



HAL
open science

**Exploring the metabolic mechanisms used by
Enterobacteriaceae to thrive in the dysbiotic gut of
patients suffering from intestinal or extra-intestinal
inflammatory diseases**

Maria Ines Moreira de Gouveia

► **To cite this version:**

Maria Ines Moreira de Gouveia. Exploring the metabolic mechanisms used by Enterobacteriaceae to thrive in the dysbiotic gut of patients suffering from intestinal or extra-intestinal inflammatory diseases. Human health and pathology. Université Clermont Auvergne, 2022. English. NNT : 2022UCFAC089 . tel-04683828

HAL Id: tel-04683828

<https://theses.hal.science/tel-04683828>

Submitted on 2 Sep 2024

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

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

**ÉCOLE DOCTORALE DES
SCIENCES DE LA VIE, SANTÉ, AGRONOMIE, ENVIRONNEMENT**
N° d'ordre :

Thèse
Présentée à l'Université Clermont-Auvergne pour l'obtention du grade de
DOCTEUR D'UNIVERSITÉ

Spécialité
Génétique, Physiologie, Pathologie, Nutrition, Microbiologie, Santé et Innovation

Soutenue le 8 Décembre 2022
MARIA INÊS MOREIRA DE GOUVEIA

**EXPLORING THE METABOLIC MECHANISMS USED BY
ENTEROBACTERIACEAE TO THRIVE IN THE DYSBIOTIC GUT
OF PATIENTS SUFFERING FROM INTESTINAL OR EXTRA-
INTESTINAL INFLAMMATORY DISEASES**

Rapporteurs	Nalini RAMA RAO, DR, UMR MICALIS, INRAE Muriel COCAIGN-BOUSQUET, DR, UMR TBI, INRAE
Membres	Julien DELMAS, MCU-PH, UMR M2iSH, UCA Thomas HINDRE, MCF, TIMC, Grenoble Alpes Julien SCANZI, PH, Centre Hospitalier Thiers
Directeur	Grégory JUBELIN, CR, UMR MEDIS, INRAE
Co-directrice	Annick BERNALIER-DONADILLE, DR, UMR MEDIS, INRAE

Para a minha Tia Lurdes

ACKNOWLEDGEMENTS

Quelle aventure... en regardant mon chemin lors de ces 3 dernières années, je me rends compte que j'ai eu une énorme chance de l'avoir parcouru si bien entourée. Le chemin n'aurait pas été pareil sans toutes ces merveilleuses personnes qui m'ont appris tellement de choses et qui m'ont poussé vers l'infini et au-delà.

J'aimerais commencer par remercier les membres du jury Nailini Rama Rao, Muriel Cocaign-Bousquet, Thomas Hindre et Julien Scanzi d'avoir accepté d'évaluer mon travail et être présents lors de ma soutenance de thèse. Je voudrais tout particulièrement remercier Julien Delmas pour son suivi pendant ces 3 ans lors de mes comités de thèse et d'avoir accepté d'être membre du jury.

Tous mes sincères remerciements vont à mon directeur de thèse Gregory Jubelin. Je n'ai pas assez de mots pour remercier tout ce que tu m'as apporté pendant ces 3 ans de dur travail. Merci pour ton encadrement, ta confiance, tes conseils scientifiques et tes grandes qualités humaines (notamment la patience). Merci de m'avoir guidé quand j'étais perdue, de m'avoir formé à de nombreuses techniques et de m'avoir soutenu. Un grand merci à ma co-directrice Annick Bernalier-Donadille pour tous tes conseils si précieux, ton optimisme et ton expertise dans le domaine. Je vous en serai toujours reconnaissante pour votre soutien et encouragements lors de ces 3 ans.

Un grand merci à Julien. Merci pour tes conseils techniques (toujours d'une grande précision), ton aide précieuse sans faille et ton soutien. Je n'oublierai pas tout le temps passé dans le P2 que tu as su rendre plus joyeux avec des imitations d'animaux, tes choix de musique parfois douteux, et les débats (parfois sans fin). Un grand merci également à Annie. Merci pour toute ton aide au cours de ces 3 ans de travail notamment sur les manips d'urgence mais aussi lors de la chasse aux tubes et produits introuvables dans le labo. Merci pour ta bonne humeur et tes mots toujours chaleureux. Je voudrais vous remercier tous les deux pour votre efficacité, sans vous ma thèse n'aurait pas avancé comme elle l'a fait. Je souhaite aussi remercier Valérie pour toute son aide lors des manips avec les « *choupinous* » à l'animalerie. Longue vie à Pirate ! Je n'oublierai pas ta bonne humeur et aussi tout ce que j'ai appris à l'animalerie avec toi, Greg et Julien.

Je voudrais remercier l'ensemble de l'équipe MEDIS mais particulièrement Gloglo, Yacine, Laurie, Lysiane, Arlette, Jeanne, Christophe, Eve et Philippe. Merci de m'avoir aidé à plusieurs reprises, merci pour tous vos conseils, votre patience avec moi, vos dépannages urgents de produits ou de kits et surtout pour votre bonne humeur et joie de vivre. J'ai beaucoup apprécié nos conversations Yacine et Laurie, qui à chaque fois me mettaient instantanément de bonne humeur. Je n'oublierai pas Evelyne et Frédérique. Merci pour vos conseils, votre sourire, bonne humeur et toutes les histoires que vous partagez avec nous. Merci à TOUS pour les bons moments passés au labo, à la pause-café, et à la cantine que je n'oublierai jamais !

Merci beaucoup Lucie, Nelly et Sabine pour votre aide quand j'ai eu des problèmes de biologie moléculaire. Je vous remercie de m'avoir transmis plein de conseils, protocoles et savoir technique dans le domaine. Je n'oublierai pas ta bonne humeur Sabine et du fait que tu m'as donné envie de déménager à Montpellier. Merci aussi Stephanie pour ton aide lors que j'en avais besoin et de m'avoir motivé à aller au yoga.

Un merci très spécial à mes colocos de bureau Bob, Raphaele et Audrey. Ils me supportent quotidiennement dans le bureau... ils ont du mérite ! Merci Bob d'avoir été là pour moi depuis le début, nous avons commencé cette aventure ensemble et ça n'aurait pas été pareil sans toi. Merci pour toutes tes petites attentions, ta bonne humeur et ton écoute. Je n'oublierai jamais tout ce que tu as fait pour moi. Je voudrais aussi remercier ma Raphou, merci infiniment pour tes mots d'encouragement et pour m'avoir guidé pendant les moments difficiles notamment l'écriture du manuscrit. Tes conseils ont été très précieux pour moi. Merci de m'avoir nourri, d'avoir chanté avec moi, dansé avec moi, miaulé avec moi, secoué ma chaise quand j'en avais besoin, et surtout merci de m'avoir amené à Disney !

Je remercie mes amis notamment Nono qui a suivi de loin ce parcours mais pourtant qui a toujours été présente quand j'avais besoin. Mon Thibaut, maintenant docteur on s'est épaulé mutuellement tout le long de cette aventure... merci d'avoir toujours été là pour moi sans faille dans les bons et mauvais moments. Un grand merci à mes soeurcières Marine et Orcylia de m'avoir soutenu inconditionnellement. Je n'oublierai jamais tout ce que vous avez fait pour moi, merci pour votre compréhension et toute la force que vous m'avez donné pendant ces 3 ans. Amélie, ton ami c'est moi ! Je n'ai pas assez de mots pour remercier tout le soutien que tu m'as apporté. Je suis si heureuse d'avoir croisé ton chemin... Merci de m'avoir écouté quand j'avais besoin, merci pour toutes tes petites attentions, ta bonne humeur, ton optimisme. Tu m'as donné le sourire quand il fallait. Merci aussi Guerric pour ta présence tout au long de ce chemin, merci de m'avoir aidé quand j'en avais besoin.

Je remercie tout particulièrement toute ma famille ici et au Portugal. Je n'aurais pas réussi sans votre soutien Mãe, Pai et João je vous aime infiniment et je ne pourrais jamais assez vous remercier. Je voudrais aussi remercier ma Tia Lurdes qui n'est plus parmi nous. Elle a fait de moi qui je suis aujourd'hui. Sans elle je ne serais pas en train d'écrire cette thèse.

“ Laughter is timeless, imagination has no age, dreams are forever. ”

- Walt Disney

Identification de propriétés métaboliques utilisées par les *Enterobacteriaceae* pour proliférer dans le tractus digestif de patients souffrant de maladies inflammatoires intestinales ou extra-intestinales

RÉSUMÉ

Un déséquilibre du microbiote intestinal appelé dysbiose a été associé à plusieurs maladies inflammatoires telles que la maladie de Crohn (MC) et la polyarthrite rhumatoïde (PR). La dysbiose est souvent caractérisée par une expansion des *Enterobacteriaceae*, une famille bactérienne qui ne représente normalement qu'un petit pourcentage du microbiote intestinal chez les individus sains (SA). Pour expliquer la prolifération de cette famille bactérienne au cours de la dysbiose, une hypothèse repose sur l'adaptation métabolique des *Enterobacteriaceae* à l'environnement intestinal perturbé. L'objectif de mon travail de doctorat était de cribler des substrats nutritionnels qui pourraient donner un avantage métabolique aux entérobactéries dans l'intestin dysbiotique de patients souffrant de MC ou de PR. À l'aide d'une technique de criblage phénotypique à haut débit, l'utilisation de 190 substrats différents comme source de carbone a été testée. La capacité à utiliser ces différents substrats a été comparée entre les souches d'entérobactérie isolées à partir d'échantillons fécaux d'SA ou de patients souffrant de MC ou de PR. Parmi les substrats identifiés, nous avons observé une capacité opposée des souches SA et MC à métaboliser la D-sérine et le saccharose, une caractéristique résultant de réarrangements génétiques dans le locus *argW* du chromosome d'*Escherichia coli*. Il a été démontré que le phénotype D-sérine⁺ des souches MC dépendait de la présence de gènes *dsdCXA* et conférait un avantage nutritionnel par rapport aux souches SA *in vitro*. Nous avons également évalué la capacité de nos collections de souches à cataboliser l'éthanolamine (EA), une petite molécule organique naturellement présente dans l'intestin humain avec une concentration accrue en cas d'inflammation intestinale. Alors qu'aucune différence spécifique n'a été observée entre les 3 collections de souches, la capacité à consommer ce substrat s'est révélée être très diversifiée entre les isolats. Les souches incapables d'utiliser l'EA présentent soit des mutations dans le locus *eut* responsable de la dégradation de l'EA, soit un défaut dans le processus de transcription de l'opéron. Les isolats consommant de l'EA peuvent utiliser ce substrat comme source d'azote uniquement ou comme source d'azote et de carbone. À l'aide d'un modèle murin de colonisation intestinale et d'un nouvel outil rapporteur de gènes développé au cours de la thèse, nous avons également démontré que les gènes responsables de la consommation de EA sont exprimés lors de la colonisation du tractus digestif et que l'utilisation de ce substrat est essentielle pour un fitness optimal des *E. coli* commensaux *in vivo*. Avec cet exemple de l'EA, mon travail met en évidence la grande diversité de phénotypes présents entre les isolats commensaux intestinaux de l'espèce *E. coli*, illustrant comment les entérobactéries sont une source majeure de la variabilité génétique constatée dans le microbiome intestinal humain. Aussi, notre étude ouvre de nouvelles pistes quant à des substrats qui pourraient favoriser l'expansion des entérobactéries dans des conditions dysbiotiques et pourraient ouvrir de nouvelles perspectives sur des approches thérapeutiques afin de restaurer l'homéostasie intestinale et la santé des patients.

