèles statistiques paramétriques. Nous avons donc opté pour l'utilisation de l'équivalent nonparamétrique des tests t (de Student), soit le test de Wilcoxon–Mann–Whitney (parfois abrégé WMW) qui permet de comparer des groupes selon la positionnalité des valeurs, un peu comme la médiane est un équivalent non-paramétrique de la moyenne. Les comparaisons ont été réalisées en paires pour chacun des trois groupes de données provenant des volontaires qui

ont pris du cefprozil. Des corrections pour tests multiples auraient pu être réalisées étant donné que nous avons réalisé trois tests différents. Cela aurait fait passer le seuil de significance de  $p < 0.05$  à  $p < 0.0167$ , mais les valeurs de valeurs-p obtenues sont loin de ces deux seuils dans tous les cas.

Pour ce qui est des données provenant des expériences in vitro de sélectomique, une approche conservatrice a aussi été adoptée. En effet, dans les 192 échantillons mis en culture, il existe plusieurs groupes qui diffèrent uniquement sur leur statut sur la présence de céfoxitine. C'est à cause des différentes conditions de culture des microbiomes complets ainsi que les deux dates de provenance des échantillons mis en culture. Toutefois, au lieu de combiner tous les échantillons appartenant à chaque condition et les comparer sur leur statut d'exposition à l'antibiotique, nous avons simplement choisi le groupe le plus représentatif et analysé celui-ci uniquement. Le groupe choisi est celui où les cultures ont été faites en environnement anaérobique parce que cela représente plus fidèlement la condition de l'intestin. De plus, nous avons exclu les échantillons provenant des cultures réalisées à partir des échantillons de la seconde journée de collecte de spécimens de matière fécale. Cela a restreint à  $n=24$  échantillons, qui ne diffèrent que sur leur statut d'exposition à l'antibiotique.

Comme nous disposions de quelques données biométriques non-identifiantes des participants à l'étude, nous avons aussi généré quelques modèles statistiques plus complexes. Par exemple, nous avons élaboré un modèle avec une correction pour l'âge des participants sur les données d'abondance de *E. lenta*. Un autre modèle prenait en compte l'âge des participants et un autre modèle de régression linéaire généralisé (GLM) prenait en compte une combinaison des deux facteurs. Toutefois, ces modèles n'amélioreraient pas beaucoup le pouvoir prédictif des modèles développés et une revue de littérature approfondie a fait ressortir que, avec une taille d'échantillon aussi basse que  $n=18$ , il n'est pas prudent d'ajouter trop de paramètres dans les modèles statistiques.

# 14 Résultats

## 14.1 Article

### **Effect of two cephalosporins on the digoxin metabolizing gut microbiota *Eggerthella Lenta*.**

Jérémie Auger; Frédéric Raymond; Steve Geoffroy; Louis-Philippe Lemieux Perreault; Sylvie Provost; Simon de Denus; Jacques Corbeil; Marie-Pierre Dubé

- **Abstract:**

#### **Background**

The prevalent gut microbiota *Eggerthella (E.) lenta* metabolizes digoxin, a cardiotonic drug with a narrow therapeutic index. Daily digoxin dosage compensates for metabolism by *E. lenta*.

#### **Materials & Methods**

In this bioinformatics project, we investigated *E. lenta*'s susceptibility to two antibiotics from the cephalosporin family. We used two publicly available datasets where whole gut microbiome communities were exposed to cefprozil and cefoxitin, *in vivo* (intestine) and *in vitro* (cultures), respectively.

#### **Results and Conclusion**

We saw a sensitivity trend of the gut bacterium *in vivo* (cefprozil) and a statistically significant lowering of its abundance *in vitro* (cefoxitin),  $p=0.046$ . Whether the antibiotics could contribute to an increase in digoxin bioavailability in clinical use will need to be further explored.

- **Keywords:**

*Eggerthella lenta* (fmr. *Eubacterium lentum*); Digoxin; Cephalosporin; Gut Microbiome; Cardiology, Drug-Drug interaction (DDi); Pharmacology, Phramacomicrobiomics;

- **Main body of text**

## **Introduction**

Digoxin is widely used as a cardiotonic glycoside for the treatment of heart failure and atrial fibrillation. The bacterial specie *Eggerthella (E.) lenta* (strain DSM 2243) was shown to metabolize digoxin *in vivo* and *in vitro* [1, 2]. A clinical trial identified that 14% of digoxin users in the study had a "high metabolizer" profile and that for these digoxin patients, over 40% of the drug was reduced to an inactive metabolite before intestinal absorption. There are at least 24 known strains of the species *E. lenta* [3] and about half of these strains possess the genetic components involved in the metabolism of digoxin. Since the complete genomic sequences of all of these 24 strains are not fully curated online, we used a database including the 5-best known *Eggerthella lenta* strains, with the only metabolizing strain being *E. lenta* DSM 2243. Additional studies have also described the metabolism of digoxin by gut bacteria in patients [4-6], but none have yet quantified *E. lenta* using microbiome whole genome sequencing technologies before and after exposure to antibiotics.

Digoxin has a notoriously narrow therapeutic range and an unusually low therapeutic dosage, one of the only drugs with a plasmatic concentration that must be maintained in the low nanograms range (recommendation 0.5 ng/mL – 0.9 ng/mL of plasma [7]). Digoxin toxicity is one of the most frequently encountered adverse drug reaction [8] and can lead to life threatening adverse effects [9]. Hospitalizations for digoxin toxicity are often due to the addition of antibiotics to patient's drug regimens and interactions with macrolide antibiotics are frequent [10]. Here, we investigated our ability to detect *E. lenta* in gut microbiome samples from 18 healthy volunteers, before and after exposure to the cephalosporin cefprozil, using whole genome shotgun sequencing of microbiome specimens. A second dataset was created by *in vitro* culture of the specimens, with and without exposure to the cephalosporin cefoxitin, followed by whole genome shotgun sequencing. The results indicate that *E. lenta*'s abundance is reduced by a cephalosporin antibiotic *in vitro* and report a trend for a decrease with *in vivo* exposure. Whether the negative effects of cephalosporin antibiotics on *E. lenta* could contribute to an increase in digoxin bioavailability in clinical use will need to be further explored.

Recent progress in next-generation sequencing (NGS) allows for the complete characterization of the microbial communities in a stool sample, without the need for bacterial culture [11]. As much as 85% to 99% of bacteria and archaeobacteria cannot be grown by *in vitro* culture using general protocols and are therefore difficult to study by conventional microbiology methods [12, 13]. NGS techniques have paved the way for the emerging field of pharmacomicrobiomics, which investigates the roles of bacteria in the pharmacokinetics (PK) and the therapeutic response to drugs [14-16], most notably by studying the gut microbiome [17]. Recent studies have demonstrated the role of the microbiome in the efficacy of a wide range of medication including chemotherapeutic agents [18-21] and metformin [22-24]. Drugs administered intravenously can also be affected by the gut microbiome's metabolic activity, as the molecules can be secreted in the gut lumen in association with bile acid prior to reabsorption in the distal ileum and transport back to the liver via portal circulation [25].

Here, we revisited a dataset of 18 healthy volunteers before and after exposure to the cephalosporin cefprozil using whole genome shotgun sequencing of fecal samples to assess the *in vivo* bactericide effect of the antibiotic on *E. lenta*. We also present the results of *in vitro* cultures of the stool specimens of 24 healthy volunteers with and without exposure to another cephalosporin antibiotic, cefoxitin, followed by whole genome shotgun sequencing.

- **Methods**

### **Study participants**

Twenty-four healthy volunteers were recruited in the Quebec City region (Canada), were aged between 21 and 35 years and had normal intestinal transit. Subjects with any of the exclusion criteria below were not eligible for participation to the study:

- > Working in a health-care facility or living with someone working in a health-care facility
- > Working on a farm or household contact in the last 2 weeks
- > Slaughterhouse worker or household contact
- > Animal care worker or household contact
- > Chronic alcohol consumption (more than one 1.5-ounce servings of 80 proof distilled spirits, five 12-ounce servings of beer or five 5-ounce servings of wine per day)
- > Antibiotic therapy or history of hospitalization (>24 h) in the past 12 months prior to the study
- > Living with someone or an animal that has been on antibiotic therapy in the last month
- > Vegetarians
- > Smokers
- > Any gastrointestinal or underlying pathology
- > Any chronic illness
- > Any infection requiring chemo/ antibiotic therapy
- > Diarrhoeal disease (World Health Organization definition) in the last 3 months prior to the study
- > Gastro-intestinal-related medication (prescription antibiotics)
- > Immunomodulating medication such as antitumour necrosis factor or steroids
- > Allergy to  $\beta$ -lactams
- > Pregnant or lactating women
- > Taking alimentary supplements
- > Body mass index abnormal <18.5 or >30 kg/m<sup>2</sup>

The clinical protocol was conducted with the approval of the ethical committee of the CHU de Québec–Université Laval, with written and informed consent obtained from the participants prior to the start of the study.

### **Study Design**

From the 24 healthy subjects enrolled in the Raymond *et al.* study [26], 18 were given 500 mg of cefprozil twice daily over the course of seven days and 6 were controls. Fecal samples were collected at the day of enrolment (D0), after the week of cefprozil exposure (D7) and after a resting period (D90). Stool samples were self-collected by participants on D0, D7 and D90 according to instructions based on the human microbiome project protocols [27]. Stool specimen collection kits comprised one Ziploc bag (Johnson and Son, Brantford, ON, Canada) and one paper bag for transport, one GasPak EZ Anaerobe sachet (Becton, Dickinson and Company, Sparks, MD, USA), one M40 Transystem 408C transport swab (Copan Italia S.P.A., Brescia, Italy), one Commode Specimen Collection System (Biomedical Polymers, Inc., Gardner, MA, USA). All samples were brought to the laboratory within 2 h of collection and placed immediately in an anaerobic chamber for processing.

We also used data from the 2019 Raymond *et al.* study [28] a secondary dataset derived from samples from the first two sampling days (D0 and D7) that were each divided into four different culture conditions, for a total of 192 test samples (4 conditions X 24 patients X 2 sampling times). Briefly, 1g of fecal matter served for inoculation of the cultures grown in MCDA broth [29]. The 4 culture conditions are pairs of these two variables: *i*) aerobic (supplemented with 5% CO<sub>2</sub>) vs. anaerobic, *ii*) No antibiotic vs. 32µg/ml of ceftiofuran. This design yields 4 the groups (as shown on figure 2) Aerobic-None; Aerobic-Ceftiofuran; Anaerobic-None; Anaerobic-Ceftiofuran. All the samples were incubated at 35°C for 7 days.

### **DNA extraction and sequencing**

The raw genomic reads are available in the European Nucleotide Archive (EBI) database under the project number PRJEB8094 and sequencing methodology is described in [26]. Briefly, bacterial genomic DNA was extracted from 500 mg aliquots of frozen stool samples, sequencing libraries were prepared using the Nextera Sample Preparation Kit (Illumina, San Diego, CA, USA) and sequencing was performed on the HiSeq 1000 sequencer (Illumina) using v3 chemistry and paired-end 101 bp reads. A total of 72 microbiome samples were sequenced with an average of 15 Gb. A total of 192 *in vitro* samples [28] were sequenced at lower depth with an average of 2.5 Gb of raw fastq sequence per sample.

### **Bioinformatics analytical Methods**

Here, we used the BWA (Burrows-Wheeler Aligner) alignment tool to align reads to a database of 5,141 bacterial genomes that represent the diversity of normal human gut microbiomes. Since the complete genomic sequences of all 24 previously reported *E. lenta* strains were not fully curated online at the time of analysis, we included in the reference database the 5-best characterized *E. lenta* strains (DSM 2243; lenta-84112; sp. 1\_3\_56FAA-665943; sp. YY7918-502558; sp. HGA1-910311), of which the only strain known to metabolize digoxin is DSM 2243. We included closely related bacterial species in the reference database, including *Slackia* spp., *Gordonibacter* spp. and



*Adlercreutzia equolifaciens* to increase accuracy in reads mapping to *E. lenta*. The reads attributed to *E. lenta* DSM 2243 with a high level of confidence were extracted and counted using closed-reference OTU picking. Only reads mapped in proper pairs on *E. lenta* DSM 2243 genome, with a minimal match of 90 nucleotides out of the 101 nucleotides full-length sequencing reads, were considered specific to this strain. The read counts were then normalized by the total number of reads per sample and expressed as number of reads per million (rpm).

The relative abundances of *E. lenta* were compared between sampling times for the 18 participants exposed to cefprozil. Data from each sampling day (D0, D7 and D90) were compared pairwise using the Wilcoxon signed-rank test (3 paired comparison tests) using the R statistical package. The p-values computed with these tests are shown on Figure 1 for the three time intervals. We plotted the data using Raw Description and Inference plots (RDI) showing median and high-density intervals computed using a Bayesian algorithm. As illustrated in the Supplementary Figures S1 and S2, there was one extreme outlier (P21) observed on D90 with values of 1,186 rpm, more than 6 times greater than the second highest value. Anecdotally, this volunteer reported using iron supplements and had to be excluded. The calculated minimally detectable mean of paired differences with 80% power using the Wilcoxon test for change in *E. lenta* rpm from D0 to D7 was estimated as a decrease of 30.2 rpm at  $P=0.05$ . The fold changes were calculated as  $F_c = \log_2(B/A)$  where "A" is the initial condition (before receiving or without antibiotics) and "B" is the treatment. The ratio (B/A) itself illustrates an augmentation when above 1 and a diminution when between 0 and 1. The  $\log_2(B/A)$  transformation allows to calculate the doubling of the microbiota's counts (a  $F_c$  of 1 is equal to a doubling of the counts and a  $F_c$  of -1 is equal to a halving of the counts) and makes the fold changes values symmetrical around zero.

We used the 'pirateplot' library implemented in R <https://github.com/ndphillips/yarr> to plot the calculated *E. lenta* relative abundances. This representation shows raw data points, a zone of high-density interval (a robust equivalent to the error bars) and an outline of the distribution (RDI plot: Raw data, Descriptive and Inferential statistics).

For the *in vitro* culture experiments, similar alignment and statistical approaches were used. Here, we report D0 samples cultured in an anaerobic environment ( $n = 24$ ) with and without ceftiofur, which were compared with a Wilcoxon signed-rank test.

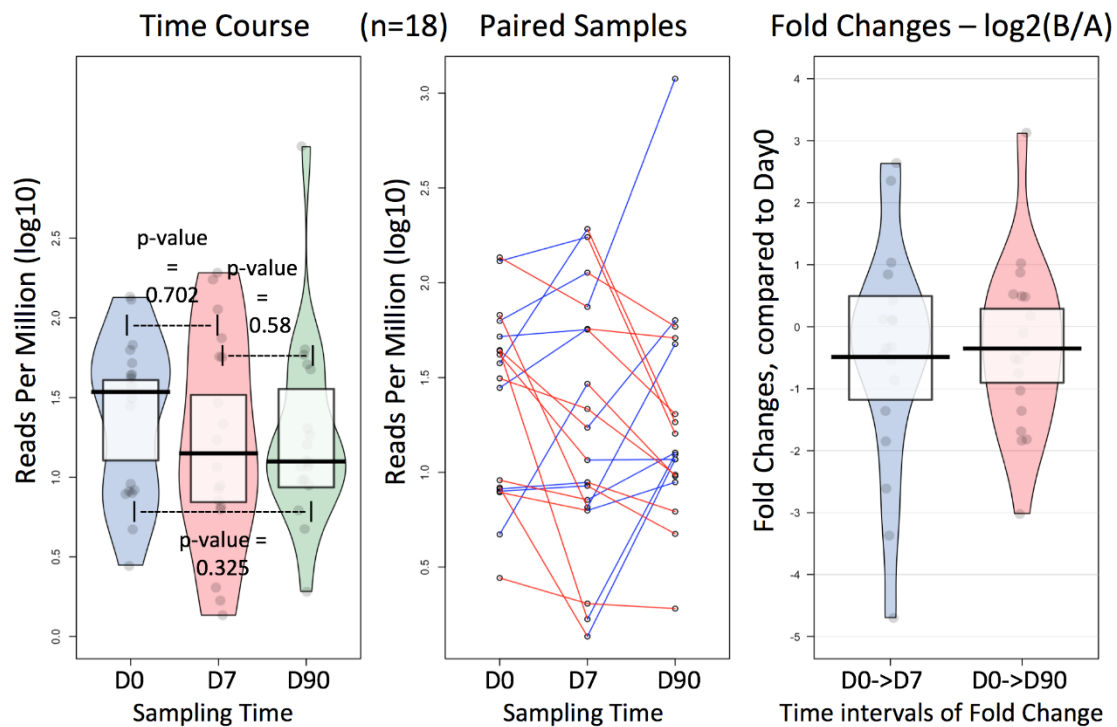
- **Results**

### **Healthy volunteers, oral cefprozil and gut microbiome**

At initial sampling day (D0) for the 18 volunteers in the cefprozil group the median number of reads mapping to *E. lenta* was 34.5 reads per million (rpm) (min. 2.77; max. 136), after twice daily 500mg of cefprozil exposure (D7), the median number of reads was 14.4 rpm (min. 1.35; max. 192), and after 90 days following study initiation (83 days exposure free) (D90), the median number of reads is 12.5 rpm (min. 1.90; max. 1185). The results are summarized in Table 1 and are graphically represented in Figure 1.

**Table 1. Summary of WGS reads mapping to *E. lenta*.**

<i>Stool samples from 18 healthy volunteers prior and after exposure to cefprozil</i>						
	<b>Reads per million mapped to <i>E. lenta</i></b>					
<b>Sampling visit</b>	<b>N</b>	<b>Median</b>	<b>Min</b>	<b>Max</b>	<b>Mean</b>	<b>SD</b>
Study initiation, prior to exposure (D0)	18	34.5	2.77	136	40.2	39.6
After 7 days of exposure to cefprozil (D7)	18	14.4	1.35	192	43.9	59.1
After 90 days post-study initiation (D90)	18	12.5	1.90	1185	86.1	275
<b>Paired sample differences</b>						
Change at D7 (D7-D0)	18	-1.2	-61.8	155	3.6	48.9
Change at D90 (D90-D0)	18	-2.6	-114	1049	45.9	252
% change at D7	18	-3.3	-179	449	10.4	142
% change at D90	18	-7.5	-331	3045	133	731
<i>Culture of stool samples from healthy volunteers</i>						
<b>Sampling visit, culture environment</b>	<b>N</b>	<b>Median</b>	<b>Min</b>	<b>Max</b>	<b>Mean</b>	<b>SD</b>
D0 samples, aerobic culture without cefoxitin	24	4,113	35.7	67,830	8,027	13,861
D0 samples, aerobic culture with cefoxitin	24	5,850	3.23	77,944	9,556	15,829
D0 samples, anaerobic culture without cefoxitin	24	8,776	56.8	64,371	14,201	15,734
D0 samples, anaerobic culture with cefoxitin	24	4,970	68.7	48,770	7,719	9,873
D7 samples, aerobic culture without cefoxitin	18	5,560	4.59	28,186	8,523	8,991
D7 samples, aerobic culture with cefoxitin	18	2,856	3.11	16,244	4,704	4,921
D7 samples, anaerobic culture without cefoxitin	18	20,458	79.7	120,654	24,751	26,928
D7 samples, anaerobic culture with cefoxitin	18	5,112	14.1	87,943	11,851	18,891
<b>Paired sample differences with and without cefoxitin (diff. Cefoxitin - None)</b>						
	<b>Differences in reads per million mapped to <i>E. lenta</i></b>					
<b>Sample origin, condition</b>	<b>N</b>	<b>Median</b>	<b>Min</b>	<b>Max</b>	<b>Mean</b>	<b>SD</b>
D0 samples, aerobic culture	24	-145	-64,686	73,727	1,529	21,319
D0 samples, anaerobic culture	24	-1,203	-59,213	43,834	-6,481	18,910
D7 samples, aerobic culture	18	-1,065	-23,622	7,952	-3,819	8,412
D7 samples, anaerobic culture	18	-11,355	-90,323	84,287	-12,899	30,466
<b>Percentage of change, compared to no ATB condition</b>						
D0 samples, aerobic culture (% change)	24	-3.5	-1573	1792	37.2	518
D0 samples, anaerobic culture (% change)	24	-13.7	-675	499	-73.8	215
D7 samples, aerobic culture (% change)	18	-19.1	-425	143	-68.7	151
D7 samples, anaerobic culture (% change)	18	-55.5	-441	412	-63.0	149



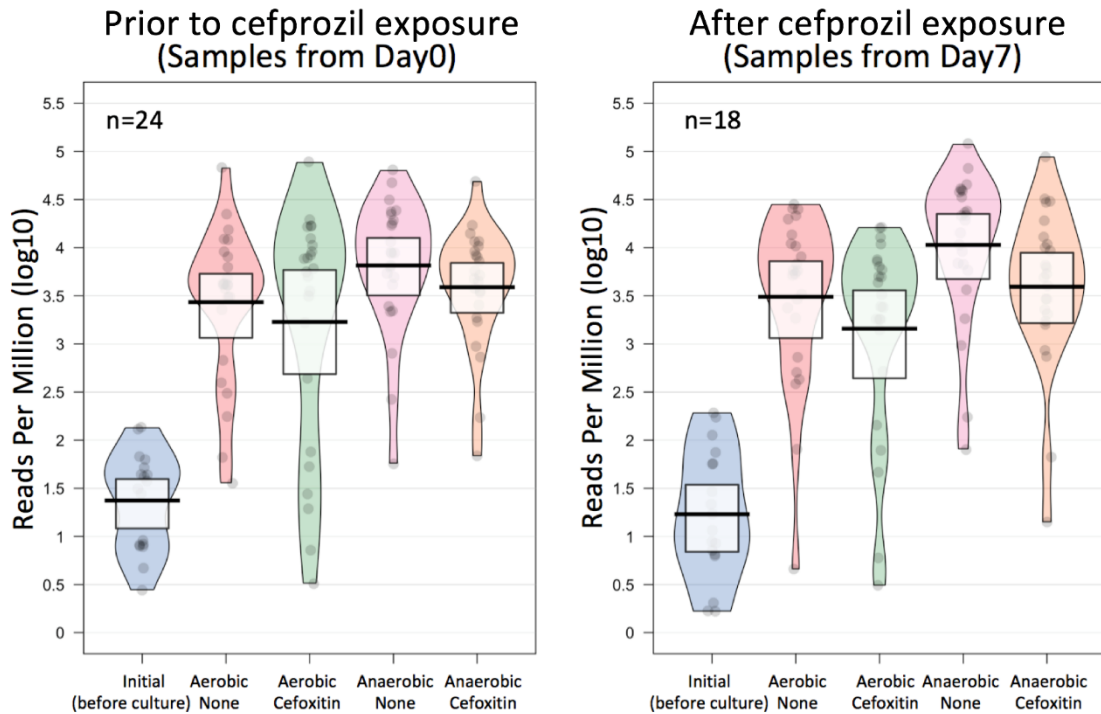
**Figure 1. Raw description and inference graph showing *E. lenta* WGS reads and fold changes for the oral cefprozil experimental data**

The median value at day 7 was 20.1 rpm smaller compared to the one at day 0, with a median decrease of 1.15 rpm from D0 to D7 in the number of reads attributed to *E. lenta*. At day 90, the median number of reads attributed to *E. lenta* remained low, with a median decrease of 2.6 rpm compared with day 0. These values, however, did not produce statistically significant differences, with  $P = 0.702$  for changes from D0 to D7, and  $P = 0.325$  for changes from D0 to D90 (Wilcoxon rank-sum test).

The right most panel in Figure 1 illustrates the fold changes ( $F_c$ ) in the counts of *E. lenta* over time for each individual, with the first interval (noted D0->D7) showing the changes in the initial time interval (before and after exposure to cefprozil) and the second interval (noted D0->D90) showing the changes between the initial levels and the final levels. The reads mapping to *E. lenta* generally decreased following exposure to cefprozil, with a calculated  $F_c = -0.33$  (min: -4.70; max: 2.64; sd: 1.83), which corresponds to a diminution in the counts by one quarter. The second  $F_c$  time intervals shows that the *E. lenta* read counts remain low even after 90 days with  $F_c = -0.45$  (min: -3.02; max: 3.13; sd: 1.39).

### Whole microbiome *in vitro* culture experiment, with and without cefoxitin

*In vitro* cultures were generated from stool samples taken from 24 healthy volunteers on day 0 (D0) and on day 7 (D7) that were grown in both aerobic and anaerobic environments and in the presence or absence of cefoxitin (Table 1, bottom section of table). The median number of reads mapping to *E. lenta* in the 24 samples prior to culture was 34.5 rpm (min. 2.77; max. 136), was 8776 rpm (min. 57; max. 64,371) when grown without cefoxitin in an anaerobic culture and 4970 rpm (min. 69; max. 48,770) with cefoxitin (anaerobic).

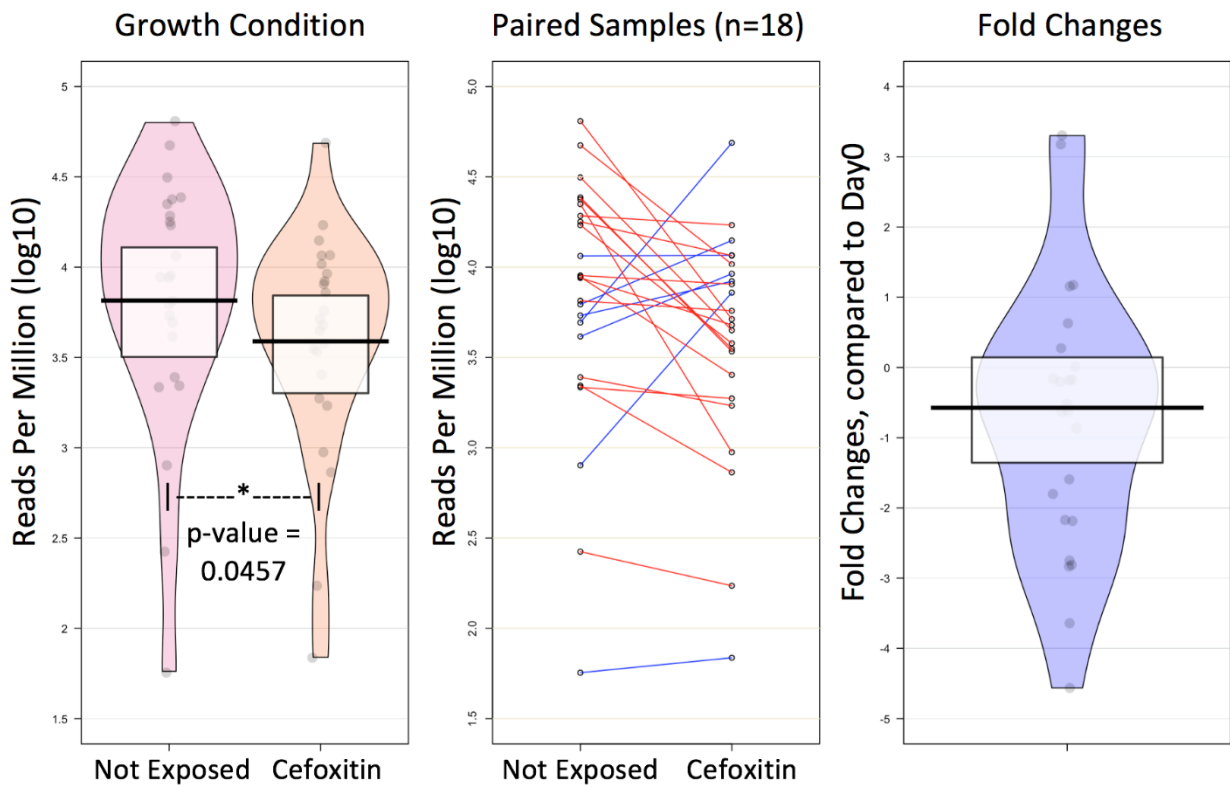


**Figure 2. Raw description and inference graph showing *E. lenta* WGS reads from the *in vitro* experiments grown in an aerobic and anaerobic environment with and without cefoxitin**

Figure 2 shows results from the culture of stool samples taken from 24 healthy volunteers unexposed to antibiotics after 7 days of culture in anaerobic conditions with and without cefoxitin. As emphasized on figure 3, the median decrease in reads mapping to *E. lenta* in the presence of cefoxitin was -1203 rpm ( $P= 0.046$ ) for the 24 samples from day 0 (no previous exposure to antibiotics) and grown in an anaerobic environment.

**Table 2: Summary of all the calculated fold changes.**

		Fold changes [ $\log_2(B/A)$ ]				
<i>Stool samples from 18 healthy volunteers before and after exposure to cefprozil</i>						
Time interval	N	Median	Min	Max	Mean	SD
D0 -> D7	18	-0.33	-4.70	2.64	-0.49	1.83
D0 -> D90	18	-0.45	-3.02	3.13	-0.35	1.39
D7 -> D90	18	-0.12	-3.44	4.00	0.14	2.14
<i>Samples from in vitro cultures</i>						
D0 samples, aerobic culture	24	-0.19	-11.06	4.21	-0.68	3.20
D0 samples, anaerobic culture	24	-0.57	-4.56	3.30	-0.75	1.93
D7 samples, aerobic culture	18	-1.19	-6.16	4.22	-1.10	2.18
D7 samples, anaerobic culture	18	-1.31	-5.67	4.59	-1.45	2.07



**Figure 3. Raw description and inference graph showing *E. lenta* WGS reads and fold changes for the *in vitro* experiments**

## Discussion

We have revisited data generated in the 2016 Raymond *et al.* study [26] which included whole genome shotgun sequencing of stool samples from 18 healthy volunteers before and after exposition to the cephalosporin cefprozil. We assessed the *in vivo* bactericide effect of the antibiotic on *E. lenta* as it could play an important role in explaining how antibiotics interfere with the pharmacokinetics of digoxin, by using data from the 2018 Raymond *et al.* study [28]. We saw a non-statistically significant reduction in the median abundance of *E. lenta*, with a median loss of 3.3% of rpm at D7 (Fc= -0.33). We also performed reanalysis of the *in vitro* dataset generated by culturing the stool specimens of 24 healthy volunteers in both aerobic and anaerobic environments with and without exposure to the cephalosporin antibiotic cefoxitin, followed by whole genome shotgun sequencing. Abundance of reads mapping to *E. lenta* increased several hundred folds in the culture environment and enabled to gain precision on the relative read estimates in the samples. We saw a significant decrease in *E. lenta* abundance in the samples grown with compared to without cefoxitin (P= 0.0457) taken at D0 in anaerobic culture.

The demonstration of *E. lenta*'s antibiotic susceptibility in the context of complete microbiome environment supports the hypothesis of a possible microbiome-mediated effect of antibiotics on digoxin. Here, sequencing reads mapping to *E. lenta* from *in vivo* specimens did not reach statistical significance, which is likely explained by the small sample size (n = 18) and the high variability of *E. lenta* read counts in the raw data, ranging from 1 rpm to 1,186 rpm. In addition, a larger sample size may have enabled to assess the impact of initial enterotypes based on prevotella spp. to bacteroides spp. ratio (P/B ratio) [30] and see whether the effect of antibiotics on *E. lenta* is modulated by the subject's initial microbiome composition.

An important limitation of the present project is that the experimental design for the study samples was not specifically developed to test the microbiome mediation of the effect of antibiotics on digoxin, as none of the volunteers received digoxin. It would be of interest in the future to correlate changes in *E. lenta* abundance with changes in the plasmatic concentration of the cardioglycoside drug. Nonetheless, the conclusions reached here provide evidence in support of effects of antibiotics on *E. lenta* abundance and offer support for the conduct of such studies in digoxin treated patients in the future.

Digoxin and its metabolites are measurable in the urine of patients, with a usual 50%-70% of the digoxin dose excreted unchanged [31]. When metabolized by *E. lenta*, digoxin's lactone ring is reduced and the drug loses its affinity for its substrate and is rapidly eliminated as DRPs (Digoxin Reduced Products) [2, 4]. A change in the ratio of digoxin excreted in its native form vs. DRPs is a good indicator of an alteration of the pharmacokinetics of digoxin mediated by a shift in the gut microbiota. Monitoring this ratio will help determine the cause of the plasmatic accumulation of digoxin.

Our *in vitro* cultures demonstrate *E. lenta*'s ability to grow in culture and probably occupy the space that other unculturable microbiota lose in the culture medium. The 24 cefoxitin pairs grown in an anaerobic environment had significantly lowered *E. lenta* counts in the antibiotic cultures, with a p-value of 0.0457 computed with a Wilcoxon-Mann Whitney test. We noted that previous exposure to a cephalosporin did not seem to impact the growth pattern of *E. lenta* (data not shown). To follow up on these findings, it would be interesting to compare the effect of different classes of antibiotics on *E. lenta*'s counts. Clinical studies indicate that macrolides are more likely to cause digoxin toxicity in patients [32] which could be supported by *in vivo* or *in vitro* tests.

## **Conclusion**

The present study offers a new look at a dataset that was not intended to demonstrate the hypothesis of the gut microbiome's implication in the pharmacokinetic of the heart failure drug digoxin. Trends seen *in vivo*, are supported by our *in vitro* experiments and indicate a decrease in the relative abundance of *E. lenta* when exposed to cephalosporin antibiotics. Together with the prior knowledge of the ability of *E. lenta* to metabolize digoxin, our results support the hypothesis of a possible microbiome-mediation of the effect of antibiotics on digoxin bioavailability.

## **Acknowledgments**

JA was supported by a scholarship from the Montreal Heart Institute Foundation. This research was enabled in part by computing infrastructures at Calcul Québec ([www.calculquebec.ca](http://www.calculquebec.ca)) and Compute Canada ([www.computeCanada.ca](http://www.computeCanada.ca)). FR is associated to the Canada Research Excellence Chair in the Microbiome-Endocannabinoidome Axis in Metabolic Health. Université Laval et Centre Recherche en Infectiologie

- **Future Perspective:**

The field of pharmacomicrobiomics is an emerging one, but promises to impose itself as a paradigm shift in personalized medicine and drug development in the coming years. In the light of all the metabolic pathways and the millions of genes that govern them, it is becoming clearer and clearer that the different human microbiomes affect the delivery and the efficiency of all drugs, as even intravenously administered drugs interact with the gut microbiome. With current knowledge, there are even cases (like with metformin) where the microbiome is the actual target of the drug. Hence, the study of microbiome's composition and functions promises to shed light on why certain drugs work for certain subsets of patients but not others. This will certainly help in the development of new drugs and the understanding of the mechanisms of action and/or of toxicity of already existing drugs.