INTRODUCTION BIBLIOGRAPHIQUE

CHAPITRE 1 : Le microbiote intestinal chez l'individu sain

Toute surface du corps humain en contact avec l'environnement est colonisée par des micro-organismes. Parmi les différents organes humains, le tractus gastro-intestinal (TGI) est le plus colonisé, avec 10^{14} **microorganismes** qui constituent le microbiote intestinal. Le TGI est composé de différentes parties : l'œsophage, l'estomac, l'intestin grêle et le côlon qui possèdent toutes un microbiote spécifique et différent (**Figure 1**).

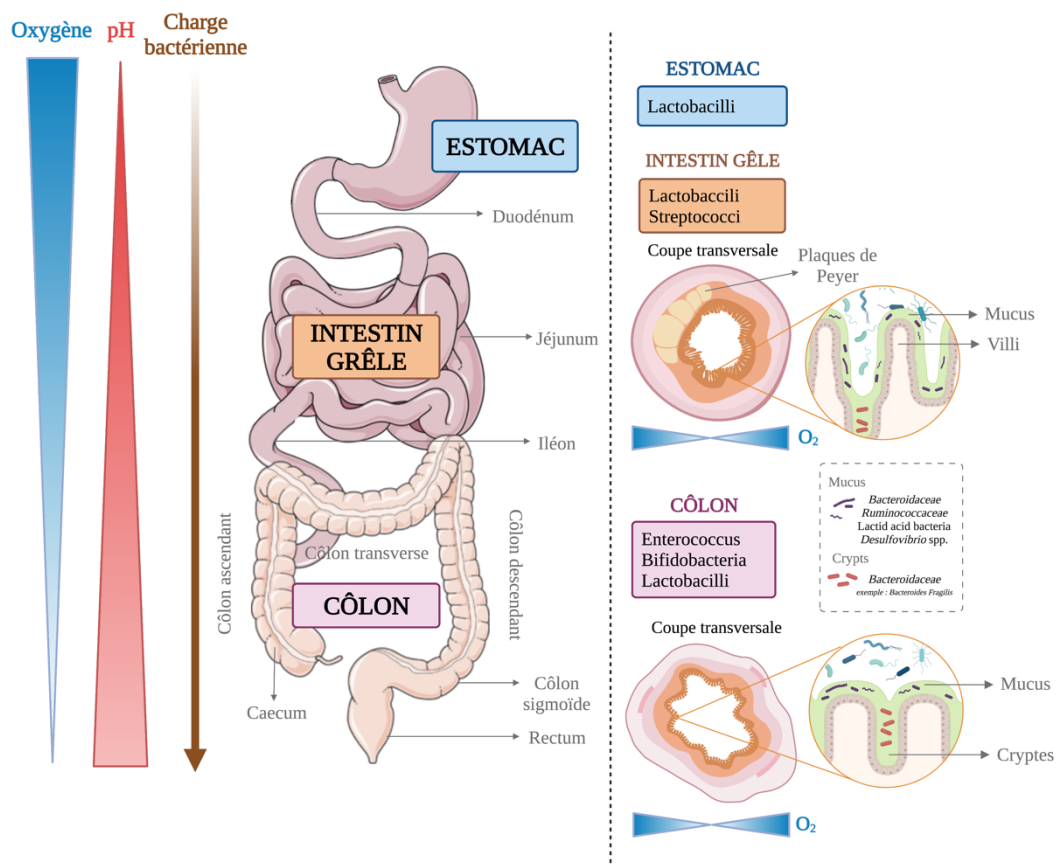


Figure 1 Répartition longitudinale et axiale du microbiote intestinal. Adapté de Pereira et al., 2017 ; Sartor, 2008 ; Donaldson et al., 2016. Figure réalisée avec biorender.com.

Le **microbiote du côlon** représente plus de 70 % de tous les microbes présents dans l'organisme avec une biomasse microbienne de 10^{11} cellules par gramme de contenu. Le côlon humain comprend le côlon ascendant, traversant, descendant et sigmoïde ainsi que le rectum [19]. Un rôle important du côlon est la résorption de l'eau et des électrolytes tandis que le microbiote contribue à la digestion des ingrédients alimentaires non digérés dans la partie supérieure de l'intestin. Les conditions physico-chimiques du côlon sont moins sévères que

celles de l'estomac ou de l'intestin grêle pour la prolifération microbienne : le pH est proche de la neutralité et la concentration en sels biliaires est plus faible. La structure de la muqueuse est également différente dans le gros intestin ce qui permet à l'hôte de tolérer une forte densité microbienne [18]. Les longs temps de transit (24-72h) dans le gros intestin peuvent avoir un impact important sur la physiologie et le métabolisme bactériens. Des études *in vivo* et *in vitro* montrent que le temps de transit dans le gros intestin entraîne une dégradation des protéines, des sucres, des acides aminés, mais surtout des fibres marquant une contribution accrue à la production d'**acides gras à chaîne courte** (AGCC) disponibles dans le côlon. Ces AGCC vont modeler la composition de l'environnement intestinal et sont une source d'énergie importante pour l'hôte [18]. Lors de cette introduction bibliographique, je ne décrirai que le **microbiote colique**, sauf indication contraire. Il est également important de noter que la majorité des études portant sur la caractérisation du microbiote intestinal reposent sur des échantillons de selles. Bien que facilement acquis, les échantillons fécaux ne représentent pas avec précision la composition du microbiote intestinal ainsi que la composition des espèces bactériennes présentes au niveau des muqueuses [27]. Cependant, il est admis que le microbiote fécal et le microbiote colique présentent des similitudes qualitatives même si des différences ont été identifiées [28].

Le microbiote colique d'un adulte sain est composé de 6 phyla principaux : **Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, Euryarcheota et Verrucomicrobia** [27]. Les phyla Firmicutes et Bacteroidetes sont toujours fortement représentés chez l'individu sain avec plus de 90% de la population bactérienne du côlon. Le phylum Bacteroidetes est composé de bactéries anaérobies à Gram négatif, les genres prédominants étant *Bacteroides* et *Prevotella*. Le genre *Bacteroides* est l'un des groupes les plus prédominants dans l'intestin représentant environ 30 % des microorganismes intestinaux, ce qui indique que ce genre bactériens a des fonctions importantes [29],[30]. Les Firmicutes sont composées de bactéries Gram-positives anaérobies obligatoires ou facultatives. Ces bactéries peuvent former des endospores, une forme résistante qui aide à survivre à des conditions défavorables dans l'intestin. Ce phylum comprend des espèces appartenant aux genres *Eubacterium*, *Clostridium*, *Ruminococcus*, *Roseburia*, *Lactobacillus* et *Butyrivibrio*.

Même si la composition des taxons au niveau du phylum est similaire entre les individus, la **diversité des espèces trouvées chez chaque individu est spécifique** [31]. Cette population spécifique individuelle est due à des facteurs individuels et environnementaux, en particulier le régime alimentaire [32]. Même s'il existe une empreinte fécale propre à chaque individu, un microbiote commun à tous les individus peut être défini. Le **core microbiote** correspond aux

espèces bactériennes partagées par la plupart des individus sains, correspondant à environ 57 espèces présentes chez 90 % des individus [33]. Plusieurs facteurs caractérisent un microbiote intestinal sain, en particulier une grande diversité et richesse taxonomique, mais aussi la présence de microorganismes connus pour améliorer le métabolisme et favoriser la résilience aux infections et à l'inflammation [1],[27],[34]. Le microbiote intestinal de l'individu sain est **majoritairement stable au cours de la vie adulte** mais cette communauté est constamment exposée à des perturbations. Les perturbations du microbiote intestinal affectent la stabilité de l'écosystème et peuvent éventuellement entraîner des états dysbiotiques associés à la maladie [35]. Différents facteurs de l'hôte ou de l'environnement peuvent impacter la composition et la diversité du microbiote intestinal. Plusieurs **facteurs de l'hôte** peuvent avoir un impact sur la composition du microbiote intestinal, tels que l'âge [36], le dysfonctionnement du système immunitaire [37] et la production de peptides antimicrobiens (AMP) [38]. Des **facteurs environnementaux** résumés dans la **Figure 2** peuvent également affecter les communautés microbiennes, tels que des changements alimentaires extrêmes [55], la prise d'antibiotiques, le mode de vie (tabagisme [56], exercice [57]), le stress physiologique et l'anxiété [58]. La plupart du temps, le microbiote intestinal altéré reviendra à son état initial après exposition à un agent de stress, c'est ce qu'on appelle la **résilience** [59]. Cependant, le microbiote intestinal peut également ne pas être en mesure de revenir à son état d'origine et ce microbiote non résilient passera à un nouvel état altéré appelé dysbiose.

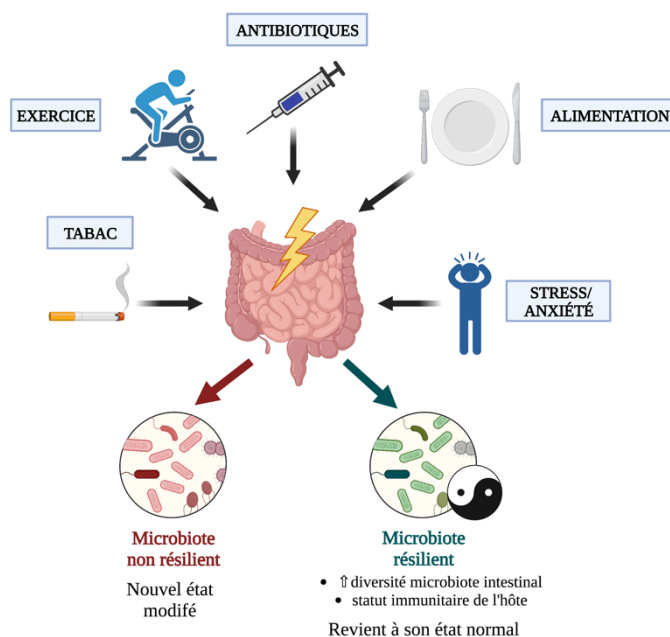


Figure 2 Différents facteurs affectant la composition du microbiote intestinal et la façon dont le microbiote résilient et non résilient réagit. Adapté de Hasans et Yang, 2019. Figure réalisée avec biorender.com.

Le microbiote intestinal joue un **rôle essentiel dans la santé** à travers des fonctions principales parmi lesquelles : la fourniture d'énergie à l'hôte, la protection contre les agents pathogènes externes, le maintien de la structure épithéliale et le développement et la maturation du système immunitaire (**Figure 3**) [72].

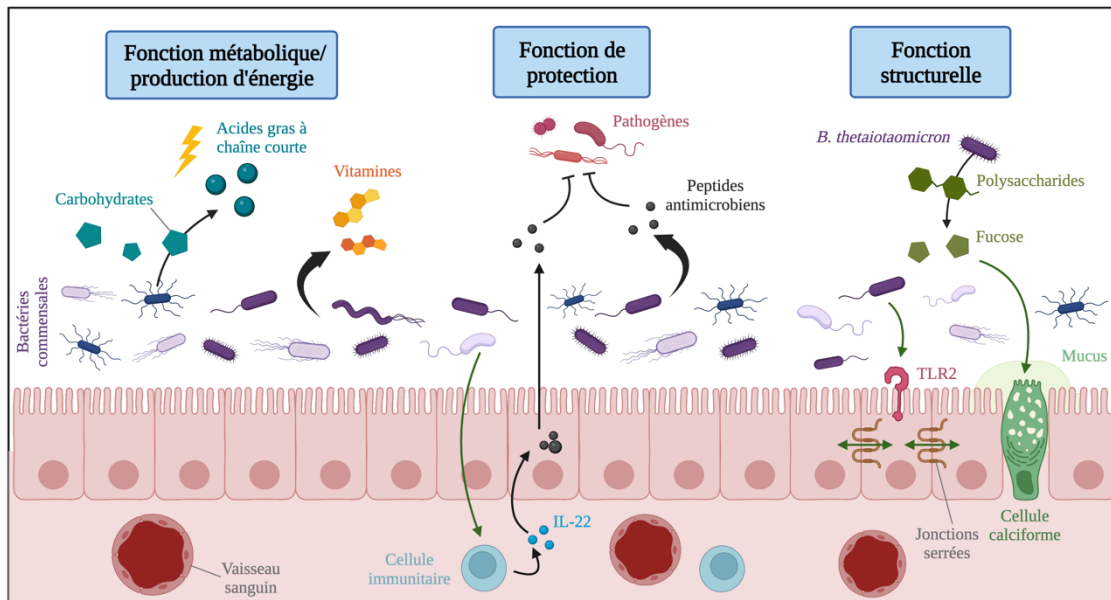


Figure 3 Représentation schématique des fonctions du microbiote intestinal. Les bactéries commensales ont des fonctions métaboliques importantes telles que la fermentation des glucides produisant des acides gras à chaîne courte et des vitamines utiles pour l'hôte. De plus, les bactéries commensales protègent l'hôte contre l'infection par des agents pathogènes en produisant des peptides antimicrobiens ou en induisant le système immunitaire. Enfin, les bactéries commensales jouent un rôle sur la fortification de la barrière par l'induction de la production de mucus par les cellules caliciformes et la stimulation de TLR2 conduisant à la fortification des jonctions serrées. Adapté de O'Hara et Shanahan, 2006. Figure réalisée avec biorender.com.

CHAPITRE 2 : Microbiote intestinal et pathologies

La dysbiose peut être définie par les caractéristiques suivantes : une diminution de la diversité et/ou de la richesse microbienne, une altération du rapport Bacteroidetes / Firmicutes, une diminution des espèces bénéfiques (i.e. *Bifidobacterium*, *Faecalibacterium prauznitsii*...) et/ou une augmentation des bactéries délétères telles que les Proteobacteria. En effet, les **Proteobacteria** représentent habituellement 1 à 5 % du microbiote intestinal sain mais ce phylum peut atteindre jusqu'à 35 % dans l'intestin des patients souffrant de diverses pathologies, notamment inflammatoires. Dans le chapitre 2, la relation entre le microbiote intestinal et deux exemples de maladies inflammatoires sont décrites : la maladie de Crohn (MC) en tant que maladie inflammatoire intestinale et la polyarthrite rhumatoïde (PR) en tant que maladie inflammatoire extra-intestinale.

Maladie de Crohn Les maladies inflammatoires chroniques peuvent être divisées en deux sous-types : la MC et la rectocolite hémorragique. Les symptômes comprennent la diarrhée, les douleurs abdominales, les saignements rectaux, la fièvre, la perte de poids et la fatigue [113]. Les maladies inflammatoires chroniques touchent environ 3,1 millions de personnes aux États-Unis [114]. L'incidence de cette maladie a augmenté dans le monde occidental et continue d'augmenter rapidement au cours des 50 dernières années [115], [116]. La MC est une maladie chronique, segmentaire et inflammatoire qui peut affecter toutes les parties du tube digestif et provoquer des lésions de la bouche à l'anus accompagnées dans certains cas de complications extra-intestinales. L'étiologie exacte est encore mal connue et résulte probablement de la combinaison de multiples facteurs tels que des troubles immunologiques, un microbiote intestinal altéré, des facteurs génétiques et des facteurs environnementaux. La dysbiose chez les patients atteints de MC est généralement caractérisée par une augmentation des Proteobacteria tandis que d'autres populations commensales sont diminuées, notamment les membres des phylums Firmicutes et Bacteroidetes [139],[140]. En effet, le phylum des Proteobacteria peut augmenter jusqu'à 10 % alors qu'il ne représente que ~1 % chez les individus sains (**Figure 4**) [141].

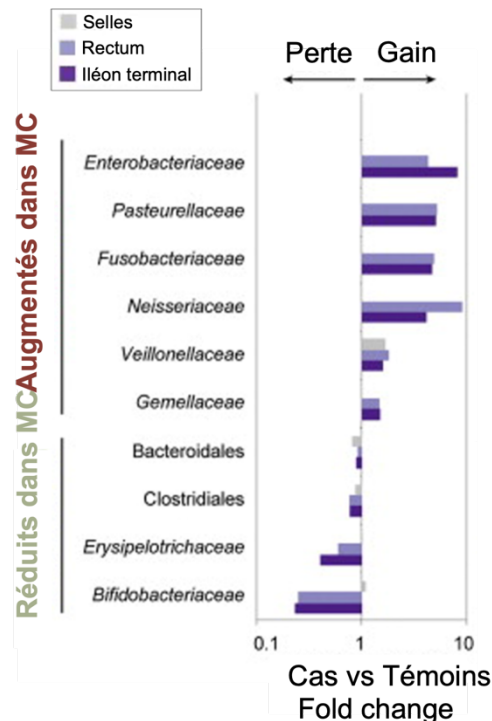


Figure 4 Différences d'abondance microbienne des familles bactériennes chez les patients atteints de MC. Le fold change pour chaque taxon a été calculé en divisant l'abondance moyenne chez les cas patients par celle des témoins. Plusieurs biomarqueurs taxonomiques ont été mesurés à différents sites. Les sites iléal et rectal se sont avérés significativement corrélés avec le phénotype de la maladie. Cependant, cette signature microbienne a été perdue dans les échantillons de selles. Les taxons du haut sont augmentés en état de maladie, tandis que les taxons du bas suivent une tendance opposée. *Adapté de Gevers et al., 2014.*

Chez les patients atteints de MC, plusieurs études ont prouvé une augmentation des protéobactéries, principalement des *E. coli* associées à la muqueuse intestinale [149]. La plupart de ces souches d'*E. coli* peuvent adhérer et envahir les cellules épithéliales intestinales et ont été dénommées **AIEC** pour *E. coli* Adhérentes et Invasives [150]. La prévalence des AIEC dans la muqueuse intestinale des patients atteints de MC iléale est de 38 % contre seulement 6 % chez les individus sains [151]. Il a été montré que ces souches d'*E. coli* étaient anormalement prévalentes dans les lésions iléales précoces et chroniques indiquant qu'elles pourraient avoir un rôle potentiel sur la pathologie précoce de la MC [152]. Les critères qui caractérisent les AIEC sont : la capacité à adhérer et à envahir les cellules épithéliales intestinales différenciées et la capacité à survivre au sein des macrophages. La plupart des souches d'AIEC n'héberge aucun des facteurs de virulence connus exprimés par d'autres pathotypes d'*E. coli* tels que l'*E. coli* entéropathogènes (EPEC) ou entérohémorragiques (EHEC) [152]. L'absence de gènes de pathogénicité conventionnels suggère que les AIEC devraient être davantage considérées comme un **pathobionte** plutôt que comme un pathogène *sensu stricto*.

Polyarthrite Rhumatoïde La PR est une maladie inflammatoire auto-immune chronique caractérisée par une réponse immunitaire constante située au niveau des articulations entraînant une inflammation du tissu synovial et une érosion subséquente du cartilage et de l'os [168]. La PR est liée à la production d'auto-antigènes (c'est-à-dire d'anticorps anti-facteur rhumatoïde ou anti-protéine citrullinée) dans les articulations, entraînant l'incapacité de reconnaître les antigènes de l'hôte. Ce processus conduit à une surproduction de cytokines par le système immunitaire de l'hôte, ce qui entraîne une inflammation persistante et une destruction des articulations et des os. Les symptômes peuvent inclure une raideur matinale des articulations touchées, de la fatigue, de la fièvre, une perte de poids, des articulations sensibles et enflées formant des nodules rhumatoïdes sous la peau [169]. La PR touche 0,5 à 1,1 % de la population mondiale avec une prévalence plus élevée dans les pays développés comme l'Europe du Nord et l'Amérique du Nord [170]. Par conséquent, l'incidence et la prévalence de la PR varient selon les populations et dépendent fortement du choix des critères pour les études de cohorte. La PR est une maladie multifactorielle qui présente à la fois des composantes auto-immunes et inflammatoires et est soumise à diverses influences génétiques, environnementales et épigénétiques résumées dans la **Figure 5** [170],[171].

Etiologie of Rheumatoid Arthritis

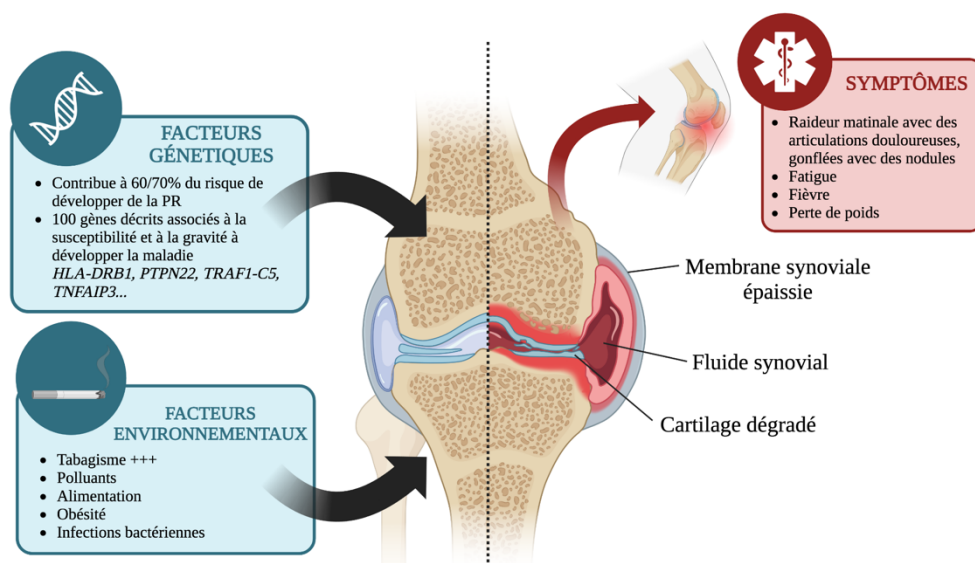


Figure 5 Différents facteurs participent à l'étiologie de la PR. Parmi les facteurs de l'hôte, les facteurs génétiques jouent un rôle important contribuant à 60 à 70 % du risque de développer une PR. L'un des gènes les plus étudiés associés au risque élevé et à la gravité de cette maladie est HLA-DRB1 responsable du déclenchement des cellules T du système immunitaire. Plusieurs facteurs environnementaux participent aux premiers stades de la maladie, l'un des plus étudiés étant la fumée de cigarette. La PR est une maladie auto-immune caractérisée par des symptômes tels que raideur matinale, articulations douloureuses, enflées et nodulaires, fatigue, fièvre et perte de poids. *Figure réalisée avec biorender.com.*

Malgré le fait que la PR soit une maladie extra-intestinale, plusieurs études ont montré que les patients présentent des différences significatives dans la composition de leur microbiote intestinal par rapport aux individus sains. Ces altérations du microbiote intestinal sont liées au risque et à la sévérité de la maladie. Des études globales ont rapporté que le microbiote intestinal des patients atteints de PR se caractérise par une diminution des *Bifidobacteria* [213],[214],[215], une augmentation des *Lactobacillus* [216] une prolifération de Proteobacteria et une forte **abondance de *Prevotella copri*** [217]. L'un des marqueurs de dysbiose intestinale associée à la PR le plus étudié est la forte abondance de *Prevotella copri* dans les échantillons fécaux des patients [217]. *P. copri* est une bactérie Gram-négative anaérobie obligatoire et non sporulée qui a été isolée pour la première fois à partir d'échantillons fécaux humains au Japon [221]. Très récemment, il a été démontré que des patients atteints de PR en phase préclinique de pays européens hébergeaient une grande abondance d'espèces de *Prevotella*, y compris *P. copri*, dans l'intestin, ce qui suggère que la dysbiose précède le développement de l'arthrite [222]. A l'inverse, l'abondance relative de *P. copri* dans l'intestin était réduite chez les patients atteints de PR chronique suggérant que cette bactérie pourrait jouer un rôle dans les premiers stades de la maladie. Cependant il apparaît que les groupes bactériens détectés sont variables et dépendent souvent de la cohorte de patients utilisée dans

l'étude. Différents paramètres tels que l'état de la maladie, l'âge et le traitement influencent la composition du microbiote intestinal et doivent être pris en compte lorsque l'on tente de définir les caractéristiques de la dysbiose.