- **Practice Points:**

1. Digoxin is widely used to control Atrial Fibrillation and Heart Failure.
2. Digoxin has a very narrow therapeutic range. Digoxin toxicity is a major cause of hospitalization in the western world.
3. Digoxin toxicity clinical manifestation: vomiting, confusion, blurry vision or altered colors that can lead to heart dysrhythmia (too fast or too slow).
4. Although falling out of popularity, digoxin is still widely used because withdrawal from the cardiotonic is associated with decompensation.
5. Antibiotics are known to interact with digoxin and modify its pharmacokinetics. These interactions often result in digoxin toxicity and hospitalizations.
6. In toxicity cases, care involves supportive measures and severe cases can be treated with digoxin-specific antibodies (Digibind).
7. Hypokalemia exacerbates digoxin toxicity.
8. Further research is needed to determine which antibiotics affect less *E. lenta*, the goal being to keep the cardiotonic's pharmacokinetics unchanged. When prescribing antibiotics for a patient treated with digoxin, the current clinical guidelines promote the use of cephalosporins (ex. cefprozil or cefoxitin) over antibiotics from the macrolide family (macrolide are associated with higher toxicity risk).



- **Figure legends**

**Figure 1. Raw description and inference graph showing *E. lenta* WGS reads and fold changes for the oral cefprozil experimental data.** The left most panel shows *E. lenta* WGS reads at the three sampling times (Day 0, Day 7 and Day 90) in 18 healthy volunteers exposed to 500 mg cefprozil twice daily from days 1 to 7. The central panel shows the pairwise changes in abundance of *E. lenta* reads (red = diminution; blue = augmentation) and the rightmost panel depicts the log<sub>2</sub> fold changes from D0 (negative = diminution, positive = augmentation). P values are from Wilcoxon paired test. The black line represents the median within a box delimiting high density intervals.

**Figure 2. Raw description and inference graph showing *E. lenta* WGS reads from the *in vitro* experiments grown in an aerobic and anaerobic environment with and without cefoxitin.** The 'initial' samples (illustrated in blue in both panels) represent the rpm values for the samples taken at D0 before the culture protocol. The black line represents the median within a box delimiting high density intervals (computed with the Bayesian algorithm).

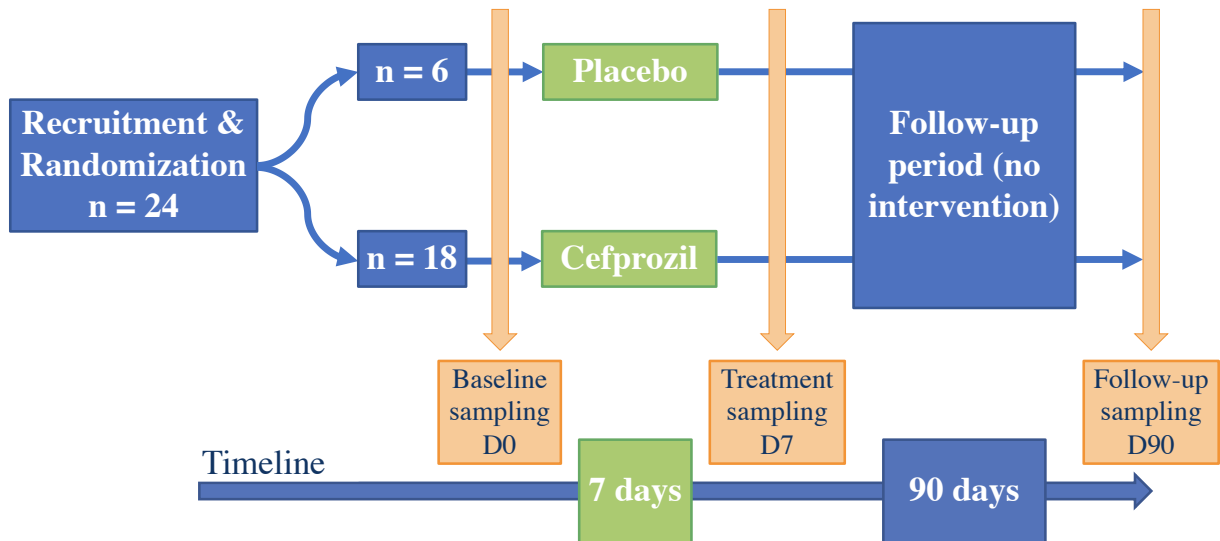
**Figure 3. Raw description and inference graph showing *E. lenta* WGS reads and fold changes for the *in vitro* experiments.** The left most panel shows the distributions for the 24 stool samples taken prior to cefprozil exposure grown in an anaerobic environment without (labelled 'Not exposed') and with (labelled 'Cefoxitin') the antibiotic cefoxitin. The central panel shows the pairwise change in abundance of *E. lenta* (red = diminution; blue = augmentation) and the rightmost panel depicts the log<sub>2</sub> fold changes. The p-value is from a Wilcoxon paired test. The black line represents the median within a box delimiting high density intervals.

- **Table Legends**

**Table 1. Summary of WGS reads mapping to *E. lenta*.** The leftmost column is a list of all the conditions for quantifying microbiota of interest. The first main horizontal section (“Stool samples...”) comprises the raw results *E. lenta*’s quantification from microbiome samples, with a second section describing the changes in these counts (Paired sample difference). For most descriptive and statistical purposes, the second data column (median) is most referred to. The second main section (“Cultures of stool...”) contains the raw numbers for the quantification of *E. lenta* following the *in vitro* cultures in different conditions. The third main section (“Paired sample differences...”) summarizes the calculations for comparing the cultures grown antibiotics-free against the cefoxitin cultures.

**Table 2: Summary of all the calculated fold changes** for the time intervals of the clinical (top section) and for the paired *in vitro* cultures diverging on the cefoxitin absence / presence (bottom section). All the fold changes have been calculated with the formula  $F_c = \log_2(B/A)$ , with either T1=A and T2=B (top half); or Atb-free culture=A and Cefoxitin culture = B (bottom half).

- **Supplementary Material:**



**Figure S-1. Clinical design used for the assessment of the effect of cefoxitin on the gut microbiome of 24 healthy participants.** Adapted from the methods of Raymond *et al.* 2016 [26]. At enrolment, participants were randomized in either Placebo group (n = 6) on cefprozil group (n = 18; 500 mg daily dose). Three stool samples were collected and analyzed through a WGS (Whole Genome Shotgun sequencing) analysis pipeline.

- **References:**

1. Saha JR, Butler VP, Neu HC, Lindenbaum J. Digoxin-inactivating bacteria: identification in human gut flora. *Science (New York, N.Y.)* 220 325-327 (1983).
2. Mathan VI, Wiederman J, Dobkin JF, Lindenbaum J. Geographic differences in digoxin inactivation, a metabolic activity of the human anaerobic gut flora. *Gut* 30 97-91 (1989).
3. Bisanz JE, Soto-Perez P, Lam KN *et al.* Illuminating the microbiome's dark matter: a functional genomic toolkit for the study of human gut Actinobacteria. *bioRxiv* 304840 (2018).
4. Lindenbaum J, Rund DG, Butler VP, Tse-Eng D, Saha JR. Inactivation of Digoxin by the Gut Flora: Reversal by Antibiotic Therapy. *New England Journal of Medicine* 305 789-794 (1981).
5. Koppel N, Bisanz JE, Pandelia M-E, Turnbaugh PJ, Balskus EP. Discovery and characterization of a prevalent human gut bacterial enzyme sufficient for the inactivation of a family of plant toxins. *eLife* 7 (2018).
6. Alkadi HO, Nooman MA, Raja'a YA. Effect of gentamicin on serum digoxin level in patients with congestive heart failure. *Pharmacy World and Science* 26(2), 107-109 (2004).
7. Goldberger ZD, Goldberger AL. Therapeutic ranges of serum digoxin concentrations in patients with heart failure. *The American journal of cardiology* 109 1818-1821 (2012).
8. Howland RD, Mycek MJ, Harvey RA, Champe PC. *Lippincott's illustrated reviews: Pharmacology*. Lippincott Williams & Wilkins Philadelphia, (2006).
9. Pincus M. Management of digoxin toxicity. *Australian prescriber* 39(1), 18 (2016).
10. Digoxin: serious drug interactions. *Prescrire international* 19 68-70 (2010).
11. Franzosa EA, Hsu T, Sirota-Madi A *et al.* Sequencing and beyond: integrating molecular 'omics' for microbial community profiling. *Nature reviews. Microbiology* 13 360-372 (2015).
12. Lok C. Mining the microbial dark matter. *Nature* 522 270-273 (2015).
13. Solden L, Lloyd K, Wrighton K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Current Opinion in Microbiology* 31 217-226 (2016).
14. Belizário JE, Napolitano M. Human microbiomes and their roles in dysbiosis, common diseases, and novel therapeutic approaches. *Frontiers in microbiology* 6 1050 (2015).
15. Elrakaiby M, Dutilh BE, Rizkallah MR, Boleij A, Cole JN, Aziz RK. Pharmacomicrobiomics: The Impact of Human Microbiome Variations on Systems Pharmacology and Personalized Therapeutics. *OMICS: A Journal of Integrative Biology* 18 402-414 (2014).
16. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nature Reviews Microbiology* 14 273-287 (2016).
17. Saad R, Rizkallah MR, Aziz RK. Gut Pharmacomicrobiomics: the tip of an iceberg of complex interactions between drugs and gut-associated microbes. *Gut Pathogens* 4 16 (2012).
18. Alexander JL, Wilson ID, Teare J, Marchesi JR, Nicholson JK, Kinross JM. Gut microbiota modulation of chemotherapy efficacy and toxicity. *Nature Reviews Gastroenterology & Hepatology* 14(6), 356 (2017).
19. Geller LT, Barzily-Rokni M, Danino T *et al.* Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science (New York, N.Y.)* 357 1156-1160 (2017).
20. Lehouritis P, Cummins J, Stanton M *et al.* Local bacteria affect the efficacy of chemotherapeutic drugs. *Scientific reports* 5 14554 (2015).
21. Takasuna K, Hagiwara T, Hirohashi M *et al.* Involvement of beta-glucuronidase in intestinal microflora in the intestinal toxicity of the antitumor camptothecin derivative irinotecan hydrochloride (CPT-11) in rats. *Cancer research* 56 3752-3757 (1996).

22. Bradley CA. Gut microbiota: Trust your gut — metformin and diabetes. *Nature Reviews Endocrinology* 13 440-440 (2017).
23. Buse JB, DeFronzo RA, Rosenstock J *et al.* The Primary Glucose-Lowering Effect of Metformin Resides in the Gut, Not the Circulation. Results From Short-term Pharmacokinetic and 12-Week Dose-Ranging Studies. *Diabetes Care* 39 dc150488 (2015).
24. Maniar K, Moideen A, Bhattacharyya R, Banerjee D. Metformin exerts anti-obesity effect via gut microbiome modulation in prediabetics: A hypothesis. *Medical Hypotheses* 104 117-120 (2017).
25. Wilson ID, Nicholson JK. Gut microbiome interactions with drug metabolism, efficacy, and toxicity. *Translational Research* 179 204-222 (2017).
26. Raymond F, Ouameur AA, Déraspe M *et al.* The initial state of the human gut microbiome determines its reshaping by antibiotics. *The ISME journal* 10(3), 707 (2016).
27. Aagaard K, Petrosino J, Keitel W *et al.* The Human Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. *The FASEB Journal* 27 1012-1022 (2013).
28. Raymond F, Boissinot M, Ouameur AA *et al.* Culture-enriched human gut microbiomes reveal core and accessory resistance genes. *Microbiome* 7(1), 56 (2019).
29. Domingo M-C, Huletsky A, Giroux R, Picard FJ, Bergeron MG. vanD and vanG-like gene clusters in a Ruminococcus species isolated from human bowel flora. *Antimicrobial agents and chemotherapy* 51 4111-4117 (2007).
30. Roager HM, Licht TR, Poulsen SK, Larsen TM, Bahl MI. Microbial enterotypes, inferred by the prevotella-to-bacteroides ratio, remained stable during a 6-month randomized controlled diet intervention with the new nordic diet. *Appl. Environ. Microbiol.* 80(3), 1142-1149 (2014).
31. Frost C, Song Y, Yu Z *et al.* The effect of apixaban on the pharmacokinetics of digoxin and atenolol in healthy subjects. *Clinical pharmacology : advances and applications* 9 19-28 (2017).
32. Gomes T, Mamdani MM, Juurlink DN. Macrolide-Induced Digoxin Toxicity: A Population-Based Study. *Clinical Pharmacology & Therapeutics* 86 383-386 (2009).

- **Reference annotations:** authors should highlight 6–8 references that are of particular significance to the subject under discussion as “\*\* of interest” or “\*\*\* of considerable interest”

[[ 4. Lindenbaum J, Rund DG, Butler VP, Tse-Eng D, Saha JR. Inactivation of Digoxin by the Gut Flora: Reversal by Antibiotic Therapy. *New England Journal of Medicine* 305 789-794 (1981).

\*\* One of the pioneering works on the subject of microbiome-mediated drug transformation. They attributed the digoxin chemical reduction to bacteria by demonstrating the following two things: *in vitro* reduction of digoxin by co-culture with stool samples (from “reducers”, but not from “non-reducers”); and reversal of *in vivo* digoxin reduction by antibiotics to healthy participants (erythromycin or tetracyclin). ]]

[[ 1. Saha JR, Butler VP, Neu HC, Lindenbaum J. Digoxin-inactivating bacteria: identification in human gut flora. *Science (New York, N.Y.)* 220 325-327 (1983).

\* First study identifying *Eggerthella lenta* (then named *Eubacterium lentum*) as the sole reducer of digoxin. From the hundreds of bacterial isolates from fecal microbiome samples from “reducers” (participants that exhibit *in vivo* digoxin reduction), only *E. lenta* transformed digoxin in *in vitro* cultures. (Secondarily, addition of arginine to the cultures inhibited that transformation.) ]]

[[ 2. Mathan VI, Wiederman J, Dobkin JF, Lindenbaum J. Geographic differences in digoxin inactivation, a metabolic activity of the human anaerobic gut flora. *Gut* 30 97-91 (1989).

\* International study comparing inter-individual levels of digoxin inactivation from three cohorts: 1. Americans from New-York 2. Indians from rural area 3. Indians from urban area; “Only 1.0% of Indians compared with 14.0% of Americans excreted large amounts of metabolites (greater than 40% DRP) in the urine”. ]]

[[ 9. Pincus M. Management of digoxin toxicity. *Australian prescriber* 39(1), 18 (2016).

\* Description of the management of digoxin toxicity cases. Thorough description of the cardiotonic drug’s mechanism of action, toxicity symptoms and management of affected patients. ]]

[[ 16. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nature Reviews Microbiology* 14 273-287 (2016).

\*\* Thorough review of different types of interaction between the gut microbiome and xenobiotics (including drugs). Nicely illustrated with a variety of examples. ]]

[[ 23. Buse JB, Defronzo RA, Rosenstock J *et al.* The Primary Glucose-Lowering Effect of Metformin Resides in the Gut, Not the Circulation. Results From Short-term Pharmacokinetic and 12-Week Dose-Ranging Studies. *Diabetes Care* 39 dc150488 (2015).

\* One of the first articles to pinpoint the target of metformin to be the gut microbiome, not a human metabolic pathway. Metformin has been used for years for its low toxicity and proved efficacy at reducing blood sugar, without knowing exactly how the drug worked. ]]

[[ 26. Raymond F, Ouameur AA, Déraspe M *et al.* The initial state of the human gut microbiome determines its reshaping by antibiotics. *The ISME journal* 10(3), 707 (2016).

\*\* Article from which the *in vivo* dataset used in this project was obtained. Our collaborators at Université Laval exposed healthy volunteers to assess how the antibiotic cefprozil affects the human gut microbiome and how the microbiome recovered after the antibiotic course. ]]

## 14.2 Résultats supplémentaires

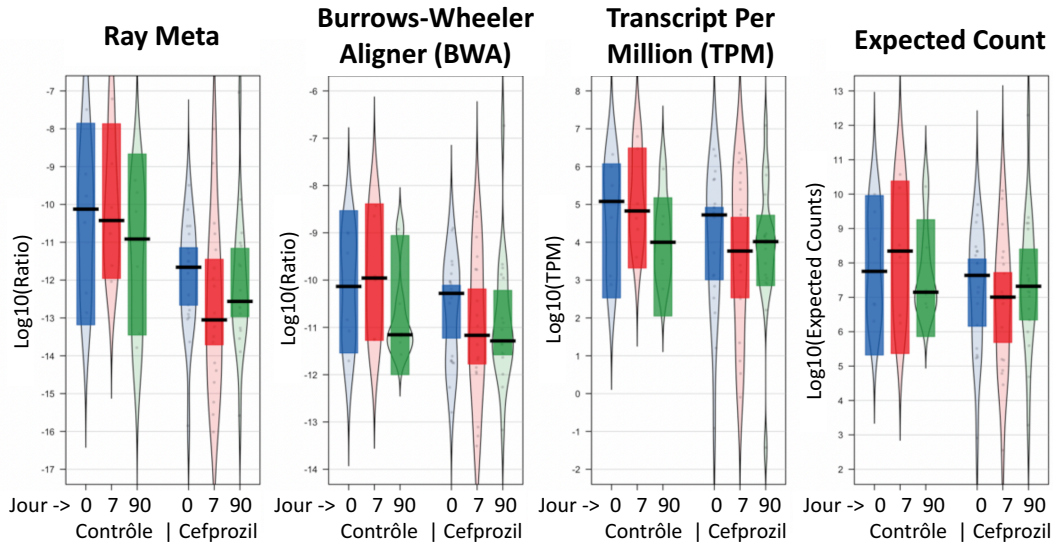


Figure S. 1. Comparaison des logiciels d'alignement de séquences pour quantifier l'abondance relative de *E. lenta*. Les alignements ont tous été faits en référence fermée sur la même base de données de 5 141 génomes bactériens de référence. Pour chaque aligneur, six groupes sont répartis en deux: les témoins (contrôles) et les volontaires exposés (céfprozil). Les deux groupes ont été échantillonnés aux jours J0, J7 et J90. Chaque méthode a généré des métriques qui leur sont propres, mais ce sont tous des abondances relatives, donc comparables. Les valeurs d'abondances ont toutes été transformées logarithmiquement pour améliorer la lisibilité des graphiques.