CHAPITRE 3 : Les Proteobacteria du microbiote intestinal

Le chapitre 3 est consacré au phylum Proteobacteria. Comme mentionné précédemment, les membres du phylum Proteobacteria sont systématiquement présents dans l'intestin d'individus sains mais en faible abondance [228]. Cependant, des preuves irréfutables montrent qu'un microbiote intestinal déséquilibré décrit dans le contexte d'une pathologie est souvent associé à une **augmentation de l'abondance du phylum Proteobacteria (Figure 6)** [229], [230], [231], [232]. Différentes études ont lié une expansion des Entérobactéries, la principale famille représentative des protéobactéries, à de nombreuses pathologies inflammatoires telles que la MC et la PR décrites au chapitre 2 [233]. Parce qu'*E. coli* est à la fois le membre le plus abondant des protéobactéries dans l'intestin et un organisme modèle de référence pour étudier et comprendre la biologie, la majorité des études liées à ce phylum dans le microbiote intestinal est consacrée à *E. coli*. Les souches commensales d'*E. coli* se trouvent principalement dans le **gros intestin**, en particulier dans le caecum et le côlon. La principale niche écologique d'*E. coli* dans l'intestin est la **couche de mucus** recouvrant les cellules épithéliales du côlon. Dans le mucus, *E. coli* se développe sous forme de **biofilm** complexe multi-espèce dans lequel la bactérie entre en compétition pour les nutriments provenant de l'alimentation ou du mucus lui-même [243],[244]. Les souches isolées de cette partie de l'intestin se développent grâce aux nutriments acquis à partir d'au moins sept sucres dérivés du mucus. Bien que les concentrations de ces sucres dans l'intestin soient faibles, *E. coli* maximise sa croissance en utilisant la **respiration micro-aérobie et anaérobie** [245], et maintient sa densité de population entre **10⁶ et 10⁹ cellules par gramme de matière fécale** [241]. Même si peu d'informations précisant les variations taxonomiques au sein du phylum des Proteobacteria dans les différentes études, la famille des *Enterobacteriaceae* semble avoir une place prépondérante pour expliquer l'expansion des Proteobacteria associée à une pathologie. À partir de ces observations, plusieurs groupes ont proposé qu'un enrichissement en protéobactéries dans l'intestin constitue un marqueur d'un microbiote intestinal instable et un critère potentiel de diagnostic de la maladie [232], [274], [231].

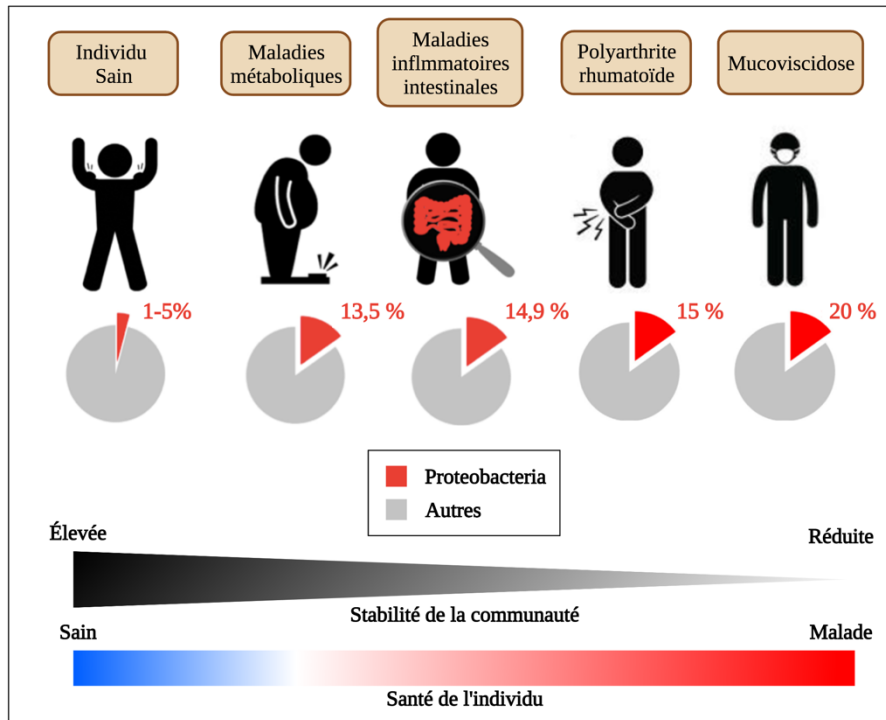


Figure 6 Expansion des Proteobacteria dans le microbiote intestinal sous différentes conditions. Adapté de Shin et al., 2015. Figure réalisée avec biorender.com.

Plusieurs mécanismes ont été proposés dans la littérature pour expliquer les proliférations d'*Enterobacteriaceae* dans l'intestin dysbiotique (**Figure 7**). Une de ces théories repose sur l'**adaptation métabolique des *Enterobacteriaceae***. La dysbiose provoque souvent des modifications importantes de l'environnement intestinal et peut influencer la diversité et les concentrations de nutriments disponibles dans la lumière que les différents membres du microbiote intestinal peuvent utiliser avec une efficacité variable. Les *Enterobacteriaceae* et particulièrement *E. coli* présentent une structure génomique dynamique conduisant à un haut niveau de plasticité génomique. De plus, les *Enterobacteriaceae* se sont révélées être la principale source de gènes variables dans le microbiome intestinal entre les individus sains alors qu'elles ne représentent qu'une population mineure, ce qui indique que cette famille apporte une variabilité dans la fonction des gènes microbiens intestinaux [288]. Conformément à ces observations, les modifications de l'environnement intestinal causées par la dysbiose associée à une pathologie peuvent entraîner une expansion des Proteobacteria en raison d'une **meilleure adaptabilité à cataboliser des substrats potentiels**. Différentes études ont exploré l'adaptation métabolique microbienne et certaines sources de nutriments disponibles pour les *Enterobacteriaceae* lors de la dysbiose intestinale.

Bien que certaines de ces recherches aient été menées avec des souches pathogènes, il est raisonnable de penser que la plupart des traits peuvent être transposés aux *Enterobacteriaceae* commensales. Ainsi, une meilleure compréhension des mécanismes d'adaptation des *Enterobacteriaceae* dans l'intestin dysbiotique pourrait aider à concevoir des stratégies d'intervention potentielles pour inverser la dysbiose et améliorer la santé.

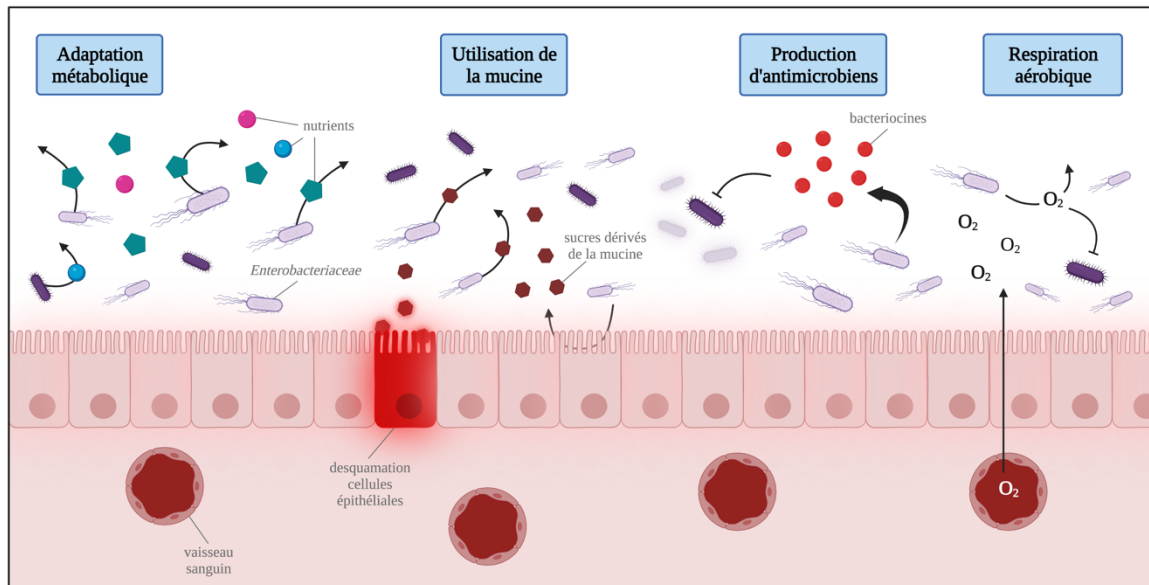


Figure 7 Mécanismes d'expansion des *Enterobacteriaceae* dans l'intestin dysbiotique. Différentes théories existent dans la littérature qui pourraient expliquer l'expansion d'*Enterobacteriaceae* dans l'intestin dysbiotique : une adaptation métabolique des *Enterobacteriaceae* à utiliser certains nutriments comme avantage nutritif, l'utilisation des sucres du mucus, la production d'antimicrobiens et l'augmentation des accepteurs d'électrons pour la respiration, *Figure réalisée avec biorender.com.*

RÉSULTATS

PARTIE 1 : Identification de propriétés métaboliques potentiellement impliquées dans l'expansion des *Enterobacteriaceae* dans l'intestin des patients atteints de MC ou de PR

Comme décrit lors de l'introduction bibliographique, une expansion d'*Enterobacteriaceae* a été démontrée comme étant signature de la dysbiose dans plusieurs maladies inflammatoires telle que la MC et la PR. Parmi les mécanismes proposés dans la littérature pour expliquer l'expansion des protéobactéries au cours de la dysbiose, l'adaptation nutritionnelle des bactéries aux changements de l'environnement intestinal peut conduire à l'expansion des protéobactéries en raison d'une meilleure adaptabilité à métaboliser les substrats présents. Le premier objectif de mon projet de thèse a été de comparer les capacités métaboliques d'entérobactéries provenant d'individus sains ou de patients présentant une pathologie associée à une expansion d'*Enterobacteriaceae*. Pour cela, nous avons constitué une collection de souches d'*Enterobacteriaceae* à partir de prélèvements fécaux d'individus SA, de patients PR ou de patients MC étalés sur milieux sélectifs. Après identification taxonomique, la plupart des *Enterobacteriaceae* obtenus appartenaient au genre *Escherichia* (>75%). Parce que le but de notre étude est de comparer les capacités métaboliques entre collections de souches et que ces propriétés peuvent varier largement selon les genres bactériens, nous avons décidé de ne conserver que les souches d'*Escherichia*. La collection de souches ainsi formée (~15 souches par groupe SA, MC ou PR) a été criblée par l'approche OmniLog permettant la caractérisation de phénotypes microbiens. Les plaques PM01 et PM02 ont été utilisées afin d'évaluer la capacité de chaque souche à utiliser 190 composés différents en tant que source de carbone. Le système OmniLog est composé d'un incubateur à plaques équipé d'une caméra quantifiant la respiration cellulaire (production de NADH) en tant que rapporteur universel. Si la souche inoculée est capable d'utiliser le substrat présent dans le puits, la cellule respire activement, réduisant ainsi un composé, le tétrazolium, pour former un produit coloré quantifié par le système OmniLog. Suite à ce criblage, la D-sérine a été identifiée comme étant préférentiellement métabolisée par les souches MC, alors que le saccharose est un substrat qui à l'inverse est préférentiellement consommé par les souches SA. Des travaux supplémentaires ont été entrepris pour caractériser et explorer le rôle potentiel de la D-sérine et du saccharose dans l'aptitude différentielle des souches du groupe SA et MC à proliférer au sein de l'intestin. Ce travail fait l'objet d'une publication scientifique en cours d'écriture dont la première page est présentée ci-après.

Identification des propriétés métaboliques potentiellement impliquées dans l'expansion des *Enterobacteriaceae* dans l'intestin de patients souffrant de maladies inflammatoires

Maria Ines Moreira de Gouveia¹, Annie Garrivier¹, Julien Daniel¹, Christophe Del'homme¹, Nicolas Barnich², Annick Bernalier-Donadille¹, Gregory Jubelin^{1*}

¹ Université Clermont Auvergne, INRAE, MEDIS UMR454, F-63000 Clermont-Ferrand, France

² Université Clermont Auvergne, Inserm U1071, USC-INRAE 2018, M2iSH, Clermont-Ferrand, France.

* auteur correspondant : Grégory Jubelin, INRAE Auvergne Rhône-Alpes, UMR454 MEDIS, site de Theix, 63122 Saint-Genès-Champanelle, France. Courriel : gregory.jubelin@inrae.fr

RÉSUMÉ

La dysbiose intestinale associée à différentes maladies intestinales ou extra-intestinales est souvent caractérisée par une expansion des *Enterobacteriaceae* qui ne représentent habituellement que 1 % du microbiote intestinal chez l'homme sain. Cette prolifération s'explique en partie par l'émergence et la prolifération de souches adaptées à l'environnement intestinal dysbiotique. Dans cette étude, nous avons réalisé un criblage pour identifier des substrats différenciellement catabolisés par des souches commensales d'*Enterobacteriaceae* isolées d'individus sains (SA) ou de patients atteints de la maladie de Crohn (MC) ou de polyarthrite rhumatoïde (PR), deux pathologies associées à une dysbiose intestinale. Parmi les différents substrats identifiés, nous avons observé une capacité opposée des souches SA et MC à métaboliser la D-sérine et le saccharose, une caractéristique résultant de réarrangements génétiques dans le locus *argW* du chromosome d'*Escherichia coli*. Nous avons démontré que le phénotype D-sérine⁺ des souches MC dépend de la présence des gènes *dscXA* et qu'il confère un avantage nutritionnel par rapport aux souches SA. Notre étude fournit de nouveaux éléments sur les substrats métaboliques potentiels qui pourraient favoriser l'expansion des *Enterobacteriaceae* en conditions dysbiotiques.

MOTS CLÉS

dysbiose intestinale, maladie de Crohn, polyarthrite rhumatoïde, *Enterobacteriaceae*, compétition nutritionnelle, D-sérine

PARTIE 2 : Le potentiel inexploré de l'utilisation de l'éthanolamine par les souches commensales d'*E. coli*

Dans la première partie de mon travail, j'ai évalué l'utilisation de 190 substrats différents comme seule source de carbone pour mettre en évidence des molécules potentielles qui pourraient être utilisées par les *Enterobacteriaceae* comme avantage compétitif dans l'intestin dysbiotique. Ces 190 substrats, correspondant principalement à des glucides et acides aminés, ne sont qu'une petite fraction des molécules présentes dans le TGI. D'après la littérature et les travaux de notre laboratoire, un composé qui n'était pas représenté dans les plaques OmniLog mais qui a un intérêt majeur dans un environnement intestinal est l'éthanolamine (EA). Cette petite molécule ($\text{H}_2\text{N}-\text{CH}_2-\text{CH}_2-\text{OH}$), composée d'une amine primaire et d'un alcool primaire, est un constituant de la phosphatidyléthanolamine, composant majeur de toute membrane cellulaire [296]. Grâce au renouvellement des cellules bactériennes et épithéliales, l'EA est naturellement libéré dans le TGI.

De façon intéressante, de nombreuses études ont déterminé que l'EA est un substrat préférentiel pour plusieurs agents pathogènes intestinaux tels que les *E. coli* pathogènes, *Listeria monocytogenes* ou *Salmonella enterica* sérotype Typhimurium [292]. Ces agents pathogènes profitent de l'avantage nutritionnel de l'EA dans l'intestin pour proliférer et/ou en tant que molécule signal pour contrôler l'expression de gènes de virulence [297]. Bien que l'EA ait été étudiée pour son rôle clé lors des interactions pathogènes hôte-intestin, peu d'études suggèrent que la consommation d'EA est spécifiquement associée à la virulence. Par exemple, une étude récente a démontré que deux souches d'*E. coli* commensales humaines, HS et Nissle, sont capables de consommer de l'EA comme source d'azote lors de la croissance *in vitro*, et même plus efficacement qu'une souche d'*E. coli* entérohémorragique (EHEC) O157:H7 [298].

Dans cette deuxième partie de mon travail de thèse, j'ai décidé de me concentrer sur le catabolisme de l'EA. Dans un premier temps, j'ai évalué la capacité des 3 groupes d'isolats d'*E. coli* (groupes SA, MC et PR) utilisés dans la partie 1 à consommer l'EA. Les tests de croissance ont été effectués dans un milieu minimal supplémenté avec l'EA comme seule source d'azote ou de carbone (**Figure 8**). Nous avons observé que la plupart des souches consomment l'EA comme source d'azote quels que soient les groupes de souches, ce qui indique que le catabolisme de l'EA est en effet répandu chez les *E. coli* commensales. De plus, deux souches de la collection SA (L3 et AA1) ont également pu utiliser l'EA comme seule source de carbone.

Ce phénotype était absent de notre collection de souches MC et PR. Certaines souches des groupes SA et PR n'étaient pas capables d'utiliser l'EA comme source d'énergie, démontrant la variabilité de l'utilisation de l'EA parmi les *E. coli* commensales.

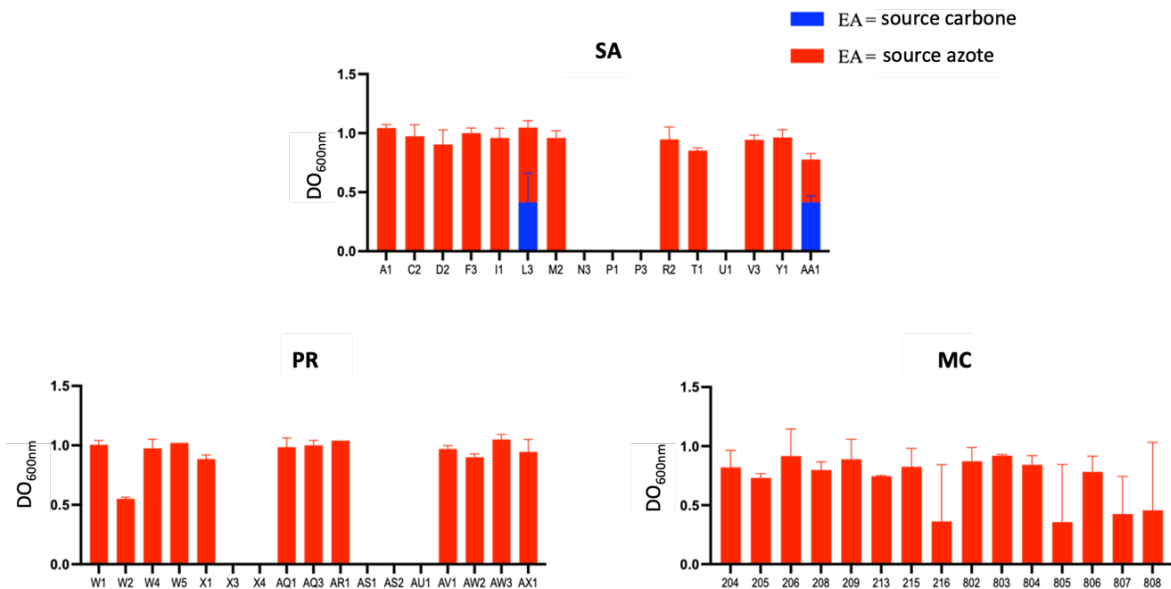


Figure 8 Profils d'utilisation de l'EA par des souches commensales d'*E. coli* de la collection SA, MC et PR. La croissance *in vitro* a été évaluée dans un milieu minimal M9 supplémenté avec l'EA comme seule source d'azote ou de carbone. Les résultats sont présentés sous forme de moyennes et d'écart types de deux à quatre expériences indépendantes. La DO_{600nm} a été mesurée après 24 h d'incubation pour les cultures en M9 supplémenté en glucose et EA (rouge) et après 72 h pour les cultures en M9 supplémenté en NH₄Cl et EA (bleu).

En résumé, ces résultats préliminaires m'ont amené à étudier la variabilité du métabolisme des EA chez les *E. coli* commensaux. Étant donné que la variabilité la plus élevée a été observée dans le groupe SA et que les souches MC et PR ont été isolées chez des patients présentant un biais possible de sélection/adaptation de souches dans l'environnement intestinal dysbiotique, nous avons choisi de poursuivre les recherches uniquement avec les souches SA. Dans ce but, la collection de souches SA a été étendue à 40 isolats au total. Nous avons d'abord étudié la diversité phénotypique de l'utilisation de l'EA parmi notre collection de souches SA. Ensuite, l'objectif principal était d'expliquer la diversité phénotypique observée dans la collection afin d'identifier la raison qui différencie les profils d'utilisation de l'EA. Ce travail a donné lieu à une publication (manuscrit N°2), qui a été acceptée dans la revue à comité de lecture *Research in Microbiology* en août 2022.

En parallèle de ce travail, nous avons également développé une étude méthodologique pour concevoir un nouveau système rapporteur pour quantifier l'expression de gènes bactériens dans des conditions complexes telles que l'écosystème intestinal. Tirant parti des processus de régulation bien connus impliqués dans la transcription des gènes associés au catabolisme de l'EA chez *E. coli*, nous avons conçu un nouvel outil assurant à la fois l'identification d'une souche d'intérêt dans des environnements complexes et le suivi de l'expression des gènes par l'utilisation combinée de deux gènes rapporteurs codant des protéines fluorescentes distinctes. L'outil a été évalué pour sa capacité à rapporter la détection de la concentration d'EA par *E. coli* *in vitro* mais aussi *in vivo* dans l'intestin de souris. Ce travail a donné lieu à une publication (manuscrit N°3) qui a été soumise à la revue *Applied Microbiology and Biotechnology*.

Avant de présenter les manuscrits, j'introduis ici quelques informations importantes concernant la voie catabolique de l'EA et les gènes associés. La capacité à cataboliser l'EA est codée par les gènes d'utilisation de l'éthanolamine (*eut*) organisés sous forme d'opéron. Par une étude de génomique comparative, Tsoy et al (2009) ont mis en évidence différentes organisations du cluster des gènes *eut* chez les bactéries. En particulier, les membres de la famille des Proteobacteria *Enterobacteriaceae*, comprenant *S. typhimurium* et *E. coli*, ont de longs opérons *eut*, tout comme les membres du phylum Firmicutes. Chez *E. coli*, cet opéron est composé de 17 gènes *eut* codant non seulement pour des enzymes clés du catabolisme de l'EA mais également pour des protéines qui forment les microcompartiments dans lesquels s'effectue le catabolisme de l'EA (**Figure 9**) [300].