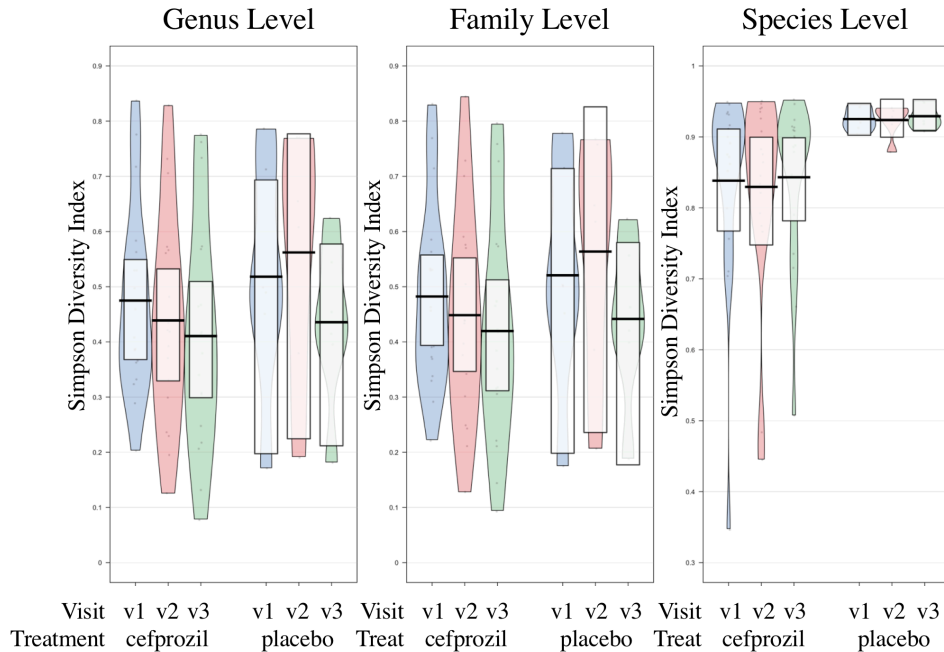


Figure S. 2. Index de Simpson calculé par échantillon après attribution taxonomique par Ray-Meta (figure générée à partir du matériel supplémentaire de Raymond et al. 2016 (217)). L'index de diversité a été calculé à trois niveaux taxonomiques, le genre, la famille et l'espèce (les trois sections principales du graphique). Pour chaque niveau, les  $n=18$  participants du groupe céfprozil (gauche) et les  $n=6$  du groupe placebo (droite) sont présentés aux trois visites (v1, v2, v3).

## 15 Discussion

### 15.1 Susceptibilité de *E. lenta* au céfprozil dans l'intestin

Nous avons évalué la susceptibilité *in vivo* de *E. lenta* à l'antibiotique céfprozil en se concentrant sur les 18 volontaires sains exposés à l'antibiotique durant une semaine. Les témoins non exposés n'ont pas été considérés, alors que l'échantillonnage du niveau initial de la bactérie permet d'utiliser chaque participant comme son propre contrôle négatif. De plus, les participants témoins constituent un groupe restreint de 6 individus et les résultats présentés à la Figure S. 1 montrent que le problème de la grande variabilité des données s'améliore entre le groupe de 6 et le groupe de 18, alors que les zones de haute densité de données sont visiblement moins étendues avec les volontaires exposés.

Malgré la variabilité mieux contrôlée avec les volontaires exposés à l'antibiotique, les valeurs-p calculées pour les trois intervalles de temps possibles de temps (entre les trois jours de collectes J0-J7-J90) n'ont donné aucun résultat de test statistiquement significatif, avec des valeurs de p associées de 0.766 (J0-J7), 0.538 (J7-J90) et 0.304 (J7-J90). Les tests non paramétriques ainsi réalisés ne permettent pas de conclure qu'il y a un effet, surtout pour l'intervalle d'intérêt J0-J7, avant et après la prise d'antibiotique, la valeur-p calculée est grande ( $0.766 \gg 0.05$ ; ce qui est très loin du seuil de significativité). Néanmoins, en observant la figure l'on peut voir que la médiane mesurée au jour 0 est très en haut de la zone de densité de probabilité calculée pour le jour 7. Si la médiane avait été à l'extérieur de la zone de haute probabilité inférée, cela aurait été un indicateur de l'inégalité des deux populations (avant et après l'antibiotique). Il est pensable que l'augmentation de la taille d'échantillon permettrait de réduire la largeur de la zone de densité calculée et ainsi éliminer ce chevauchement des distributions et éventuellement prouver la diminution de *E. lenta* statistiquement.

Pour ce qui est du panneau central, où chacune des données provenant de chaque patient sont reliées par une ligne (bleue en cas d'augmentation et rouge en cas de diminution) donne un

autre regard sur la variabilité des données. En effet, cela fait ressortir que dans le premier intervalle de temps (J0-J7), plusieurs patients qui avaient une abondance moyenne de *E. lenta* ont vu cette quantité fortement diminuée suite à la prise d'antibiotiques. Inversement, quelques autres populations de *E. lenta* ont fortement augmenté suite à la prise de l'antibiotique, possiblement. Cette différence est probablement due au fait que ce sont des échantillons de microbiome complet, donc ce n'est pas seulement l'effet de l'antibiotique sur *E. lenta*, mais sur toutes les populations bactériennes présentes. Donc, ce n'est pas seulement la bactérie d'intérêt qui augmente, mais peut-être que c'est un compétiteur écologique de *E. lenta* qui est plus fortement affectée et donc, la bactérie d'intérêt prend sa place et augmente. Pour résumer, ce n'est pas seulement la susceptibilité de la bactérie d'intérêt, mais la différence de susceptibilité avec les autres bactéries présentes dans le microbiome à ce moment.

Le troisième panneau de la figure illustre les  $\log_2(\text{foldchange})$ , c'est-à-dire une mesure de chaque différence à partir du jour initial, rapportée sur le taux initial de *E. lenta*. Dans ce type d'interprétation des données, une valeur de -1 correspond à une diminution de moitié par rapport au seuil initial et une valeur de -2 correspond à une diminution par un facteur de 4 (à cause de l'échelle logarithmique). Une valeur positive correspond à une augmentation par rapport au taux initial alors qu'une valeur négative est synonyme d'une diminution. Pour les deux intervalles de temps illustrés, la barre médiane est sous zéro (environ -0.5 dans les deux cas), ce qui signifie que pour les intervalles de temps J0->J7 et pour J0->J90, la majorité des échantillons ont vu une diminution de l'abondance de *E. lenta* suite à la prise de céfprozil et que ces taux restent bas même après la période de retour à la normale. Donc dans l'ensemble, bien que les statistiques ne soient pas conclusives, les résultats tendent à démontrer une diminution de la population de *E. lenta* suivant la prise de l'antibiotique céfprozil par des volontaires sains.

## **15.2 Susceptibilité de *E. lenta* à la céfoxitine *in vitro***

Nous avons aussi étudié l'effet d'une autre céphalosporine de seconde génération, la céfoxitine, sur le microbiome fécal complet cultivé *in vitro*. Ces analyses ont généré quelques



résultats intéressants, dont une diminution statistiquement significative de la bactérie d'intérêt lorsque cultivée avec l'antibiotique. Dans un premier temps, la figure 2 présente tous les résultats de sélectomique, sauf ceux provenant des 6 témoins qui n'ont pas été exposés à l'antibiotique, qui sont absents du panneau de droite de la figure. La figure présente deux panneaux qui présentent chacun 5 groupes de lectures d'abondance de *E. lenta*. Le premier de ces 5 groupes, à gauche en bleu dans les deux cas, représente le taux initial de la bactérie d'intérêt dans les échantillons avant leur mise en culture. Les quatre autres groupes sont les résultats de la culture *in vitro* et il est facile de constater que toutes les abondances de *E. lenta* sont beaucoup plus élevées suite à la culture qu'au seuil initial de l'échantillon, signifiant que peu importe la condition de culture, la bactérie *E. lenta* se développe très bien en laboratoire (minimalement, qu'elle se développe mieux que les autres bactéries avec lesquelles son abondance est comparée dans les calculs d'abondance relative).

Ensuite, la différence entre le panneau de gauche et de droite est l'exposition préalable à l'antibiotique cefprozil. Le panneau de droite (B.) présente les échantillons mis en culture provenant des volontaires après avoir été exposés à l'antibiotique durant une semaine, alors que celui de gauche (A.) présente les échantillons "naïfs", mis en culture sans avoir été exposés à quelque antibiotique préalablement à la mise en culture. Ainsi, ces résultats pourraient illustrer un effet préparateur sur les communautés bactériennes du microbiome intestinal des 18 participants à l'étude, alors que l'exposition à une première céphalosporine (le céfprozil) aurait pu activer des mécanismes pour prémunir les microbiotes contre l'exposition à une seconde céphalosporine *in vitro*, ici le céfprozil. Toutefois, la comparaison visuelle des deux panneaux (A./B.) de la figure 2 permet de constater que dans les 4 conditions de culture, le microbiote *E. lenta* a proliféré de manière très similaire. Des comparaisons numériques ont aussi été réalisées et confirment que la croissance observée n'est pas affectée par la préexposition au céfprozil.

La figure 3 présente un sous-ensemble de 24 échantillons de culture *in vitro* provenant des 24 échantillons des volontaires au jour de collecte J0, soit avant l'exposition à l'antibiotique. Les analyses statistiques des expériences *in vitro* sont ciblées sur les 48 échantillons mis en

culture en condition anaérobique, alors que cet environnement reproduit plus fidèlement la condition de l'intestin, avec les 24 échantillons cultivés en parallèle, avec ou sans l'antibiotique céfoxitine ajouté au milieu de culture. Ces échantillons mis en culture sont donc, non pré-exposés aux céphalosporines et cultivés sans oxygène, pour un total de n=24.

Le panneau de gauche de la figure 3 présente les 24 échantillons décrits au paragraphe précédent, cultivés sans et avec céfoxitine. La médiane est plus basse dans le cas de culture avec l'antibiotique, alors que la croissance de *E. lenta* est diminuée de façon statistiquement significative par rapport à son homologue cultivé sans antibiotique, avec une valeur de p de 0.0457 calculée avec un test de Wilcoxon-Mann-Withney. D'ailleurs, à la figure 2, toutes les paires de 24 échantillons mis en culture sans/avec céfoxitine démontrent la même tendance à l'inhibition de la croissance de *E. lenta*, même après la préexposition à une autre céphalosporine, ce qui soutient l'observation de la sensibilité de la bactérie aux céphalosporines. Les autres résultats de culture *in vitro* n'ont pas été considérés pour les tests statistiques présentés ici pour éviter de bonifier artificiellement les résultats des tests statistiques en incluant des échantillons qui sont presque des répliqués techniques.

Encore dans la figure 3, le panneau central présente les 24 échantillons décrits dans le paragraphe précédent, mais cette fois en les reliant en paires selon duquel participant à l'étude provient l'échantillon de microbiome mis en culture. Il est intéressant de constater que la majorité des échantillons connaissent une baisse significative de l'abondance de *E. lenta* en présence du céfoxitine (lignes rouges), mais que pour 7 échantillons c'est une augmentation de l'abondance qui est constatée (avec 4 échantillons qui augmentent par un facteur d'au moins 2). La tendance est quand même à la diminution, alors que les 17 autres échantillons ont vu une diminution d'abondance de la bactérie d'intérêt par rapport au taux sans antibiotique, avec 7 échantillons qui ont connu une diminution d'abondance par un facteur d'au moins 4. Finalement, le panneau de droite de la figure 3 présente les changements d'abondance relative avec antibiotique par rapport au niveau de la culture sans antibiotique sous la forme de  $\log_2(\text{Fold Change})$ , comme avec les résultats de la partie *in vivo* avec les volontaires. Comme décrit plus

tôt, un résultat négatif indique une diminution alors qu'un nombre positif témoigne d'une augmentation de l'abondance de la bactérie quantifiée. Cette distribution de valeurs illustre que, comme décrit plus tôt, certains échantillons ont proliféré positivement dans un milieu ajouté d'antibiotique, mais que pour la majorité, c'est une diminution d'abondance qui est observée. La valeur médiane de  $\log_2(\text{FC})$  est égale à -0.573, ce qui correspond à une diminution de l'abondance par un facteur de 1.49 fois par rapport à l'abondance initiale. Cela signifie que, en culture d'échantillons complets de microbiome fécal, la prolifération de la bactérie *E. lenta* est négativement affectée par l'exposition à l'antibiotique céfoxitine.

### **15.3 Forces et faiblesses de l'étude**

La présente étude est intéressante alors qu'elle revisite vieille une hypothèse de pharmacomicrobiomique, mais à une époque où les technologies permettent d'étudier efficacement le microbiome intestinal humain. En effet, comme décrit plus tôt, de nombreux facteurs ont mené à la marginalisation des évidences du rôle de *Eggerthella lenta* dans la pharmacocinétique de la digoxine. En démontrant la sensibilité du microbiote d'intérêt à deux céphalosporines, cette étude valide une étape préliminaire dans l'étude du rôle du microbiome intestinal dans les nombreux cas de toxicité à la digoxine qui surviennent suite à l'administration d'antibiotiques à des patients ayant des problèmes cardiaques.

Un second point intéressant de l'étude est le fait d'utiliser des données de séquençage génomique complets (WGS, décrit dans matériel et méthodes), ce qui a permis de raffiner la distinction entre les souches de bactéries en utilisant les génomes complets des différentes souches de *E. lenta* dont les génomes sont disponibles au NCBI. Cela a permis de jeter un coup d'œil nouveau sur un jeu de données déjà existant et d'identifier la variation de la souche métabolisatrice DSM 2243 suite à l'exposition aux antibiotiques. Il sera aussi éventuellement intéressant d'utiliser une approche de PCR ciblée sur le gène métabolisateur, comme dans l'article récent de Bisanz de 2018 (13), afin de suivre de plus près l'évolution du potentiel de dégradation de la digoxine par le microbiome suivant l'exposition à des antibiotiques. En effet,

il serait très intéressant de voir si le potentiel génétique pour la dégradation de la digoxine est affecté par la prise d'antibiotiques de la même façon que l'abondance de la souche métabolisatrice est affectée.

La limitation majeure du projet vient du fait que le design expérimental n'a pas été développé avec les présentes hypothèses en tête. En effet, aucun des volontaires n'ont reçu de digoxine et il est donc impossible de confirmer que l'antibiotique agit sur la pharmacocinétique du médicament cardiotonique. Les tests *in vitro* ont aussi été réalisés sans digoxine et les interactions entre les deux classes de molécules n'ont pas pu être observées dans ce contexte non plus. Ces limitations ne réduisent pas l'importance des conclusions du présent projet, soit la démonstration de la sensibilité de la bactérie métabolisatrice de la digoxine en présence d'antibiotiques de la famille des céphalosporines. Ce projet *post-hoc* d'analyses valide des prémisses pour l'élaboration de projets visant à confirmer le rôle de *E. lenta* dans la pharmacocinétique de la digoxine et son implication possible dans les cas de toxicité due aux interactions antibiotiques-digoxine.

## 15.4 Perspectives futures

Dans un premier temps, il est certain que la meilleure façon de tester le rôle de *Eggerthella lenta* dans les cas de toxicité à la digoxine serait de recruter les patients sous thérapie au cardiotonique lorsqu'ils se font prescrire des antibiotiques pour des infections non reliées. Il serait possible de prendre des échantillons pour analyser les deux paramètres principaux, soit l'abondance de *E. lenta* (échantillonnage métagénomique du microbiome) et quantifier l'élimination de la digoxine et de ses métabolites secondaires, surtout les DRP urinaires. Il serait ainsi possible de mesurer plusieurs paramètres très intéressants pour la théorie explorée dans ce projet. Dans un premier temps, il serait possible de quantifier la variation d'abondance de *E. lenta* dans des conditions physiologiques humaines et de possiblement même de suivre cette progression selon l'effet de divers antibiotiques (et ainsi comparer leurs effets respectifs). Ensuite, en échantillonnant la formation des métabolites secondaires, il serait possible de relier

la composition du microbiome intestinal *in vivo* avec le niveau de transformation de la digoxine en DRPs (Digoxin Reduced Products). Une analyse de type WGS sur génomes bactériens complet permettrait aussi de confirmer ou découvrir les gènes corrélés avec le niveau de transformation de la digoxine, comme ceux de la liste de gènes identifiés par Bisanz *et al.* en 2018 (13). Ces découvertes permettraient de concevoir des meilleurs tests pour l'échantillonnage et la caractérisation de *E. lenta* par rapport à son rôle dans la pharmacocinétique de la digoxine.