Le locus *eut* permet la synthèse du facteur de transcription EutR qui détecte l'EA ainsi que la vitamine B₁₂ (co-facteur) afin d'activer directement la transcription de l'opéron *eut* [301],[302]. Les gènes centraux dans la transformation de l'EA sont *eutB* et *eutC*, dont les produits protéiques forment l'éthanolamine ammoniac lyase. L'éthanolamine ammoniac lyase EutBC transforme l'EA en acétaldéhyde et en ammoniac, une précieuse source d'énergie azotée. L'acétaldéhyde est ensuite converti en éthanol par EutG ou en acétyl-CoA par EutE, qui peuvent être utilisés dans de nombreux processus cellulaires (cycle de Krebs, biosynthèse des lipides, ou autres processus). L'acétyl-CoA peut également être converti en acétate (**Figure 9**) [304].

Voie catabolique de l'éthanolamine

A. Constitution de l'opéron *eut* (chez *E. coli*)



B. Voie catabolique

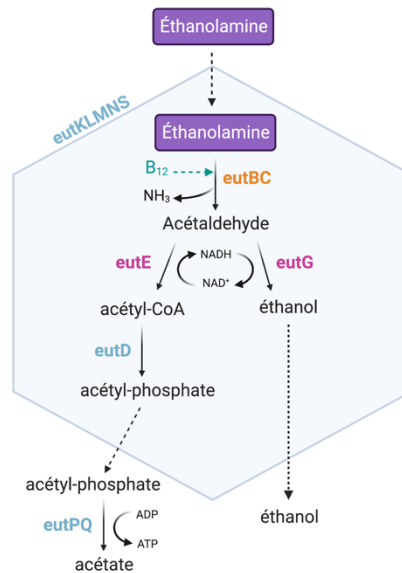


Figure 9 Catabolisme de l'EA et organisation de l'opéron *eut* chez *E. coli*. Représentation schématique de (A) l'organisation de l'opéron *eut* chez *E. coli* et (B) du catabolisme de l'éthanolamine dans la structure du microcompartiment. *Figure réalisée avec biorender.com.*

L'opéron *eut* code également pour des protéines structurales (EutS, EutM, EutK, EutL et EutN) qui s'auto-assemblent dans des microcompartiments bactériens formant une structure protéique icosaédrique sélectivement perméable [305]. Toutes les protéines impliquées dans ces microcompartiments, à l'exception d'EutK, forment des unités hexagonales comme les protéines précédemment décrites du carboxysome [306]. Il est important de noter que l'opéron *eut* ne code pas systématiquement pour ces protéines structurales de microcompartiments, mais pour les souches qui les expriment, ces structures apportent différents avantages au catabolisme de l'EA. En particulier, les microcompartiments assurent la concentration des produits intermédiaires du catabolisme de l'EA et favorise ainsi l'efficacité des réactions enzymatiques [307],[308]. De plus, l'acétaldéhyde, un intermédiaire volatil et toxique, est retenu dans les microcompartiments, limitant sa toxicité pour la bactérie.

Diversité de l'utilisation de l'éthanolamine par des *Escherichia coli* commensales de l'intestin humain

Maria Ines Moreira de Gouveia^a, Julien Daniel^a, Annie Garrivier^a, Annick Bernalier-Donadille^a, Gregory Jubelin^{a*}

^a Université Clermont Auvergne, INRAE, MEDIS UMR454, F-63000 Clermont-Ferrand, France

* auteur correspondant

RÉSUMÉ

L'éthanolamine (EA) est un substrat naturellement présent dans l'intestin humain et son catabolisme par les bactéries repose sur la présence de gènes *eut* codant des enzymes métaboliques spécifiques et des protéines accessoires. À ce jour, l'utilisation de l'EA a été principalement étudiée chez les pathogènes bactériens intestinaux. Le but de cette étude était d'évaluer la capacité des isolats d'*Escherichia coli* commensaux de l'intestin humain à utiliser l'EA comme source d'azote et/ou de carbone. Bien que la capacité à consommer l'EA soit hétérogène entre les 40 souches de notre collection, nous avons déterminé que la plupart d'entre elles pouvaient dégrader l'EA pour générer de l'ammoniac, une ressource azotée utile pour la croissance. Trois isolats ont également pu exploiter l'EA en tant que source de carbone. Nous avons également mis en évidence que l'incapacité de certaines souches à cataboliser l'EA s'explique soit par des mutations de l'opéron *eut*, soit par un défaut de transcription des gènes. Enfin, nous avons démontré l'importance de l'utilisation de l'EA pour le fitness optimal des *E. coli* commensaux *in vivo*. Notre étude fournit de nouvelles informations sur la diversité des *E. coli* commensales pour utiliser l'EA comme nutriment dans l'intestin et ouvre des portes à de nouvelles recherches dans le domaine des interactions entre l'hôte, le microbiote intestinal et les agents pathogènes.

MOTS CLÉS

Éthanolamine; opéron *eut*; *E. coli* commensales; intestin humain ; compétition nutritionnelle

Conception et validation d'un système rapporteur à double fluorescence pour suivre l'expression de gènes bactériens *in situ*

Maria Ines Moreira de Gouveia¹, Audrey Reuter¹, Annie Garrivier¹, Julien Daniel¹, Annick Bernalier-Donadille¹, Gregory Jubelin^{1*}

¹ Université Clermont Auvergne, INRAE, MEDIS UMR454, F-63000 Clermont-Ferrand, France

* auteur correspondant : Grégory Jubelin, INRAE Auvergne Rhône-Alpes, UMR454 MEDIS, site de Theix, 63122 Saint-Genès-Champanelle, France. Courriel : gregory.jubelin@inrae.fr

RÉSUMÉ

Les systèmes rapporteurs basés sur la fluorescence sont des outils précieux pour étudier la dynamique de l'expression génique dans les cellules vivantes. Cependant, les stratégies disponibles pour suivre l'expression de gènes chez les bactéries au sein de leur écosystème naturel, qui peuvent être généralement riches et complexes, sont rares. Dans ce travail, nous avons conçu et développé un nouvel outil assurant à la fois l'identification d'une souche d'intérêt dans des environnements complexes et le suivi de l'expression de gènes grâce à l'utilisation combinée de deux gènes rapporteurs codant des protéines fluorescentes distinctes. L'outil a été validé chez *Escherichia coli* pour suivre l'expression des gènes *eut* impliqués dans le catabolisme de l'éthanolamine. Nous avons démontré que la souche rapportrice construite répond de façon corrélée à l'augmentation de la concentration en éthanolamine lors de cultures *in vitro*. La souche rapportrice a ensuite été inoculée chez des souris et une stratégie de cytométrie en flux a été développée pour détecter spécifiquement la souche rapportrice parmi le microbiote dense des échantillons intestinaux prélevés, permettant l'analyse du niveau d'expression des gènes *eut in vivo*. Ce nouveau système rapporteur à double fluorescence facilitera l'évaluation des processus de transcription chez les bactéries au sein de leur environnement complexe.

DISCUSSION

Au cours des années, l'accumulation des connaissances sur la composition et la fonction du microbiote intestinal a mis en évidence son importance et son impact sur la santé d'un individu. En effet, plusieurs études dans la littérature ont établi un lien entre la dysbiose intestinale et les pathologies intestinales et extra-intestinales telles que la MC et la PR. Pour les patients atteints de MC, plusieurs études ont montré une augmentation des Proteobacteria, principalement des *E. coli* associées à la muqueuse intestinale [149],[143],[139],[311]. Pour les patients atteints de PR, l'expansion des *Enterobacteriaceae* semble être plus variable mais a été détectée dans certaines études ainsi que dans notre collection d'isolats d'*E. coli* [230]. Une expansion d'entérobactéries semble être une signature de dysbiose pour ces deux maladies inflammatoires.

Plusieurs théories (décrites au chapitre 3 de mon introduction bibliographique) ont été proposées pour expliquer cette expansion. L'une de ces théories est basé sur l'adaptation métabolique des entérobactéries à l'environnement intestinal dysbiotique. Les conditions changeantes du tractus gastro-intestinal en cas de maladie pourraient influencer les composés disponibles dans la lumière intestinale qui peuvent être utilisés par les bactéries commensales. L'objectif principal de mon travail de thèse était d'identifier des substrats qui pourraient conférer aux *Enterobacteriaceae* un avantage nutritionnel pour favoriser leur expansion en cas de dysbiose associée à la MC et à la PR.

PARTIE 1 : Identification de propriétés métaboliques potentiellement impliquées dans l'expansion des *Enterobacteriaceae* dans l'intestin des patients atteints de MC ou de PR

Le premier objectif de mon projet de thèse a été d'identifier les substrats consommés de manière différentielle par les *Enterobacteriaceae* isolées de l'intestin d'individus sains ou de l'intestin de patients atteints de MC ou de PR. Dans cette optique, un criblage de notre collection de souches d'*Enterobacteriaceae* SA, MC et PR a été réalisé avec le système OmniLog pour comparer l'utilisation de 190 substrats différents.

L'analyse OmniLog a montré que les souches PR avaient une meilleure capacité à métaboliser l'acide α -cétobutyrique et l'acide D-malique que les souches SA. Même si aucun lien n'existe entre l'acide D-malique et la PR dans la littérature, peu d'études relient la PR à l'acide α -cétobutyrique. Ce cétoacide appelé acide α -cétobutyrique ou 2-oxobutanoate, est un produit de dégradation de plusieurs acides aminés tels que la méthionine et la thréonine. Ce composé peut être transformé en propanoyl-CoA qui peut en outre participer au cycle de Krebs.

Très récemment, Jutley et al. (2021) ont analysé le métabolome sérique et urinaire de patients atteints de PR [312]. Cette étude a montré une corrélation positive entre la protéine C-réactive et la présence d'acide kétobutyrique. La protéine C-réactive est une protéine dont les concentrations augmentent en réponse à l'inflammation, suggérant un lien entre l'inflammation systémique et l'augmentation de la concentration de ce cétoacide dans le plasma de patients atteints de PR précoce. De plus, Wang et al. ont analysé le profil métabolique du plasma de rats qui ont développé la goutte, une forme d'arthrite très douloureuse. Le profil métabolique plasmatique a révélé 22 métabolites endogènes associés à la goutte, dont l'acide α -cétobutyrique [313].

Même si la PR est une maladie inflammatoire extra-intestinale, les altérations du microbiote intestinal sont bien décrites dans la littérature. Cependant, les mécanismes impliqués dans l'axe intestin-articulations restent inconnus. Il existe des interactions complexes entre les réponses immunitaires innées et acquises, impliquant les cellules présentatrices d'antigène, la formation de lymphocytes T autoréactifs et la production d'auto-anticorps dirigés contre les cellules hôtes. Ces auto-anticorps sont souvent détectés dans le sang, avant tout signe d'inflammation des articulations, suggérant que le déclenchement de l'auto-immunité pourrait se produire à différents endroits des articulations, par exemple dans le tractus gastro-intestinal [314]. De plus, Kalinkovich et al. ont démontré dans une étude de 2019 que la dysbiose intestinale précède l'auto-immunité et l'inflammation des articulations [315]. Ces données suggèrent que la dysbiose intestinale pourrait jouer un rôle important dans l'apparition de la PR, ce qui signifie que la dysbiose pourrait être plus une cause qu'une conséquence de la maladie. Si la dysbiose précède l'inflammation au niveau des articulations, on pourrait supposer que l'expansion des Proteobacteria pourrait être détectable au début de la PR. Notre étude ne peut pas répondre à cette question car les souches de notre collection de PR ont été isolées chez des individus souffrant de PR depuis plusieurs années avec plusieurs traitements. Il faut également garder à l'esprit que les médicaments pourraient également biaiser la sélection des populations microbiennes dans l'intestin des patients, en plus des effets induits par la pathologie elle-même. La PR est caractérisée par une réponse immunitaire constante localisée au niveau des articulations entraînant une inflammation du tissu synovial et une érosion subséquente du cartilage et des os [316]. Mais l'inflammation n'est pas limitée aux articulations, Nissinen et al. ont démontré une augmentation de l'expression de l'ARNm des marqueurs inflammatoires montrant que les cellules inflammatoires sont également activées dans l'intestin suggérant que l'intestin des patients atteints de PR est enflammé [317]. Cependant, cette étude doit être

interprétée avec précaution car les résultats pourraient avoir été influencée par des traitements anti-inflammatoires qui pourraient avoir biaisé les résultats.