Pour faire les tests cliniques décrits au paragraphe précédent, il serait nécessaire d'obtenir un échantillon au jour zéro, soit avant la prise d'antibiotiques, afin d'établir le niveau de base de métabolisme et utiliser les patients comme leur propres contrôles négatifs. Cela est très important parce qu'il existe beaucoup de variabilité au niveau de l'activité de *E. lenta* entre les patients, comme illustré dans ce projet et dans quelques autres publications (12). Toutefois, il faudrait que le protocole soit bien implanté et ne puisse pas nuire à la qualité des soins prodigués, par exemple au niveau du temps d'administration des doses d'antibiotiques. Le résultat attendu serait une diminution des DRPs urinaires par rapport au jour initial d'échantillonnage (et donc une augmentation de la digoxine éliminée inchangée dans l'urine et les selles). Toutefois, la mise en place d'un protocole clinique est généralement très onéreux et d'autres résultats plus simples à obtenir pourraient aussi servir à confirmer le rôle du microbiote dans la pharmacocinétique de la digoxine.

Un autre étude clinique intéressante pourrait être réalisée avec des volontaires sains à qui l'on administre les deux médicaments (digoxine + antibiotique) afin de caractériser la cinétique de l'élimination du médicament de la manière décrite dans les derniers paragraphes. Une telle étude serait un croisement entre l'étude de Raymond *et al.* de 2016 (217) celle de Hiderling et Hartmann (221), la première ayant administré des antibiotiques et la seconde de la digoxine à des volontaires sains. Une telle approche serait plus facile à mettre en place, alors que le protocole ne serait pas tributaire de la facilité à recruter des patients qui sont dans la situation spécifique d'être sous thérapie à la digoxine et d'avoir besoin d'un traitement aux antibiotiques. Pour être très rigoureux lors de la réalisation de cet essai clinique, il serait

pertinent d'entamer la prise de la digoxine plusieurs jours avant l'administration de l'antibiotique étudié. Cela permettra de stabiliser la pharmacocinétique de la digoxine et donnera une période d'adaptation au microbiome intestinal, alors que la bactérie d'intérêt réagit en activant l'expression des gènes métaboliseurs seulement en présence de digoxine (215). Comme avec la procédure expérimentale décrite plus tôt avec les patients d'insuffisance cardiaque, un tel protocole permettrait d'identifier les gènes responsables de la transformation, prédire le taux de métabolisme et quantifier le changement d'abondance de la bactérie lors de l'administration d'antibiotiques.

Comme discuté dans l'introduction, la pharmacomicrobiomique n'a pas pu être sérieusement étudiée avant l'avènement des nouvelles technologies de séquençage qui permettent de caractériser un microbiome complet sans avoir recours aux techniques de culture des bactéries. L'heure est à l'exploration des fondements des interactions entre les bactéries et les médicaments, comme l'étude de Maier *et al.* parue dans Nature en 2018 (222) où le groupe de recherche a exposé *in vitro* des isolats de souches bactériennes importantes dans le microbiome intestinal humain normal pour évaluer l'impact de médicaments non antibiotiques sur leur croissance. Cette étude a testé un grand nombre de souches bactériennes et d'autant plus de médicaments avec des technologies de cultures en plaques à haut débit (high throughput). En plus des études de réponse du microbiome (souches isolées ou échantillons de microbiome complets) aux médicaments et autres xénobiotiques, il sera intéressant de faire d'autres tests *in vitro* pour évaluer à des frais minimums la transformation de médicaments par les microbiotes.

Dans cet esprit de tests de laboratoire *in vitro*, quelques expériences pourraient suivre les tests de sélectomique présentés ici, dans le but d'explorer le rôle de *E. lenta* dans la pharmacocinétique de la digoxine. Par exemple, il serait intéressant de dresser un profil de résistance complet de la bactérie par rapport à différentes classes d'antibiotiques, ce qui pourrait éventuellement aider à établir un guide clinique pour les patients traités avec le cardiotonique qui ont besoin de traitement aux antibiotiques. Ainsi, il serait possible d'identifier les antibiotiques qui sont le moins susceptible de perturber la pharmacocinétique de la digoxine et

suggérer ceux-ci pour la clinique. Comme élaboré plus tôt, dans une telle expérience, on s'attendrait à ce que la survie de *E. lenta* soit plus fortement affectée par des antibiotiques de la classe des macrolides (comme la clarithromycine) que les antibiotiques de la classe des céphalosporines (comme le céfprozil, la céfoxitine et le céfuroxime). Si la clarithromycine affecte plus sévèrement la survie de *E. lenta* que les céphalosporines, cela concorderait avec l'hypothèse du rôle du microbiome dans les cas de toxicité à la digoxine observés suivant la prise d'antibiotiques.

Finalement, il serait intéressant d'observer la réduction de la digoxine par le microbiome *in vitro* en exposant les bactéries au cardiotonique en laboratoire. Par exemple, il serait possible de caractériser plus en détail des isolats des 24 souches de *Eggerthella lenta* récemment répertoriées (13), et de quantifier précisément la vitesse de transformation de la digoxine en DRPs dans le milieu de culture. En faisant une telle expérience, il serait possible de mieux relier les souches avec les niveaux de métabolisation et même contribuer à l'amélioration de la suite logicielle ElenMatchR. Ensuite, il serait pertinent de refaire des cultures de microbiomes fécaux complets et d'exposer ces cultures à la digoxine. Cela permettrait de voir l'effet des antibiotiques sur ces microbiomes complets en culture et leur niveau de transformation de la digoxine. Encore ici, il serait très intéressant de comparer l'effet de différents antibiotiques sur la prolifération de *E. lenta* ainsi que sur son niveau de métabolisme de la digoxine. Ces expériences permettront de mieux comprendre les fondements pharmacocinétique de la digoxine et éventuellement réduire les risques d'intoxication pour les patients qui utilisent ce médicament pour traiter leurs conditions cardiaques problématiques.

## 16 Conclusion

Le présent projet a permis de démontrer la sensibilité de la bactérie *Eggerthella lenta* à deux céphalosporines de seconde génération, le céfprozil (*in vivo*) et la céfoxitine (*in vitro*). La démonstration de l'effet des antibiotiques sur la survie de ce microbiote intestinal est importante dans la démonstration du rôle du microbiome intestinal dans la pharmacocinétique de la digoxine chez l'humain et ses implications dans la gestion clinique du cardiotonique.

La susceptibilité de *E. lenta* à l'antibiotique céfprozil a été étudiée avec un jeu de données WGS (Whole Genome Shotgun) sur l'ensemble des bactéries présentes dans des échantillons de microbiome fécal de 18 volontaires sains. Le ratio standardisé de lectures de PCR attribuées à la souche métabolisatrice de *E. lenta* (DSM 2243) a effectivement chuté entre le jour initial de collecte de données (J0) et la seconde journée de collecte (J7), soit après une semaine de prise quotidienne de céfprozil. La valeur d'abondance médiane de la bactérie a chuté par 58,3% durant la période d'exposition à l'antibiotique, mais cette diminution n'a pas donné lieu à des résultats statistiquement significatifs de valeur-p ( $0.766 \gg 0.05$ ). Cette non-significativité des résultats est due à la grande variabilité interindividuelle au niveau de la composition des microbiomes et au petit nombre d'échantillons étudiés dans le volet impliquant des sujets humains (n=18). Intéressamment, l'abondance du microbiote d'intérêt continue à rester basse, et même à diminuer un petit peu, 90 jours après la cessation de la prise d'antibiotiques. La raison pour laquelle *E. lenta* ne retrouve pas son abondance originale suivant l'arrêt du traitement est inconnue et a probablement à voir avec le fait que ce sont des échantillons complets de microbiome fécal. Nous n'avons pas mesuré une abondance absolue de *E. lenta*, mais une abondance relative, l'abondance de *E. lenta* et toujours comparée à la prolifération des autres microbiotes intestinaux présents.



Dans le second volet de recherche, les échantillons de microbiome fécal complet mis en culture *in vitro* ont permis de démontrer la sensibilité de *E. lenta* de façon statistiquement significative avec une valeur-p de 0.0457 (test de Wilcoxon Mann-Whitney). L'ajout des autres résultats de cultures *in vitro* aurait résulté en l'ajout de réplicats techniques quasi similaires aux n=24 considérés dans les tests statistiques (sur les 192 échantillons mis en culture). Il vaut de noter que dans toutes les autres paires d'échantillons mis en culture, la même tendance s'observe, avec des valeurs de comptes normalisés attribués à *E. lenta* diminuées dans les échantillons supplémentés avec la céfoxitine comparativement à ceux incubés sans antibiotiques.

Les résultats calculés d'abondance de *E. lenta* dans les 192 échantillons mis en culture *in vitro* ont été représentés graphiquement et permettent de dégager quelques autres conclusions plus sommaires. Par exemple, compte tenu de la grande différence entre les niveaux initiaux d'abondance de la bactérie (aux jours de collecte J0 et J7) et tous les niveaux finaux (après culture), il est clair que la bactérie *E. lenta* prolifère bien *in vitro* (niveaux médians 230 fois plus élevé après la semaine de culture qu'avant). L'expérience aurait pu démontrer un effet préparateur de l'exposition au céfprozil dans l'intestin et montrer une réaction différente de *E. lenta* lorsque cultivée subséquemment en laboratoire. Toutefois, cet effet n'est pas observé, alors que le microbiote d'intérêt prolifère largement peu importe la condition de culture. En fait, la condition de culture est un meilleur prédicteur de l'abondance finale de *E. lenta* que l'exposition préalable au céfprozil. Cela signifie que l'exposition intra-intestinale préalable à une céphalosporine n'affecte pas la prolifération *in vitro* du microbiome, même lorsque soumis à une seconde céphalosporine en culture.

Les conclusions qui peuvent être tirées de ces analyses sont importantes parce qu'elles vont dans le même sens que la théorie selon laquelle les interactions médicamenteuses problématiques entre les antibiotiques et la digoxine sont explicable par le rôle du microbiome dans la pharmacocinétique de la digoxine. Par exemple, des tests *in vitro* supplémentaires

permettront de déterminer quels antibiotiques sont le moins susceptibles d'être problématiques en clinique. D'autres tests *in vitro* permettront aussi de mieux caractériser les nombreuses souches de *Eggerthella lenta* et améliorer des logiciels comme ElenMatchR et prédire le niveau de métabolisme de la digoxine selon les analyses métagénomiques. Des tests plus poussés sur pharmacocinétique de la digoxine avec des volontaires sains, semblables à celle de Hinderling et Hartmann (221) en 1991, mais avec la prise sélective d'antibiotiques pourraient valider ou infirmer hors de tout doute l'implication de *E. lenta* dans les cas de toxicité à la digoxine observés en clinique suivant la prise d'antibiotiques.

## 17 Bibliographie

1. Lok C. Mining the microbial dark matter. *Nature* [Internet]. 2015 Jun 16 [cited 2018 Mar 28];522(7556):270–3. Available from: <http://www.nature.com/doi/10.1038/522270a>
2. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Evolution* (N Y) [Internet]. 1990 [cited 2017 Dec 1];87:4576–9. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC54159/pdf/pnas01037-0173.pdf>
3. Woo PCY, Lau SKP, Teng JLL, Tse H, Yuen K-Y. Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clin Microbiol Infect* [Internet]. 2008 Oct [cited 2018 Jun 4];14(10):908–34. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18828852>
4. Lagier J-C, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, et al. Culturing the human microbiota and culturomics. *Nat Rev Microbiol* [Internet]. 2018 Sep 24 [cited 2019 Apr 20];16(9):540–50. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29937540>
5. Kaeberlein T, Lewis K, Epstein SS. Isolating “uncultivable” microorganisms in pure culture in a simulated natural environment. *Science* [Internet]. 2002 May 10 [cited 2019 Apr 20];296(5570):1127–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12004133>
6. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P, Rolain JM, et al. Ongoing Revolution in Bacteriology: Routine Identification of Bacteria by Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry. *Clin Infect Dis* [Internet]. 2009 Aug 15 [cited 2019 Apr 20];49(4):543–51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19583519>
7. Croxatto A, Prod’hom G, Greub G. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev* [Internet]. 2012 Mar [cited 2019 Apr 20];36(2):380–407. Available from:

<http://www.ncbi.nlm.nih.gov/pubmed/22092265>

8. Burton JH, Johnson M, Johnson J, Hsia DS, Greenway FL, Heiman ML. Addition of a Gastrointestinal Microbiome Modulator to Metformin Improves Metformin Tolerance and Fasting Glucose Levels. *J Diabetes Sci Technol* [Internet]. 2015 Jul 23 [cited 2018 Feb 22];9(4):808–14. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25802471>
9. Forslund K, Hildebrand F, Nielsen T, Falony G, Le Chatelier E, Sunagawa S, et al. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* [Internet]. 2015 Dec 2 [cited 2018 Feb 22];528(7581):262–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26633628>
10. Maniar K, Moideen A, Bhattacharyya R, Banerjee D. Metformin exerts anti-obesity effect via gut microbiome modulation in prediabetics: A hypothesis. *Med Hypotheses* [Internet]. 2017 Jul [cited 2017 Sep 8];104:117–20. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28673568>
11. Saha JR, Butler VP, Neu HC, Lindenbaum J. Digoxin-inactivating bacteria: identification in human gut flora. *Science* [Internet]. 1983 Apr 15 [cited 2017 Sep 8];220(4594):325–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/6836275>
12. Mathan VI, Wiederman J, Dobkin JF, Lindenbaum J. Geographic differences in digoxin inactivation, a metabolic activity of the human anaerobic gut flora. *Gut* [Internet]. 1989 [cited 2017 May 9];30:97–1. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1434295/pdf/gut00220-0093.pdf>
13. Bisanz JE, Soto-Perez P, Lam KN, Bess EN, Haiser HJ, Allen-Vercoe E, et al. Illuminating the microbiome’s dark matter: a functional genomic toolkit for the study of human gut Actinobacteria. *bioRxiv* [Internet]. 2018;304840. Available from: <https://www.biorxiv.org/content/early/2018/05/01/304840?%3Fcollection=>
14. Clark MA, Finkel R (Richard S., Rey JA, Whalen K. *Lippincott’s illustrated reviews. Pharmacology.* 612 p.
15. Adams KF, Ghali JK, Herbert Patterson J, Stough WG, Butler J, Bauman JL, et al. A perspective on re-evaluating digoxin’s role in the current management of patients with chronic systolic heart failure: targeting serum concentration to reduce hospitalization and improve safety profile. *Eur J Heart Fail* [Internet]. 2014 May [cited 2016 Sep 20];16(5):483–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24574198>

16. Sender R, Fuchs S, Milo R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* [Internet]. 2016 Jan 28 [cited 2017 Nov 14];164(3):337–40. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26824647>
17. Sender R, Fuchs S, Milo R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biol* [Internet]. 2016 Aug 19 [cited 2017 Nov 10];14(8):e1002533. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27541692>
18. Ley RE, Peterson DA, Gordon JI. Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine. *Cell* [Internet]. 2006 Feb 24 [cited 2017 Nov 10];124(4):837–48. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16497592>
19. Prescott SL. History of medicine: Origin of the term microbiome and why it matters. *Hum Microbiome J* [Internet]. 2017 Jun 1 [cited 2018 Feb 14];4:24–5. Available from: <https://www.sciencedirect.com/science/article/pii/S245223171730012X>
20. Gao R, Gao Z, Huang L, Qin H. Gut microbiota and colorectal cancer. *Eur J Clin Microbiol Infect Dis* [Internet]. 2017 May 7 [cited 2017 Nov 14];36(5):757–69. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28063002>
21. O’Hara AM, Shanahan F. The gut flora as a forgotten organ. *EMBO Rep* [Internet]. 2006 Jul [cited 2017 Nov 14];7(7):688–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16819463>
22. Ruestow EG (Edward G. The microscope in the Dutch Republic: the shaping of discovery [Internet]. Cambridge University Press; 1996 [cited 2018 Jan 26]. 348 p. Available from: <http://www.cambridge.org/ca/academic/subjects/history/history-science-and-technology/microscope-dutch-republic-shaping-discovery?format=PB&isbn=9780521528634>
23. Robertson LA. van Leeuwenhoek microscopes—where are they now? *FEMS Microbiol Lett* [Internet]. 2015 May 1 [cited 2018 Jan 25];362(9). Available from: <http://academic.oup.com/femsle/article/doi/10.1093/femsle/fnv056/549359/van-Leeuwenhoek-microscopeswhere-are-they-now>
24. Smith J. The inverted microscope (a new form of microscope.) With the description of a new eye-piece micrometer, and a new form of goniometer for measuring the angles of crystals under the microscope. [Internet]. [New York]; 1852 [cited 2018 Jan 26]. Available from: <http://www.worldcat.org/title/inverted-microscope-a-new-form-of->

- microscope-with-the-description-of-a-new-eye-piece-micrometer-and-a-new-form-of-goniometer-for-measuring-the-angles-of-crystals-under-the-microscope/oclc/5325620
25. Rousseau C, Poilane I, De Pontual L, Maherault A-C, Le Monnier A, Collignon A. Clostridium difficile Carriage in Healthy Infants in the Community: A Potential Reservoir for Pathogenic Strains. Clin Infect Dis [Internet]. 2012 Nov 1 [cited 2018 Jan 22];55(9):1209–15. Available from: <https://academic.oup.com/cid/article-lookup/doi/10.1093/cid/cis637>
  26. Lepage P, Leclerc MC, Joossens M, Mondot S, Blottière HM, Raes J, et al. A metagenomic insight into our gut's microbiome. Gut [Internet]. 2013 Jan [cited 2018 Jan 25];62(1):146–58. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22525886>
  27. Marteau P. Bacterial Flora in Inflammatory Bowel Disease. Dig Dis [Internet]. 2009 [cited 2018 Jan 25];27(1):99–103. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20203504>
  28. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, et al. The minimal gene complement of Mycoplasma genitalium. Science [Internet]. 1995 Oct 20 [cited 2018 Jan 22];270(5235):397–403. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7569993>
  29. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science [Internet]. 1995 Jul 28 [cited 2018 Jan 22];269(5223):496–512. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7542800>
  30. Suzuki TA, Nachman MW. Spatial Heterogeneity of Gut Microbial Composition along the Gastrointestinal Tract in Natural Populations of House Mice. Zoetendal EG, editor. PLoS One [Internet]. 2016 Sep 26 [cited 2018 Jan 25];11(9):e0163720. Available from: <http://dx.plos.org/10.1371/journal.pone.0163720>
  31. Ciuti G, Menciassi A, Dario P. Capsule Endoscopy: From Current Achievements to Open Challenges. IEEE Rev Biomed Eng [Internet]. 2011 [cited 2018 Jan 25];4:59–72. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22273791>
  32. Moglia A, Menciassi A, Schurr MO, Dario P. Wireless capsule endoscopy: from diagnostic devices to multipurpose robotic systems. Biomed Microdevices [Internet]. 2007 Apr 2 [cited 2018 Jan 25];9(2):235–43. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/17160703>
33. Fisher LR, Hasler WL. New vision in video capsule endoscopy: current status and future directions. *Nat Rev Gastroenterol Hepatol* [Internet]. 2012 Jul 8 [cited 2018 Jan 25];9(7):392–405. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22565098>
  34. Amoako-Tuffour Y, Jones ML, Shalabi N, Labbe A, Vengallatore S, Prakash S. Ingestible gastrointestinal sampling devices: state-of-the-art and future directions. *Crit Rev Biomed Eng* [Internet]. 2014 [cited 2018 Jan 25];42(1):1–15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25271356>
  35. Carding S, Verbeke K, Vipond DT, Corfe BM, Owen LJ. Dysbiosis of the gut microbiota in disease. *Microb Ecol Health Dis* [Internet]. 2015 [cited 2018 Jan 29];26:26191. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25651997>
  36. Gong D, Gong X, Wang L, Yu X, Dong Q. Involvement of Reduced Microbial Diversity in Inflammatory Bowel Disease. *Gastroenterol Res Pract* [Internet]. 2016 Dec 15 [cited 2018 Jan 29];2016:1–7. Available from: <https://www.hindawi.com/journals/grp/2016/6951091/>
  37. Nowak P, Troseid M, Avershina E, Barqasho B, Neogi U, Holm K, et al. Gut microbiota diversity predicts immune status in HIV-1 infection. *AIDS* [Internet]. 2015 Nov 28 [cited 2018 Jan 29];29(18):2409–18. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26355675>
  38. Abrahamsson TR, Jakobsson HE, Andersson AF, Björkstén B, Engstrand L, Jenmalm MC. Low diversity of the gut microbiota in infants with atopic eczema. *J Allergy Clin Immunol* [Internet]. 2012 Feb [cited 2018 Jan 29];129(2):434–440.e2. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22153774>
  39. Scher JU, Ubeda C, Artacho A, Attur M, Isaac S, Reddy SM, et al. Decreased Bacterial Diversity Characterizes the Altered Gut Microbiota in Patients With Psoriatic Arthritis, Resembling Dysbiosis in Inflammatory Bowel Disease. *Arthritis Rheumatol* [Internet]. 2015 Jan [cited 2018 Jan 29];67(1):128–39. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25319745>
  40. Mercer DK, Scott KP, Melville CM, Glover LA, Flint HJ. Transformation of an oral bacterium via chromosomal integration of free DNA in the presence of human saliva. *FEMS Microbiol Lett* [Internet]. 2001 Jun 25 [cited 2018 Jan 29];200(2):163–7.

Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11425469>

41. Mercer DK, Scott KP, Bruce-Johnson WA, Glover LA, Flint HJ. Fate of free DNA and transformation of the oral bacterium *Streptococcus gordonii* DL1 by plasmid DNA in human saliva. *Appl Environ Microbiol* [Internet]. 1999 Jan [cited 2018 Jan 29];65(1):6–10. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9872752>
42. Bailey JK, Pinyon JL, Anantham S, Correspondence RMH, Hall RM. Commensal *Escherichia coli* of healthy humans: a reservoir for antibiotic-resistance determinants. [cited 2018 May 10]; Available from: <http://www.microbiologyresearch.org/docserver/fulltext/jmm/59/11/1331.pdf?expires=1525995772&id=id&acname=guest&checksum=7838EF06E081E0611D6412FF78D4491F>
43. Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, Floyd R, et al. Defining operational taxonomic units using DNA barcode data. *Philos Trans R Soc Lond B Biol Sci* [Internet]. 2005 Oct 29 [cited 2018 Jan 29];360(1462):1935–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16214751>
44. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* [Internet]. 2010 Oct 1 [cited 2018 Jan 29];26(19):2460–1. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btq461>
45. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* [Internet]. 2012 Dec 1 [cited 2018 Jan 29];28(23):3150–2. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bts565>
46. Koslicki D, Foucart S, Rosen G. WGSQuikr: Fast Whole-Genome Shotgun Metagenomic Classification. Liles MR, editor. *PLoS One* [Internet]. 2014 Mar 13 [cited 2018 Jan 29];9(3):e91784. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24626336>
47. Gorvitovskaia A, Holmes SP, Huse SM. Interpreting *Prevotella* and *Bacteroides* as biomarkers of diet and lifestyle. *Microbiome* [Internet]. 2016 Dec 12 [cited 2018 Jan 31];4(1):15. Available from: <http://www.microbiomejournal.com/content/4/1/15>
48. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* [Internet]. 2011 Oct 7 [cited



- 2018 Jan 31];334(6052):105–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21885731>
49. Jeffery IB, Claesson MJ, O’Toole PW, Shanahan F. Categorization of the gut microbiota: enterotypes or gradients? *Nat Rev Microbiol* [Internet]. 2012 Sep 1 [cited 2018 Jan 31];10(9):591–2. Available from: <http://www.nature.com/articles/nrmicro2859>
50. Knights D, Ward TL, McKinlay CE, Miller H, Gonzalez A, McDonald D, et al. Rethinking “Enterotypes.” *Cell Host Microbe* [Internet]. 2014 Oct 8 [cited 2018 Jan 31];16(4):433–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25299329>
51. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. *Nature* [Internet]. 2011 May 12 [cited 2018 Jan 31];473(7346):174–80. Available from: <http://www.nature.com/doi/10.1038/nature09944>
52. Su L-J, Liu Y-Q, Liu H, Wang Y, Li Y, Lin H-M, et al. Linking lignocellulosic dietary patterns with gut microbial Enterotypes of *Tsaietmies ampliceps* and comparison with *Mironasutietmies shangchengensis*. *Genet Mol Res* [Internet]. 2015 Oct 30 [cited 2018 Feb 1];14(4):13954–67. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26535711>
53. Li J, Powell JE, Guo J, Evans JD, Wu J, Williams P, et al. Two gut community enterotypes recur in diverse bumblebee species. *Curr Biol* [Internet]. 2015 Aug 3 [cited 2018 Feb 1];25(15):R652–3. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26241138>
54. Moeller AH, Degnan PH, Pusey AE, Wilson ML, Hahn BH, Ochman H. Chimpanzees and humans harbour compositionally similar gut enterotypes. *Nat Commun* [Internet]. 2012 Jan 13 [cited 2018 Feb 1];3(1):1179. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23149725>
55. Bushman FD, Lewis JD, Wu GD. Diet, Gut Enterotypes and Health: Is There a Link? In: Nestle Nutrition Institute workshop series [Internet]. 2013 [cited 2018 Feb 1]. p. 65–73. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24107497>
56. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* [Internet]. 2016 Jan [cited 2018 Feb 1];65(1):57–62. Available

- from: <http://www.ncbi.nlm.nih.gov/pubmed/26069274>
57. Sebastián Domingo JJ, Sánchez Sánchez C. From the intestinal flora to the microbiome. *Rev Española Enfermedades Dig* [Internet]. 2017 Jan [cited 2018 Feb 1];110(1):51–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29271225>
  58. Costea PI, Hildebrand F, Manimozhayan A, Bäckhed F, Blaser MJ, Bushman FD, et al. Enterotypes in the landscape of gut microbial community composition. *Nat Microbiol* [Internet]. 2018 Jan 18 [cited 2018 Feb 1];3(1):8–16. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29255284>
  59. Kang C, Zhang Y, Zhu X, Liu K, Wang X, Chen M, et al. Healthy Subjects Differentially Respond to Dietary Capsaicin Correlating with Specific Gut Enterotypes. *J Clin Endocrinol Metab* [Internet]. 2016 Dec [cited 2018 Feb 1];101(12):4681–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27676396>
  60. Zoetendal EG, de Vos WM. Effect of diet on the intestinal microbiota and its activity. *Curr Opin Gastroenterol* [Internet]. 2014 Mar [cited 2018 Feb 1];30(2):189–95. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24457346>
  61. Lim MY, Rho M, Song Y-M, Lee K, Sung J, Ko G. Stability of Gut Enterotypes in Korean Monozygotic Twins and Their Association with Biomarkers and Diet. *Sci Rep* [Internet]. 2015 May 8 [cited 2018 Feb 1];4(1):7348. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25482875>
  62. Koliada A, Syzenko G, Moseiko V, Budovska L, Puchkov K, Perederiy V, et al. Association between body mass index and Firmicutes/Bacteroidetes ratio in an adult Ukrainian population. *BMC Microbiol* [Internet]. 2017 May 22 [cited 2018 Feb 2];17(1):120. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28532414>
  63. Hjorth MF, Roager HM, Larsen TM, Poulsen SK, Licht TR, Bahl MI, et al. Pre-treatment microbial Prevotella-to-Bacteroides ratio, determines body fat loss success during a 6-month randomized controlled diet intervention. *Int J Obes* [Internet]. 2017 Oct 10 [cited 2018 Feb 5]; Available from: <http://www.nature.com/doi/10.1038/ijo.2017.220>
  64. Hodin CM, Visschers RGJ, Rensen SS, Boonen B, Damink SWMO, Lenaerts K, et al. Total Parenteral Nutrition Induces a Shift in the Firmicutes to Bacteroidetes Ratio in Association with Paneth Cell Activation in Rats<sup>1,2</sup>. *J Nutr* [Internet]. 2012 Dec 1 [cited 2018 Feb 2];142(12):2141–7. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/23096015>
65. Roager HM, Licht TR, Poulsen SK, Larsen TM, Bahl MI. Microbial enterotypes, inferred by the Prevotella-to-Bacteroides ratio, remained stable during a 6-month randomized controlled diet intervention with the new nordic diet. *Appl Environ Microbiol* [Internet]. 2014 Feb 1 [cited 2018 Feb 2];80(3):1142–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24296500>
  66. Mariat D, Firmesse O, Levenez F, Guimarães V, Sokol H, Doré J, et al. The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiol* [Internet]. 2009 Jun 9 [cited 2018 Feb 2];9(1):123. Available from: <http://bmcmicrobiol.biomedcentral.com/articles/10.1186/1471-2180-9-123>
  67. Ursell LK, Metcalf JL, Parfrey LW, Knight R. Defining the human microbiome. *Nutr Rev* [Internet]. 2012 Aug [cited 2017 Dec 8];70(Suppl 1):S38–44. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22861806>
  68. Shade A, Handelsman J. Beyond the Venn diagram: the hunt for a core microbiome. *Environ Microbiol* [Internet]. 2012 Jan [cited 2018 Feb 2];14(1):4–12. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22004523>
  69. Huse SM, Ye Y, Zhou Y, Fodor AA. A Core Human Microbiome as Viewed through 16S rRNA Sequence Clusters. Ahmed N, editor. *PLoS One* [Internet]. 2012 Jun 13 [cited 2018 Feb 2];7(6):e34242. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22719824>
  70. Brown T, Otero C, Grajales A, Rodriguez E, Rodriguez-Lanetty M. Worldwide exploration of the microbiome harbored by the cnidarian model, *Exaiptasia pallida* (Agassiz in Verrill, 1864) indicates a lack of bacterial association specificity at a lower taxonomic rank. *PeerJ* [Internet]. 2017 [cited 2017 Dec 7];5:e3235. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28533949>
  71. Muller EM, Fine M, Ritchie KB. The stable microbiome of inter and sub-tidal anemone species under increasing pCO<sub>2</sub>. *Sci Rep* [Internet]. 2016 Dec 23 [cited 2017 Dec 7];6(1):37387. Available from: <http://www.nature.com/articles/srep37387>
  72. Ryu T, Seridi L, Moitinho-Silva L, Oates M, Liew YJ, Mavromatis C, et al. Hologenome analysis of two marine sponges with different microbiomes. *BMC Genomics* [Internet]. 2016 Dec 29 [cited 2017 Dec 7];17(1):158. Available from:

<http://www.biomedcentral.com/1471-2164/17/158>

73. Ramanan R, Kim B-H, Cho D-H, Oh H-M, Kim H-S. Algae–bacteria interactions: Evolution, ecology and emerging applications. *Biotechnol Adv* [Internet]. 2016 Jan 1 [cited 2018 Feb 5];34(1):14–29. Available from: <https://www.sciencedirect.com/science/article/pii/S0734975015300586>
74. Augustin R, Schröder K, Murillo Rincón AP, Fraune S, Anton-Erxleben F, Herbst E-M, et al. A secreted antibacterial neuropeptide shapes the microbiome of Hydra. *Nat Commun* [Internet]. 2017 Dec 26 [cited 2017 Dec 7];8(1):698. Available from: <http://www.nature.com/articles/s41467-017-00625-1>
75. Schopf JW, Kitajima K, Spicuzza MJ, Kudryavtsev AB, Valley JW. SIMS analyses of the oldest known assemblage of microfossils document their taxon-correlated carbon isotope compositions. *Proc Natl Acad Sci U S A* [Internet]. 2018 Jan 2 [cited 2018 Feb 5];115(1):53–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29255053>
76. Bromham L, Rambaut A, Fortey R, Cooper A, Penny D. Testing the Cambrian explosion hypothesis by using a molecular dating technique. *Proc Natl Acad Sci U S A* [Internet]. 1998 Oct 13 [cited 2018 Feb 5];95(21):12386–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9770496>
77. Javaux EJ. The Early Eukaryotic Fossil Record. In: *Eukaryotic Membranes and Cytoskeleton* [Internet]. New York, NY: Springer New York; 2007 [cited 2018 Feb 5]. p. 1–19. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17977455>
78. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A* [Internet]. 1998 Jun 9 [cited 2018 Feb 5];95(12):6578–83. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9618454>
79. Singh Y, Ahmad J, Musarrat J, Ehtesham NZ, Hasnain SE. Emerging importance of holobionts in evolution and in probiotics. *Gut Pathog* [Internet]. 2013 May 22 [cited 2018 Feb 5];5(1):12. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23694677>
80. Dheilly NM. Holobiont–Holobiont Interactions: Redefining Host–Parasite Interactions. Rall GF, editor. *PLoS Pathog* [Internet]. 2014 Jul 3 [cited 2018 Feb 5];10(7):e1004093. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24992663>
81. Grasis JA. The Intra-Dependence of Viruses and the Holobiont. *Front Immunol* [Internet]. 2017 [cited 2018 Feb 5];8:1501. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/29170664>
82. Guerrero R, Margulis L, Berlanga M. Symbiogenesis: the holobiont as a unit of evolution. *Int Microbiol* [Internet]. 2013 Sep [cited 2018 Feb 5];16(3):133–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24568029>
  83. Six DL. The Bark Beetle Holobiont: Why Microbes Matter. *J Chem Ecol* [Internet]. 2013 Jul 12 [cited 2018 Feb 5];39(7):989–1002. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23846183>
  84. Zilber-Rosenberg I, Rosenberg E. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol Rev* [Internet]. 2008 Aug [cited 2018 Feb 5];32(5):723–35. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18549407>
  85. Moran NA, Sloan DB. The Hologenome Concept: Helpful or Hollow? *PLOS Biol* [Internet]. 2015 Dec 4 [cited 2018 Feb 5];13(12):e1002311. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26636661>
  86. Kopac SM, Klassen JL. Can They Make It on Their Own? Hosts, Microbes, and the Holobiont Niche. *Front Microbiol* [Internet]. 2016 [cited 2018 Feb 5];7:1647. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27818648>
  87. Douglas AE, Werren JH. Holes in the Hologenome: Why Host-Microbe Symbioses Are Not Holobionts. *MBio* [Internet]. 2016 Mar 31 [cited 2018 Feb 5];7(2):e02099. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27034285>
  88. Stephens JC, Reich DE, Goldstein DB, Shin HD, Smith MW, Carrington M, et al. Dating the Origin of the CCR5-Δ32 AIDS-Resistance Allele by the Coalescence of Haplotypes. *Am J Hum Genet* [Internet]. 1998 Jun [cited 2018 Feb 5];62(6):1507–15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9585595>
  89. Duncan SR, Scott S, Duncan CJ. Reappraisal of the historical selective pressures for the CCR5-32 mutation. *J Med Genet* [Internet]. 2005 Mar 1 [cited 2018 Feb 5];42(3):205–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15744032>
  90. Cohn SK, Weaver LT. The Black Death and AIDS: CCR5-32 in genetics and history. *QJM* [Internet]. 2006 Aug 1 [cited 2018 Feb 5];99(8):497–503. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16880184>
  91. Newburg DS. Glycobiology of human milk. *Biochem* [Internet]. 2013 Jul 13 [cited 2018