Contrairement à la comparaison des souches SA *versus* PR, les résultats OmniLog ont montré plus de différences pour la comparaison entre les souches SA et MC. L'état de l'inflammation intestinale est assez différent entre ces deux maladies, et l'inflammation est connue pour avoir un impact profond sur l'environnement intestinal. C'est peut-être la raison pour laquelle nous avons trouvé plus de composés différentiels pour la comparaison SA *versus* MC. L'analyse OmniLog avec les souches MC a révélé une meilleure capacité à métaboliser la D-sérine par ce groupe tandis que les souches SA avaient une meilleure capacité à métaboliser le saccharose. L'utilisation de la D-sérine est possible grâce à la présence du locus *dsdCXA* codant pour l'activateur DsdC, le transporteur membranaire interne DsdX et DsdA qui est la D-sérine désaminase/déshydratase [319]. Le locus *dsdCXA* est localisé dans la région chromosomique *argW* qui correspond à la position 53' de la carte génétique d'*E. coli* K-12 MG1655 [320]. Le locus *argW* est un hotspot pour l'intégration d'éléments génétiques mobiles. De façon intéressante, une étude de 2002 a mis en évidence le lien opposé entre le catabolisme de la D-sérine et celui du saccharose. En effet, en comparant les régions *argW* de plusieurs souches d'*E. coli*, les auteurs ont trouvé un niveau de diversité surprenant. Alors que la souche *E. coli* K-12 possède le locus *dsdCXA* dans le site *argW*, d'autres souches, y compris les souches pathogènes O157:H7, ont le locus *csc* responsable du catabolisme du saccharose [321]. Les souches de notre collection présentent également une relation opposée entre la présence des clusters *dsd* et *csc* [321], [319]. En effet, la plupart des souches MC possèdent le locus *dsd* alors que la plupart des souches SA hébergent le locus *csc*. De plus, la capacité à utiliser la D-sérine par les souches MC (et certaines souches SA) était systématiquement associée à la présence du locus *dsd* sans exception. Par ailleurs, certaines souches (205, V3, D2, F3, M2) présentaient un locus *dsd* mais étaient capables de consommer à la fois la D-sérine et le saccharose. Cela pourrait s'expliquer par le fait que le catabolisme du saccharose n'est pas strictement dépendant de la présence du locus *csc*. En effet, le gène *scrB* codant pour une sucros-6-phosphate hydrolase, peut être retrouvé chez les *Enterobacteriaceae* dont *E. coli* soit sur des plasmides soit sur des éléments mobiles au sein du chromosome.

La concentration en D-sérine dans l'intestin est d'environ 1 μ M et que l'apport quotidien de D acides-aminés provenant des aliments est d'environ 100 mg par jour [319]. Lors de ma thèse, nous avons démontré que la plupart des souches d'*E. coli* isolées de l'intestin des patients MC sont capables de cataboliser la D-sérine et que ce phénotype confère un avantage compétitif de la souche MC 204 sur les souches SA T1, R2 et A1 *in vitro*, bien que nous n'ayons pas réussi à

démontrer cet avantage *in vivo* dans notre modèle de rats axéniques. À partir de ces données, nous pouvons supposer que, si la concentration en D-sérine est plus élevée dans l'intestin des patients MC, cela sélectionnerait des souches ayant la capacité de métaboliser la D-sérine et pourrait conduire à une expansion de ces souches dans l'intestin des patients MC. De plus, une étude a révélé que la L-sérine conférait un avantage compétitif aux AIEC, mais uniquement dans l'intestin inflammé [294]. Comme pour la L-sérine, l'utilisation de l'énantiomère D par *E. coli* pourrait être associée à des conditions inflammatoires de l'intestin, expliquant pourquoi nous n'avons pas observé davantage de fitness de la souche MC dans notre modèle expérimental de rat. À l'avenir, il serait intéressant d'effectuer des tests de compétition chez des rongeurs avec une colite ainsi que de quantifier la D-sérine dans le tube digestif pour évaluer une corrélation potentielle entre l'inflammation et la concentration de D-sérine dans l'intestin. Nos travaux mettent en évidence une augmentation de la prévalence du cluster *dsd* dans les souches MC et suggèrent que ce trait phénotypique confère potentiellement un avantage concurrentiel sur les autres membres du microbiote intestinal, contribuant à la prolifération des *Enterobacteriaceae*. D'autres travaux sont cependant requis pour confirmer une telle hypothèse.

PARTIE 2 : Le potentiel inexploré de l'utilisation de l'éthanolamine par les souches commensales d'*E. coli*

Le deuxième objectif de mon projet de thèse a été d'explorer la diversité phénotypique du catabolisme de l'EA dans les souches commensales d'*E. coli*. Notre travail a mis en évidence l'existence de 3 profils différents d'utilisation d'EA parmi la collection de souches SA montrant une grande diversité phénotypique d'utilisation de l'EA. En effet, nos découvertes sur les phénotypes EA chez *E. coli* commensal reflètent la grande variabilité d'*E. coli* en termes de traits phénotypiques. Cette bactérie a une grande plasticité génomique et la famille des *Enterobacteriaceae* s'est avérée être le plus grand contributeur à la diversité génétique dans l'intestin humain [249].

Dans cette partie de mon travail de thèse, je me suis concentré uniquement sur les *E. coli* commensaux de la collection SA. Parmi notre collection, les souches qui ne pouvaient pas cataboliser l'EA avaient une structure du locus *eut* altérée ou un défaut de son expression. Cependant, notre étude n'a pas pu démontrer ce qui diffère entre les souches capables d'utiliser l'EA comme seule source d'azote et les 3 souches capables d'utiliser l'EA comme source d'azote et de carbone. Aucune différence n'a été détectée sur l'organisation du locus *eut* ou l'expression des gènes testés. Il serait intéressant d'explorer si d'autres mécanismes non encore identifiés sont impliqués dans le métabolisme de l'EA par exemple sur la séquestration du composé lui-

même et/ou du cofacteur vitamine B₁₂. Comme il existe une grande diversité phénotypique du catabolisme des EA chez les *E. coli* commensaux, la structure des microcompartiments bactériens pourrait varier entre les profils d'utilisation de l'EA. Mais ceci ne reste qu'une hypothèse et des travaux complémentaires dans ce domaine sont nécessaires pour répondre à ces questions. Notre étude a également démontré que l'utilisation *in vivo* d'EA apporte un avantage de fitness aux *E. coli* commensaux. En utilisant une nouvelle approche développée au cours de ma thèse impliquant un système rapporteur à double fluorescence (Manuscrit 3), d'autres études pourraient être menées *in vivo* pour suivre l'utilisation de l'EA dans l'intestin de la souris à la fois dans le temps et dans l'espace, dans des conditions homéostasiques ou dysbiotiques.

Plusieurs études dans la littérature et des travaux antérieurs dans notre laboratoire avaient porté sur l'utilisation de l'EA par des souches pathogènes d'*E. coli*. En effet, l'EA a été principalement décrite pour son rôle central dans l'adaptation des souches pathogènes dans l'intestin [325],[326]. L'EA peut servir comme source de carbone et/ou d'azote pour favoriser la croissance ainsi que servir de signal pour influencer la virulence lors de l'infection de l'hôte [327]. Bien que la présence du locus *eut* soit répandue chez les bactéries, moins d'attention a été accordée au rôle de ce composé chez les souches commensales d'*E. coli*. Une étude récente de Rowley et al. a démontré que deux souches commensales sont capables d'utiliser l'EA comme source d'azote [298]. Ces résultats couplés à mon travail suggèrent un rôle plus étendu et inconnu pour l'utilisation de l'EA par les souches commensales d'*E. coli*. Une meilleure compréhension du catabolisme de l'EA dans les *E. coli* commensaux pourrait ouvrir des portes sur de nouvelles approches pour limiter l'infection intestinale par des pathogènes pour lesquels l'utilisation de l'EA est importante tels que les EHEC, *Salmonella* ou *Listeria*. En effet, les souches commensales EA⁺ identifiées grâce à mes travaux, pourraient être évaluées à l'avenir en tant que souches compétitrices pour limiter la colonisation intestinale par les pathogènes intestinaux et les maladies associées.

Même si l'EA n'était pas présent dans les plaques PM01 et PM02 testées avec la technique OmniLog, nous avons testé notre collection de souches MC et PR pour l'utilisation de ce composé. Nous avons observé que la plupart des souches consomment l'EA comme source d'azote quels que soient les groupes de souches mais le phénotype capable d'utiliser l'EA comme source de carbone était absent parmi notre collection de souches MC et PR. Cependant, nous devons garder à l'esprit que cette absence résulte potentiellement du nombre réduit de souches dans ces collections. De façon intéressante, l'EA est un composé qui a déjà été mis en évidence dans la littérature pour avoir son importance pour les deux pathologies. De plus, l'EA est très

abondante dans l'intestin inflammé à cause de la desquamation importante des cellules épithéliales dans la lumière intestinale. Dans une étude de 2019, Ormsby et al. ont déterminé que l'acide propionique agit comme un signal pour les AIEC et modifie leur métabolisme et augmente leur utilisation de l'EA [326]. À ce jour, la réplication rapide des AIEC dans les macrophages est inexplicée, mais il semble que la croissance peut être facilitée par la présence de niveaux extracellulaires d'EA comparables à ceux de l'intestin humain. Aussi, Shan et al. (2018) ont identifié l'EA comme un marqueur différentiel présent dans l'intestin des rats développant une PR par rapport au groupe témoin [328]. Cette étude a même mis en évidence une corrélation entre l'augmentation des *Enterobacteriaceae* observée dans le modèle *in vivo* et la variation de la concentration en EA dans l'intestin. Le rôle de l'EA semble donc avoir son importance dans la physiopathologie des deux maladies.

En résumé, l'objectif de mon travail de thèse était d'identifier des capacités métaboliques différentielles entre les entérobactéries isolées de SA ou de patients atteints de MC ou de PR comme hypothèse expliquant leur expansion dans l'intestin dysbiotique. Parmi les substrats utilisés différemment entre les souches SA et celles isolées de patients MC ou PR, l'utilisation préférentielle de la D-sérine par les souches MC est une hypothèse prometteuse à suivre. D'autres études sont cependant nécessaires, afin d'évaluer si la D-sérine confère un avantage nutritionnel aux souches MC dans l'intestin et dans quelles conditions cela se produit. De même, il serait de grand intérêt d'analyser l'efficacité des souches MC et PR à utiliser l'EA, une caractéristique qui pourrait participer à un meilleur fitness de ces souches par rapport aux souches SA. Comprendre les mécanismes par lesquels les *Enterobacteriaceae* prolifèrent dans l'intestin des patients est un préalable pour envisager des stratégies visant à limiter leur expansion et par extension, les conséquences néfastes associées pour l'hôte.