- Feb 5];78(7):771–85. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24010840>
92. Underwood MA, Gaerlan S, De Leoz MLA, Dimapasoc L, Kalanetra KM, Lemay DG, et al. Human milk oligosaccharides in premature infants: absorption, excretion and influence on the intestinal microbiota. *Pediatr Res* [Internet]. 2015 Dec 31 [cited 2018 Feb 5];78(6):670–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26322410>
  93. LaTuga M, Stuebe A, Seed P. A Review of the Source and Function of Microbiota in Breast Milk. *Semin Reprod Med* [Internet]. 2014 Jan 3 [cited 2018 Feb 5];32(01):068–73. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24390923>
  94. Bode L. Human milk oligosaccharides: Every baby needs a sugar mama. *Glycobiology* [Internet]. 2012 Sep 1 [cited 2018 Feb 5];22(9):1147–62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22513036>
  95. Reynolds JD, Morris B. The evolution and involution of Peyer’s patches in fetal and postnatal sheep. *Eur J Immunol* [Internet]. 1983 Aug [cited 2018 Feb 5];13(8):627–35. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/6884422>
  96. Cresci GA, Bawden E. Gut Microbiome. *Nutr Clin Pract* [Internet]. 2015 Dec 8 [cited 2018 Feb 5];30(6):734–46. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26449893>
  97. Belkaid Y, Harrison OJ. Homeostatic Immunity and the Microbiota. *Immunity* [Internet]. 2017 Apr 18 [cited 2018 Feb 5];46(4):562–76. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28423337>
  98. Nikoopour E, Singh B. Reciprocity in microbiome and immune system interactions and its implications in disease and health. *Inflamm Allergy Drug Targets* [Internet]. 2014 [cited 2018 Feb 5];13(2):94–104. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24678760>
  99. Kuhn KA, Stappenbeck TS. Peripheral education of the immune system by the colonic microbiota. *Semin Immunol* [Internet]. 2013 Nov 30 [cited 2018 Feb 5];25(5):364–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24169518>
  100. Lin L, Zhang J. Role of intestinal microbiota and metabolites on gut homeostasis and human diseases. *BMC Immunol* [Internet]. 2017 Dec 6 [cited 2018 Feb 5];18(1):2. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28061847>
  101. Meyer F, Paarmann D, D’Souza M, Olson R, Glass E, Kubal M, et al. The metagenomics

- RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* [Internet]. 2008 Sep 19 [cited 2018 Feb 12];9(1):386. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18803844>
102. Archibald JM. Endosymbiosis and Eukaryotic Cell Evolution. *Curr Biol* [Internet]. 2015 Oct 5 [cited 2018 Feb 8];25(19):R911–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26439354>
  103. O’Malley MA. Endosymbiosis and its implications for evolutionary theory: Table 1. *Proc Natl Acad Sci* [Internet]. 2015 Aug 18 [cited 2018 Feb 8];112(33):10270–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25883268>
  104. Lane N. Serial endosymbiosis or singular event at the origin of eukaryotes? *J Theor Biol* [Internet]. 2017 Dec 7 [cited 2018 Feb 8];434:58–67. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28501637>
  105. Emelyanov V V. Rickettsiaceae, rickettsia-like endosymbionts, and the origin of mitochondria. *Biosci Rep* [Internet]. 2001 Feb [cited 2018 Feb 8];21(1):1–17. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11508688>
  106. Mizrahi-Man O, Davenport ER, Gilad Y. Taxonomic Classification of Bacterial 16S rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. White BA, editor. *PLoS One* [Internet]. 2013 Jan 7 [cited 2018 Feb 7];8(1):e53608. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23308262>
  107. Singh AK, Kumar R, Singh M, Mishra AK, Chauhan UK, Baisvar VS, et al. Mitochondrial 16S rRNA gene-based evolutionary divergence and molecular phylogeny of *Barilius* spp. *Mitochondrial DNA* [Internet]. 2015 Jan 2 [cited 2018 Feb 7];26(1):41–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23883181>
  108. Mitani T, Akane A, Tokiyasu T, Yoshimura S, Okii Y, Yoshida M. Identification of animal species using the partial sequences in the mitochondrial 16S rRNA gene. *Leg Med* [Internet]. 2009 Apr [cited 2018 Feb 7];11:S449–50. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19289300>
  109. Kambhampati S, Smith PT. PCR primers for the amplification of four insect mitochondrial gene fragments. *Insect Mol Biol* [Internet]. 1995 Nov [cited 2018 Feb 7];4(4):233–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8825760>
  110. Yang L, Tan Z, Wang D, Xue L, Guan M-X, Huang T, et al. Species identification

- through mitochondrial rRNA genetic analysis. *Sci Rep* [Internet]. 2014 Feb 13 [cited 2018 Feb 7];4(1):4089. Available from: <http://www.nature.com/articles/srep04089>
111. Guo F, Ju F, Cai L, Zhang T. Taxonomic Precision of Different Hypervariable Regions of 16S rRNA Gene and Annotation Methods for Functional Bacterial Groups in Biological Wastewater Treatment. Hudson AO, editor. *PLoS One* [Internet]. 2013 Oct 16 [cited 2018 Feb 15];8(10):e76185. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24146837>
  112. Yang B, Wang Y, Qian P-Y. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* [Internet]. 2016 Dec 22 [cited 2017 Dec 6];17(1):135. Available from: <http://www.biomedcentral.com/1471-2105/17/135>
  113. Miyazaki K, Sato M, Tsukuda M. PCR Primer Design for 16S rRNAs for Experimental Horizontal Gene Transfer Test in *Escherichia coli*. *Front Bioeng Biotechnol* [Internet]. 2017 [cited 2018 Feb 13];5:14. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28293553>
  114. Bodilis J, Nsigure-Meilo S, Besaury L, Quillet L. Variable Copy Number, Intra-Genomic Heterogeneities and Lateral Transfers of the 16S rRNA Gene in *Pseudomonas*. Kao KC, editor. *PLoS One* [Internet]. 2012 Apr 24 [cited 2018 Feb 13];7(4):e35647. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22545126>
  115. Schirrmeister BE, Dalquen DA, Anisimova M, Bagheri HC. Gene copy number variation and its significance in cyanobacterial phylogeny. *BMC Microbiol* [Internet]. 2012 Aug 15 [cited 2018 Feb 13];12(1):177. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22894826>
  116. Kembel SW, Wu M, Eisen JA, Green JL. Incorporating 16S Gene Copy Number Information Improves Estimates of Microbial Diversity and Abundance. von Mering C, editor. *PLoS Comput Biol* [Internet]. 2012 Oct 25 [cited 2018 Feb 13];8(10):e1002743. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23133348>
  117. Angly FE, Dennis PG, Skarshewski A, Vanwonderghem I, Hugenholtz P, Tyson GW. CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome* [Internet]. 2014 [cited 2018 Feb 13];2(1):11. Available from:



- <http://www.ncbi.nlm.nih.gov/pubmed/24708850>
118. Tyler AD, Christianson S, Knox NC, Mabon P, Wolfe J, Van Domselaar G, et al. Comparison of Sample Preparation Methods Used for the Next-Generation Sequencing of *Mycobacterium tuberculosis*. PLoS One [Internet]. 2016 [cited 2018 Feb 9];11(2):e0148676. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26849565>
  119. Commins J, Toft C, Fares MA. Computational biology methods and their application to the comparative genomics of endocellular symbiotic bacteria of insects. Biol Proced Online [Internet]. 2009 Mar 11 [cited 2018 Feb 26];11:52–78. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19495914>
  120. Tyson GH, McDermott PF, Li C, Chen Y, Tadesse DA, Mukherjee S, et al. WGS accurately predicts antimicrobial resistance in *Escherichia coli*. J Antimicrob Chemother [Internet]. 2015 Oct [cited 2018 Feb 7];70(10):2763–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26142410>
  121. Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, Earl AM, et al. Relating the metatranscriptome and metagenome of the human gut. Proc Natl Acad Sci U S A [Internet]. 2014 Jun 3 [cited 2018 Feb 12];111(22):E2329-38. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24843156>
  122. Chun J, Lee J-H, Jung Y, Kim M, Kim S, Kim BK, et al. EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. Int J Syst Evol Microbiol [Internet]. 2007 Oct 1 [cited 2018 Feb 15];57(Pt 10):2259–61. Available from: <http://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.64915-0>
  123. Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. CABIOS Appl NOTE [Internet]. 1997 [cited 2018 Feb 16];13(5):555–6. Available from: <https://pdfs.semanticscholar.org/b550/be9ddf9579b120ad475bbcc654d2c00e7675.pdf>
  124. Guindon S, Lethiec F, Duroux P, Gascuel O. PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res [Internet]. 2005 Jul 1 [cited 2018 Feb 16];33(Web Server):W557–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15980534>
  125. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the

- performance of PhyML 3.0. *Syst Biol* [Internet]. 2010 May [cited 2014 Jul 9];59(3):307–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20525638>
126. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* [Internet]. 2014 May 1 [cited 2018 Feb 16];30(9):1312–3. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24451623>
  127. Hordijk W, Gascuel O. Improving the efficiency of SPR moves in phylogenetic tree search methods based on maximum likelihood. *Bioinformatics* [Internet]. 2005 Dec 15 [cited 2018 Feb 16];21(24):4338–47. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bti713>
  128. Chaudhary N, Sharma AK, Agarwal P, Gupta A, Sharma VK. 16S classifier: a tool for fast and accurate taxonomic classification of 16S rRNA hypervariable regions in metagenomic datasets. *PLoS One* [Internet]. 2015 [cited 2018 Feb 15];10(2):e0116106. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25646627>
  129. Ghosh TS, Gajjala P, Mohammed MH, Mande SS. C16S — A Hidden Markov Model based algorithm for taxonomic classification of 16S rRNA gene sequences. *Genomics* [Internet]. 2012 Apr [cited 2018 Feb 15];99(4):195–201. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22326741>
  130. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl Environ Microbiol* [Internet]. 2006 Jul 1 [cited 2018 Feb 16];72(7):5069–72. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16820507>
  131. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* [Internet]. 2014 Jan [cited 2018 Feb 16];42(Database issue):D633–42. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24288368>
  132. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* [Internet]. 2012 Nov 27 [cited 2018 Feb 16];41(D1):D590–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23193283>
  133. Martinez X, Pozuelo M, Pascal V, Campos D, Gut I, Gut M, et al. MetaTrans: an open-

- source pipeline for metatranscriptomics. *Sci Rep* [Internet]. 2016 Sep 23 [cited 2018 Feb 20];6(1):26447. Available from: <http://www.nature.com/articles/srep26447>
134. Westreich ST, Korf I, Mills DA, Lemay DG. SAMSA: a comprehensive metatranscriptome analysis pipeline. *BMC Bioinformatics* [Internet]. 2016 Dec 29 [cited 2018 Feb 20];17(1):399. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27687690>
  135. Ni Y, Li J, Panagiotou G. COMAN: a web server for comprehensive metatranscriptomics analysis. *BMC Genomics* [Internet]. 2016 [cited 2018 Feb 20];17(1):622. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27515514>
  136. Anderson JW, Baird P, Davis Jr RH, Ferreri S, Knudtson M, Koraym A, et al. Health benefits of dietary fiber. *Nutr Rev* [Internet]. 2009 Apr [cited 2017 Nov 20];67(4):188–205. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19335713>
  137. Marlett JA, McBurney MI, Slavin JL, American Dietetic Association. Position of the American Dietetic Association: health implications of dietary fiber. *J Am Diet Assoc* [Internet]. 2002 Jul [cited 2017 Nov 20];102(7):993–1000. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12146567>
  138. Trowell H. Ischemic heart disease and dietary fiber. *Am J Clin Nutr* [Internet]. 1972 Sep 1 [cited 2017 Nov 20];25(9):926–32. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/4559894>
  139. Braaten JT, Wood PJ, Scott FW, Wolynetz MS, Lowe MK, Bradley-White P, et al. Oat beta-glucan reduces blood cholesterol concentration in hypercholesterolemic subjects. *Eur J Clin Nutr* [Internet]. 1994 Jul [cited 2017 Nov 20];48(7):465–74. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7956987>
  140. Bell LP, Hectorne K, Reynolds H, Balm TK, Hunninghake DB. Cholesterol-lowering effects of psyllium hydrophilic mucilloid. Adjunct therapy to a prudent diet for patients with mild to moderate hypercholesterolemia. *JAMA* [Internet]. 1989 Jun 16 [cited 2017 Nov 20];261(23):3419–23. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2724486>
  141. Anderson JW, Davidson MH, Blonde L, Brown W V, Howard WJ, Ginsberg H, et al. Long-term cholesterol-lowering effects of psyllium as an adjunct to diet therapy in the treatment of hypercholesterolemia. *Am J Clin Nutr* [Internet]. 2000 Jun [cited 2017 Nov

- 20];71(6):1433–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10837282>
142. Anderson JW, Zeigler JA, Deakins DA, Floore TL, Dillon DW, Wood CL, et al. Metabolic effects of high-carbohydrate, high-fiber diets for insulin-dependent diabetic individuals. *Am J Clin Nutr* [Internet]. 1991 Nov [cited 2017 Nov 20];54(5):936–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1659172>
143. Anderson JW, Story L, Sieling B, Chen WJ, Petro MS, Story J. Hypocholesterolemic effects of oat-bran or bean intake for hypercholesterolemic men. *Am J Clin Nutr* [Internet]. 1984 Dec [cited 2017 Nov 20];40(6):1146–55. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/6095635>
144. Naumann E, van Rees AB, Onning G, Oste R, Wydra M, Mensink RP. Beta-glucan incorporated into a fruit drink effectively lowers serum LDL-cholesterol concentrations. *Am J Clin Nutr* [Internet]. 2006 Mar [cited 2017 Nov 20];83(3):601–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16522906>
145. Morita S. Dietary Fiber and Heavy Metal. *SEIKATSU EISEI (Journal Urban Living Heal Assoc)* [Internet]. 1983 Jan 10 [cited 2017 Nov 20];27(1):22–9. Available from: [https://www.jstage.jst.go.jp/article/seikatsueisei1957/27/1/27\\_1\\_22/\\_article](https://www.jstage.jst.go.jp/article/seikatsueisei1957/27/1/27_1_22/_article)
146. Rose HE, Quarterman J. Dietary fibers and heavy metal retention in the rat. *Environ Res* [Internet]. 1987 Feb [cited 2017 Nov 20];42(1):166–75. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/3026796>
147. Nawirska A, Kwaśniewska M. Dietary fibre fractions from fruit and vegetable processing waste. *Food Chem* [Internet]. 2005 Jun 1 [cited 2017 Nov 20];91(2):221–5. Available from: <http://www.sciencedirect.com/science/article/pii/S0308814603005314>
148. Karczewski J, Poniedziałek B, Adamski Z, Rzymiski P. The effects of the microbiota on the host immune system. *Autoimmunity* [Internet]. 2014 Dec 14 [cited 2017 Nov 20];47(8):494–504. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25019177>
149. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Nageshwar Reddy D. Role of the normal gut microbiota. *World J Gastroenterol* [Internet]. 2015 Aug 7 [cited 2017 Nov 20];21(29):8787. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26269668>
150. Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA, Wolter M, et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and

- Enhances Pathogen Susceptibility. *Cell* [Internet]. 2016 Nov 17 [cited 2017 Nov 10];167(5):1339-1353.e21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27863247>
151. Belzer C, de Vos WM. Microbes inside—from diversity to function: the case of *Akkermansia*. *ISME J* [Internet]. 2012 Aug 22 [cited 2018 Feb 26];6(8):1449–58. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22437156>
  152. Foretz M, Hébrard S, Leclerc J, Zarrinpashneh E, Soty M, Mithieux G, et al. Metformin inhibits hepatic gluconeogenesis in mice independently of the LKB1/AMPK pathway via a decrease in hepatic energy state. *J Clin Invest* [Internet]. 2010 Jul 1 [cited 2017 Nov 22];120(7):2355–69. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20577053>
  153. Sliwinska A, Drzewoski J. Molecular action of metformin in hepatocytes: an updated insight. *Curr Diabetes Rev* [Internet]. 2015 [cited 2017 Nov 22];11(3):175–81. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25808533>
  154. Bradley CA. Gut microbiota: Trust your gut — metformin and diabetes. *Nat Rev Endocrinol* [Internet]. 2017 Jun 2 [cited 2017 Sep 8];13(8):440–440. Available from: <http://www.nature.com/doi/10.1038/nrendo.2017.72>
  155. Buse JB, DeFronzo RA, Rosenstock J, Kim T, Burns C, Skare S, et al. The Primary Glucose-Lowering Effect of Metformin Resides in the Gut, Not the Circulation. Results From Short-term Pharmacokinetic and 12-Week Dose-Ranging Studies. *Diabetes Care* [Internet]. 2015 Aug 18 [cited 2017 Sep 8];39(2):dc150488. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26285584>
  156. Wu H, Esteve E, Tremaroli V, Khan MT, Caesar R, Mannerås-Holm L, et al. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. *Nat Med* [Internet]. 2017 May 22 [cited 2018 Feb 22];23(7):850–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28530702>
  157. Napolitano A, Miller S, Nicholls AW, Baker D, Van Horn S, Thomas E, et al. Novel Gut-Based Pharmacology of Metformin in Patients with Type 2 Diabetes Mellitus. Zhu Z, editor. *PLoS One* [Internet]. 2014 Jul 2 [cited 2018 Feb 22];9(7):e100778. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24988476>
  158. Shin N-R, Lee J-C, Lee H-Y, Kim M-S, Whon TW, Lee M-S, et al. An increase in the

- Akkermansia* spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. Gut [Internet]. 2014 May [cited 2017 Nov 22];63(5):727–35. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23804561>
159. Lee H, Ko G. Effect of Metformin on Metabolic Improvement and Gut Microbiota. Griffiths MW, editor. Appl Environ Microbiol [Internet]. 2014 Oct 1 [cited 2018 Feb 22];80(19):5935–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25038099>
160. Langan EA, Griffiths CEM, Solbach W, Knobloch JK, Zillikens D, Thaçi D. The role of the microbiome in psoriasis: moving from disease description to treatment prediction? Br J Dermatol [Internet]. 2017 Oct 26 [cited 2018 Feb 26]; Available from: <http://doi.wiley.com/10.1111/bjd.16081>
161. Zhou XJ, Rahmani R. Preclinical and clinical pharmacology of vinca alkaloids. Drugs [Internet]. 1992 [cited 2017 Nov 21];44 Suppl 4:1–16; discussion 66-9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1283846>
162. Huber R, Hartmann M, Bliesath H, Lühmann R, Steinijans VW, Zech K. Pharmacokinetics of pantoprazole in man. Int J Clin Pharmacol Ther [Internet]. 1996 May [cited 2017 Nov 21];34(1 Suppl):S7-16. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8793599>
163. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. Nat Rev Microbiol [Internet]. 2016 Mar 14 [cited 2017 Nov 27];14(5):273–87. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26972811>
164. Caruso G. Antibiotic Resistance in *Escherichia coli* from Farm Livestock and Related Analytical Methods: A Review. J AOAC Int [Internet]. 2018 Jul 1 [cited 2019 Apr 19];101(4):916–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29554996>
165. Ruuskanen M, Muurinen J, Meierjohan A, Pärnänen K, Tamminen M, Lyra C, et al. Fertilizing with Animal Manure Disseminates Antibiotic Resistance Genes to the Farm Environment. J Environ Qual [Internet]. 2016 Mar [cited 2019 Apr 19];45(2):488. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27065395>
166. Hu H-W, Han X-M, Shi X-Z, Wang J-T, Han L-L, Chen D, et al. Temporal changes of antibiotic-resistance genes and bacterial communities in two contrasting soils treated with cattle manure. Smalla K, editor. FEMS Microbiol Ecol [Internet]. 2016 Feb [cited 2019

- Apr 19];92(2):fiv169. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26712351>
167. Zhu Y-G, Zhao Y, Li B, Huang C-L, Zhang S-Y, Yu S, et al. Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat Microbiol* [Internet]. 2017 Apr 30 [cited 2019 Apr 20];2(4):16270. Available from: <http://www.nature.com/articles/nmicrobiol2016270>
168. Almakki A, Jumas-Bilak E, Marchandin H, Licznar-Fajardo P. Antibiotic resistance in urban runoff. *Sci Total Environ* [Internet]. 2019 Jun 1 [cited 2019 Apr 19];667:64–76. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30826682>
169. Manaia CM, Rocha J, Scaccia N, Marano R, Radu E, Biancullo F, et al. Antibiotic resistance in wastewater treatment plants: Tackling the black box. *Environ Int* [Internet]. 2018 Jun [cited 2019 Apr 19];115:312–24. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0160412017322328>
170. Rizzo L, Manaia C, Merlin C, Schwartz T, Dagot C, Ploy MC, et al. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Sci Total Environ* [Internet]. 2013 Mar 1 [cited 2019 Apr 19];447:345–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23396083>
171. Issa I, Moucari R. Probiotics for antibiotic-associated diarrhea: Do we have a verdict? *World J Gastroenterol* [Internet]. 2014 Dec 21 [cited 2019 Apr 20];20(47):17788–95. Available from: <http://www.wjgnet.com/1007-9327/full/v20/i47/17788.htm>
172. Banerjee S, Adcock L. Probiotics for Antibiotic-Associated Diarrhea and *Clostridium difficile* Infection: A Review of Guidelines [Internet]. Probiotics for Antibiotic-Associated Diarrhea and *Clostridium difficile* Infection: A Review of Guidelines. Canadian Agency for Drugs and Technologies in Health; 2018 [cited 2019 Apr 20]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30817102>
173. Bignardi GE. Risk factors for *Clostridium difficile* infection. *J Hosp Infect* [Internet]. 1998 Sep [cited 2019 Apr 20];40(1):1–15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9777516>
174. Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci* [Internet]. 2011 Mar 15 [cited 2019 Apr 19];108(Supplement\_1):4554–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20847294>

175. Dethlefsen L, Huse S, Sogin ML, Relman DA. The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing. Eisen JA, editor. *PLoS Biol* [Internet]. 2008 Nov 18 [cited 2019 Apr 19];6(11):e280. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19018661>
176. Francino MP. Antibiotics and the Human Gut Microbiome: Dysbioses and Accumulation of Resistances. *Front Microbiol* [Internet]. 2016 Jan 12 [cited 2019 Apr 8];6:1543. Available from: <http://journal.frontiersin.org/Article/10.3389/fmicb.2015.01543/abstract>
177. Kantele A, Laaveri T, Mero S, Vilkinan K, Pakkanen SH, Ollgren J, et al. Antimicrobials Increase Travelers' Risk of Colonization by Extended-Spectrum Betalactamase-Producing Enterobacteriaceae. *Clin Infect Dis* [Internet]. 2015 Mar 15 [cited 2019 Apr 19];60(6):837–46. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25613287>
178. Björkholm B, Bok CM, Lundin A, Rafter J, Hibberd ML, Pettersson S. Intestinal Microbiota Regulate Xenobiotic Metabolism in the Liver. Bereswill S, editor. *PLoS One* [Internet]. 2009 Sep 9 [cited 2018 Feb 26];4(9):e6958. Available from: <http://dx.plos.org/10.1371/journal.pone.0006958>
179. Alexander JL, Wilson ID, Teare J, Marchesi JR, Nicholson JK, Kinross JM. Gut microbiota modulation of chemotherapy efficacy and toxicity. *Nat Rev Gastroenterol Hepatol* [Internet]. 2017 Mar 8 [cited 2017 Sep 8];14(6):356–65. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28270698>
180. Zheng X, Zhao A, Xie G, Chi Y, Zhao L, Li H, et al. Melamine-Induced Renal Toxicity Is Mediated by the Gut Microbiota. *Sci Transl Med* [Internet]. 2013 Feb 13 [cited 2017 Nov 24];5(172):172ra22-172ra22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23408055>
181. Scimone C, Donato L, Rinaldi C, Sidoti A, D'Angelo R. First case of Currarino syndrome and trimethylaminuria: two rare diseases for a complex clinical presentation. *J Dig Dis* [Internet]. 2016 Sep [cited 2017 Nov 27];17(9):628–32. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27335202>
182. Sabir N, Jones EA, Padmakumar B. Trimethylaminuria. *BMJ Case Rep* [Internet]. 2016 Apr 26 [cited 2017 Nov 27];2016:bcr2015213742. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27118741>
183. Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I, Rosenberg E. Commensal



- bacteria play a role in mating preference of *Drosophila melanogaster*. *Proc Natl Acad Sci* [Internet]. 2010 Nov 16 [cited 2017 Dec 5];107(46):20051–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21041648>
184. Sharon G, Segal D, Zilber-Rosenberg I, Rosenberg E. Symbiotic bacteria are responsible for diet-induced mating preference in *Drosophila melanogaster*, providing support for the hologenome concept of evolution. *Gut Microbes* [Internet]. [cited 2017 Dec 5];2(3):190–2. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21804354>
  185. Friend DG. Digitalis after two centuries (William Withering). *Arch Surg* [Internet]. 1976 Jan [cited 2018 Feb 28];111(1):14–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1106355>
  186. Lee TC. Van Gogh's vision. Digitalis intoxication? *JAMA* [Internet]. 1981 Feb 20 [cited 2018 Feb 28];245(7):727–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7007674>
  187. Elliott DB, Skaff A. Vision of the famous: the artist's eye. *Ophthalmic Physiol Opt* [Internet]. 1993 Jan [cited 2018 Feb 28];13(1):82–90. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8510952>
  188. Wallick D, Stuesse SL, Masuda Y, Levy MN. Effects of digoxin on the control of heart rate and atrioventricular conduction in the dog. *Cardiovasc Res* [Internet]. 1983 Jul [cited 2018 Mar 5];17(7):400–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/6883415>
  189. MacLeod-Glover N, Mink M, Yarema M, Chuang R. Digoxin toxicity: Case for retiring its use in elderly patients? *Can Fam Physician* [Internet]. 2016 Mar [cited 2018 Mar 5];62(3):223–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26975913>
  190. Bismuth C, Gaultier M, Conso F, Efthymiou ML. Hyperkalemia in Acute Digitalis Poisoning: Prognostic Significance and Therapeutic implications. *Clin Toxicol* [Internet]. 1973 Jan 25 [cited 2018 Mar 5];6(2):153–62. Available from: <http://www.tandfonline.com/doi/full/10.3109/15563657308990513>
  191. Kanji S, MacLean RD. Cardiac Glycoside Toxicity. *Crit Care Clin* [Internet]. 2012 Oct [cited 2018 Mar 5];28(4):527–35. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22998989>
  192. Levine M, Nikkanen H, Pallin DJ. The Effects of Intravenous Calcium in Patients with

- Digoxin Toxicity. *J Emerg Med* [Internet]. 2011 Jan [cited 2018 Mar 5];40(1):41–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19201134>
193. Roever C, Ferrante J, Gonzalez EC, Pal N, Roetzheim RG. Comparing the toxicity of digoxin and digitoxin in a geriatric population: should an old drug be rediscovered? *South Med J* [Internet]. 2000 Feb [cited 2018 Mar 5];93(2):199–202. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10701788>
  194. Pincus M. Management of digoxin toxicity. *Aust Prescr* [Internet]. 2016 Feb [cited 2017 Sep 14];39(1):18–20. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27041802>
  195. Moffett BS, Garner A, Zapata T, Orcutt J, Niu M, Lopez KN. Serum digoxin concentrations and clinical signs and symptoms of digoxin toxicity in the paediatric population. *Cardiol Young* [Internet]. 2016 Mar 27 [cited 2018 Mar 5];26(03):493–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25912244>
  196. van VELDHUISEN DJ, de GRAEFF PA, REMME WJ, LIE KI. Value of Digoxin in Heart Failure and Sinus Rhythm: New Features of an Old Drug? *J Am Coll Cardiol* [Internet]. 1996 Oct 1 [cited 2018 Mar 5];28(4):813–9. Available from: <https://www.sciencedirect.com/science/article/pii/S0735109796002471>
  197. Group TDI. The Effect of Digoxin on Mortality and Morbidity in Patients with Heart Failure. *N Engl J Med* [Internet]. 1997 Feb 20 [cited 2018 Mar 5];336(8):525–33. Available from: <http://www.nejm.org/doi/abs/10.1056/NEJM199702203360801>
  198. Cosmi F, Tarquini B, Mariottoni B, Cosmi D. La digitale: un farmaco da rottamare? *doi.org* [Internet]. 2017 Feb [cited 2017 Sep 8];18(2017Febbraio):121–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28398365>
  199. Gheorghide M, Patel K, Filippatos G, Anker SD, van Veldhuisen DJ, Cleland JGF, et al. Effect of oral digoxin in high-risk heart failure patients: a pre-specified subgroup analysis of the DIG trial. *Eur J Heart Fail* [Internet]. 2013 May [cited 2018 Mar 5];15(5):551–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23355060>
  200. Hopper I, Skiba M, von Lueder TG, Watanabe M, Funston R, Tonkin A, et al. Digoxin Withdrawal Worsens Clinical Status in Stable Patients With Heart Failure Receiving Optimal Contemporaneous Therapy—A Randomized Controlled Trial. *J Card Fail* [Internet]. 2015 Sep [cited 2018 Mar 5];21(9):779–81. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26169443>

201. Packer M, Gheorghide M, Young JB, Costantini PJ, Adams KF, Cody RJ, et al. Withdrawal of Digoxin from Patients with Chronic Heart Failure Treated with Angiotensin-Converting-Enzyme Inhibitors. *N Engl J Med* [Internet]. 1993 Jul 1 [cited 2018 Mar 5];329(1):1–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8505940>
202. Fair JF. Supervised withdrawal of long-term digoxin therapy. *Fam Pract* [Internet]. 1990 Mar [cited 2018 Mar 5];7(1):56–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2318375>
203. Rathore SS, Curtis JP, Wang Y, Bristow MR, Krumholz HM. Association of serum digoxin concentration and outcomes in patients with heart failure. *JAMA* [Internet]. 2003 Feb 19 [cited 2018 Mar 5];289(7):871–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12588271>
204. Goldberger ZD, Goldberger AL. Therapeutic ranges of serum digoxin concentrations in patients with heart failure. *Am J Cardiol* [Internet]. 2012 Jun 15 [cited 2018 Mar 5];109(12):1818–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22502901>
205. Ahmed A, Gambassi G, Weaver MT, Young JB, Wehrmacher WH, Rich MW. Effects of Discontinuation of Digoxin Versus Continuation at Low Serum Digoxin Concentrations in Chronic Heart Failure. *Am J Cardiol* [Internet]. 2007 Jul 15 [cited 2018 Mar 5];100(2):280–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17631083>
206. Digoxin: serious drug interactions. *Prescrire Int* [Internet]. 2010 Apr [cited 2018 Mar 5];19(106):68–70. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20568489>
207. Juliano RL, Ling V. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim Biophys Acta - Biomembr* [Internet]. 1976 Nov 11 [cited 2018 Mar 5];455(1):152–62. Available from: <https://www.sciencedirect.com/science/article/pii/0005273676901607?via%3Dihub>
208. Dobkin JF, Saha JR, Butler VP, Neu HC, Lindenbaum J. Inactivation of digoxin by *Eubacterium lentum*, an anaerobe of the human gut flora. *Trans Assoc Am Physicians* [Internet]. 1982 [cited 2018 Mar 6];95:22–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7182977>
209. Haiser HJ, Seim KL, Balskus EP, Turnbaugh PJ. Mechanistic insight into digoxin inactivation by *Eggerthella lenta* augments our understanding of its pharmacokinetics.

- Gut Microbes [Internet]. 2014 [cited 2017 Nov 23];5(2):233–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24637603>
210. Burton ME. Applied pharmacokinetics & pharmacodynamics : principles of therapeutic drug monitoring [Internet]. Lippincott Williams & Wilkins; 2006 [cited 2018 Mar 7]. 867 p. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1805474/>
211. Liston HL, Markowitz JS, DeVane CL. Drug glucuronidation in clinical psychopharmacology. *J Clin Psychopharmacol* [Internet]. 2001 Oct [cited 2018 Mar 7];21(5):500–15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11593076>
212. Meyer JM. Individual changes in clozapine levels after smoking cessation: results and a predictive model. *J Clin Psychopharmacol* [Internet]. 2001 Dec [cited 2018 Mar 7];21(6):569–74. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11763003>
213. Scheline RR. CRC handbook of mammalian metabolism of plant compounds. CRC PRESS; 2017.
214. Lindenbaum J, Rund DG, Butler VP, Tse-Eng D, Saha JR. Inactivation of Digoxin by the Gut Flora: Reversal by Antibiotic Therapy. *N Engl J Med* [Internet]. 1981 Oct 1 [cited 2017 Feb 1];305(14):789–94. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7266632>
215. Haiser HJ, Gootenberg DB, Chatman K, Sirasani G, Balskus EP, Turnbaugh PJ. Predicting and manipulating cardiac drug inactivation by the human gut bacterium *Eggerthella lenta*. *Science* [Internet]. 2013 Jul 19 [cited 2018 Mar 8];341(6143):295–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23869020>
216. Lu L, Wu Y, Zuo L, Luo X, Large PJ. Intestinal microbiome and digoxin inactivation: meal plan for digoxin users? *World J Microbiol Biotechnol* [Internet]. 2014 Mar 9 [cited 2018 Mar 8];30(3):791–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24105082>
217. Raymond F, Ouameur AA, Déraspe M, Iqbal N, Gingras H, Dridi B, et al. The initial state of the human gut microbiome determines its reshaping by antibiotics. *ISME J* [Internet]. 2016 Mar [cited 2017 Aug 30];10(3):707. Available from: </pmcc/articles/PMC4817689/?report=abstract>
218. Aagaard K, Petrosino J, Keitel W, Watson M, Katancik J, Garcia N, et al. The Human

- Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. *FASEB J* [Internet]. 2013 Mar [cited 2018 Mar 19];27(3):1012–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23165986>
219. Ma J, Prince A, Aagaard KM. Use of whole genome shotgun metagenomics: a practical guide for the microbiome-minded physician scientist. *Semin Reprod Med* [Internet]. 2014 Jan [cited 2016 Sep 19];32(1):5–13. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24390915>
220. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol* [Internet]. 2011 Jan [cited 2018 May 14];29(1):24–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21221095>
221. Hinderling PH, Hartmann D. Pharmacokinetics of digoxin and main metabolites/derivatives in healthy humans. *Ther Drug Monit* [Internet]. 1991 Sep [cited 2018 Jan 23];1316(5):381–401. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1957331>
222. Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* [Internet]. 2018;555(7698):623–8. Available from: <http://dx.doi.org/10.1038/nature25979>