TABLE OF CONTENTS

TABLE OF CONTENTS.....	1
TABLE OF FIGURES.....	3
TABLE OF TABLES	4
LIST OF ABBREVIATIONS	5
LITERATURE REVIEW.....	7
The intestinal microbiota of the healthy individual.....	9
1. Anatomo-physiology of the GIT and microbiota distribution	10
1.1 Longitudinal distribution of the microbiota along the GIT.....	10
1.2 Sectional distribution of the microbiota along the GIT	12
2. Structure and composition of the human gut microbiota	13
2.1 Composition of the gut microbiota of the healthy adult	13
2.2 Factors affecting gut microbial composition	15
A) Host factors.....	15
B) Environmental factors	17
3. Function of the normal gut microbiota	19
3.1 Role of the gut microbiota in metabolism.....	19
3.2 Structural role of the gut microbiota.....	21
3.3 Antimicrobial protection against pathogens	21
Gut microbiota and diseases	25
1. Intestinal inflammatory diseases : example of Crohn's disease	25
1.1 Etiology and pathophysiology of Crohn's disease.....	25
1.2 Dysbiosis in Crohn's disease.....	27
2. Extraintestinal diseases : example of Rheumatoid Arthritis	30
2.1 Etiology and pathophysiology of Rheumatoid Arthritis	30
2.2 Dysbiosis in Rheumatoid Arthritis.....	34
Proteobacteria in the gut microbiota.....	39
1. Proteobacteria in the healthy human gut.....	39
1.1 Focus on <i>E. coli</i> in the healthy human gut.....	40
A) Genetic and phylogenetic diversity of <i>E. coli</i>	41
B) <i>E. coli</i> residency in the GIT	43
C) Functional roles of <i>E. coli</i> in the GIT	44
a) Oxygen consumption to maintain an efficient anaerobic environment in the gut	45
b) Production of vitamins	45
c) Protection against gut pathogen infections	46
2. Proteobacteria expansion in dysbiosis associated with human disease	47
2.1 Proteobacteria bloom : a marker of gut microbiota in disease	47
2.2 Causes of Proteobacteria proliferation in dysbiotic conditions	49
A) Increase of electron acceptors for respiration	49
B) Metabolic adaptation to dysbiotic conditions	51
a) Glycerol in cystic fibrosis patient	52

b) Intestinal mucosa-derived substrates.....	52
c) Diet-derived substrates.....	53
RESULTS	55
Identification of metabolic properties potentially involved in the expansion of <i>Enterobacteriaceae</i> in the gut of patients suffering from CD or RA.....	57
MANUSCRIPT 1.....	61
The unexplored potential of Ethanolamine utilization by commensal <i>E. coli</i> strains	85
MANUSCRIPT 2.....	89
MANUSCRIPT 3.....	119
DISCUSSION.....	145
SCIENTIFIC CONTRIBUTIONS.....	155
ANNEX.....	157
REFERENCES	167

TABLE OF FIGURES

Figure 1 Major phyla of the human gut microbiota and their predominant genera.....	10
Figure 2 Longitudinal and axial bacterial distribution of the gut microbiota.....	11
Figure 3 Phylogenetic tree of the human gastrointestinal microbiota.....	14
Figure 4 Evolution of the composition of the gut microbiota from the womb to infancy.....	16
Figure 5 Different factors affecting the gut microbiota composition and how resilient and non-resilient microbiota reacts.....	18
Figure 6 Schematic representation of the functions of the gut microbiota.....	19
Figure 7 Differences in microbial abundance of bacterial families among CD patients.....	28
Figure 8 Different factors participate in the etiology of RA.....	31
Figure 9 Organization of the Proteobacteria phylum with several examples of Class, Order, Family, Genus and Species.....	40
Figure 10 Phylogenetic history of <i>Escherichia coli</i>	42
Figure 11 <i>E. coli</i> pan-genome and core genome evolution according to the number of sequenced genomes.....	43
Figure 12 Expansion of Proteobacteria in gut microbiota under different host conditions....	47
Figure 13 Changes in the gut microbiota associated with inflammation of the lower GI tract..	48
Figure 14 Mechanisms of expansion of <i>Enterobacteriaceae</i> in the dysbiotic gut.....	49
Figure 15 Phylum-level changes in the microbiota composition on intestinal inflammation...	50
Figure 16 <i>Enterobacteriaceae</i> isolated on McConkey medium represented by CFU/g feces..	57
Figure 17 Invasion tests on epithelial cells using our HI and CD collection of strains.....	58
Figure 18 HS outcompetes EHEC specifically during growth on EA.....	85
Figure 19 EA utilization profiles of commensal <i>E. coli</i> strains from HI, CD and RA collection.....	86
Figure 20 EA catabolism and organisation of the <i>eut</i> operon in <i>E. coli</i>	88
Figure 21 The gut-articulation axis.....	148
Figure 22 Schematic illustration of the differences between the <i>argW</i> region and D-serine catabolism pathway.....	150

TABLE OF TABLES

Table 1 Alterations of gut microbiota related with RA.....	35
Table 2 Murine models of RA known to be correlated with the gut microbiota.....	37
Table 3 Prevalence of the main <i>E. coli</i> groups in humans.....	44
Table 4 Studies available reporting nutrient sources available to <i>Enterobacteriaceae</i>	51

LIST OF ABBREVIATIONS

AA : Amino acids	IBD : Inflammatory bowel disease
ACPA : Anti-citrullinated protein antibody	IL : Interleukine
AIEC : Adherent and invasive <i>E. coli</i>	LPS : Lipopolysaccharide
AMP : Antimicrobial peptides	NOD : nucleotide-binding oligomerization domain
AUC : Area under the curve	PCR : Polymerase chain reaction
CD : Crohn's disease	RA : Rheumatoid arthritis
CFU : Colony-forming unit	rRNA : Ribosomal ribonucleic acid
DNA : Deoxyribonucleic acid	RF : Rheumatoid factor
EA : Ethanolamine	RT-qPCR : Reverse transcription quantitative real-time PCR
ECOR : <i>E. coli</i> reference	SPF : Specific-pathogen-free
EHEC : Enterohemorrhagic <i>E. coli</i>	TLR : Toll-like receptor
EPEC : Enteropathogenic <i>E. coli</i>	TNF- α : Tumor necrosis factor type α
eut : Ethanolamine utilization	SCFA : Short-chain fatty acid
GF : Germfree	SE : Shared epitope
GIT : Gastrointestinal tract	UC : Ulcerative colitis
GWAS : Genome-wide association study	
HI : Healthy individual	
HLA : Human leukocyte antigen	

LITTERATURE REVIEW

CHAPTER 1

The intestinal microbiota of the healthy individual

All multicellular organisms including humans, live in close association with microbes. Indeed, every surface of the human body that is in contact with the environment is colonized by microorganisms. Among the different human organs, the gastrointestinal tract (GIT) is the most colonized with 10^{14} microorganisms, referred as the gut microbiota, which encompasses ~10 times more microbes than the total number of human cells. The gut microbiota is mainly composed of bacteria but also includes viruses, fungi, archaea and protozoans [1]. Even though the first studies of this ecosystem date from the mid-80s, complexity of the gut microbiota was discovered with the onset of molecular biology techniques in the 2000s. The pioneer programs MetaHIT and HMP gave essential information concerning the composition of the gut microbiota [2],[3]. These studies determined the microbial richness and composition of a limited set of individuals by 16S rRNA gene sequencing and metagenomic profiles generated by whole-genome shotgun sequencing. Indeed, the 16S rRNA gene can be found in all prokaryotes and it is at once conserved enough for an accurate alignment but also distinct enough to detect variation for taxonomic analysis.

To date, more than 1000 different bacterial species inhabiting the gut have been described, with about 160 species being found in the gut of any individual [4]. Taxonomically, a limited number of bacterial phyla are represented in the human gut: Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, Euryarcheota, and Verrucomicrobia. The predominant genera of each phylum are indicated in [Figure 1](#). In addition to the discovery of the composition of the healthy gut, several studies highlighted the important function of the gut microbiota on the wellbeing of an individual. A well-balanced gut microbiota plays an important role on human health participating on immunological and gut protective functions in the healthy individual. Indeed, the gut microbiota imparts specific functions such as nutrient metabolism, xenobiotic and drug metabolism, maintenance of structural integrity of the gut mucosal barrier, immunomodulation, and protection against pathogens that play key roles on human health [5],[6].

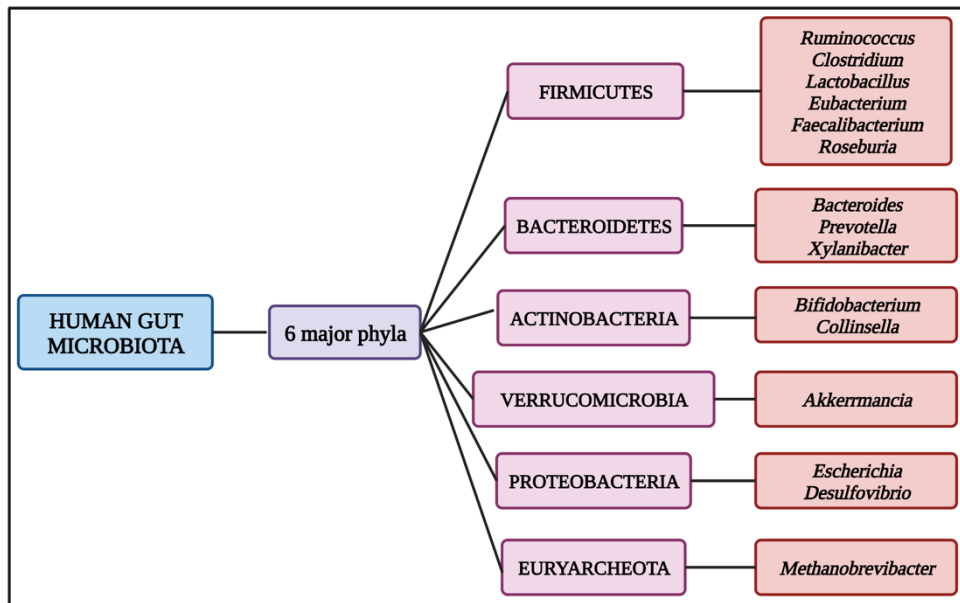


Figure 1 Major phyla of the human gut microbiota and their predominant genera. Adapted from Bliss and Whiteside, 2018. Figure made with biorender.com.

1. Anatomico-physiology of the GIT and microbiota distribution

1.1 Longitudinal distribution of the microbiota along the GIT

The GIT is composed of different parts: the esophagus, the stomach, the small intestine and the colon which all possess specific and different microbiota (Figure 2).

The Stomach The conditions were thought to be inhospitable for microorganisms in the stomach due to low pH and other antimicrobial factors, but it has been demonstrated that acid resistant bacterial populations can be found [7]. The microbial mass in the stomach is relatively low estimated at 10^3 CFU/g. In the stomach, the gastric juice composed of proteolytic enzymes and hydrochloric acid, has an important digestive and protective role [8]. The low pH of 1-2 is the main restrictive component of the gastric juice [9]. This condition facilitates denaturation of proteins and nutrient absorption but limits colonization and survival of bacteria. The transit time of food in the stomach is around 3 hours in total [10]. Culture-based studies found out that the most prevalent phyla in the gastric ecosystem are Firmicutes, followed by Proteobacteria and Bacteroidetes [11]. In terms of genera, the most commonly found are *Streptococcus*, *Lactobacillus*, *Bacteroides*, *Staphylococcus*, *Veillonella*, *Corynebacterium*, and *Neisseria* [12].

The small intestine The small intestine is divided into 3 parts the duodenum, jejunum and ileum and has a less acidic pH than the stomach (pH = 5-7). It is responsible for approximately 90% of the body's total nutrient absorption from the diet and plays an important role in maintaining immune homeostasis [13]. The small intestine is characterized by a short transit time (2-4h), a regular flow of digestive enzymes and bile and high oxygen concentrations [14]. For these several reasons, the microbial biomass is relatively low around 10^4 – 10^5 cells per milliliter of effluent [15],[16]. The host limits microbial growth for instances by secretion of bile salts into the duodenum which are bactericidal [17]. Nonetheless, we can find a bacterial community specific to the small intestine. Bacterial genera commonly found in the small intestine is composed of *Bacteroides*, *Clostridium*, *Streptococcus* and some species of *Lactobacillus* [18].

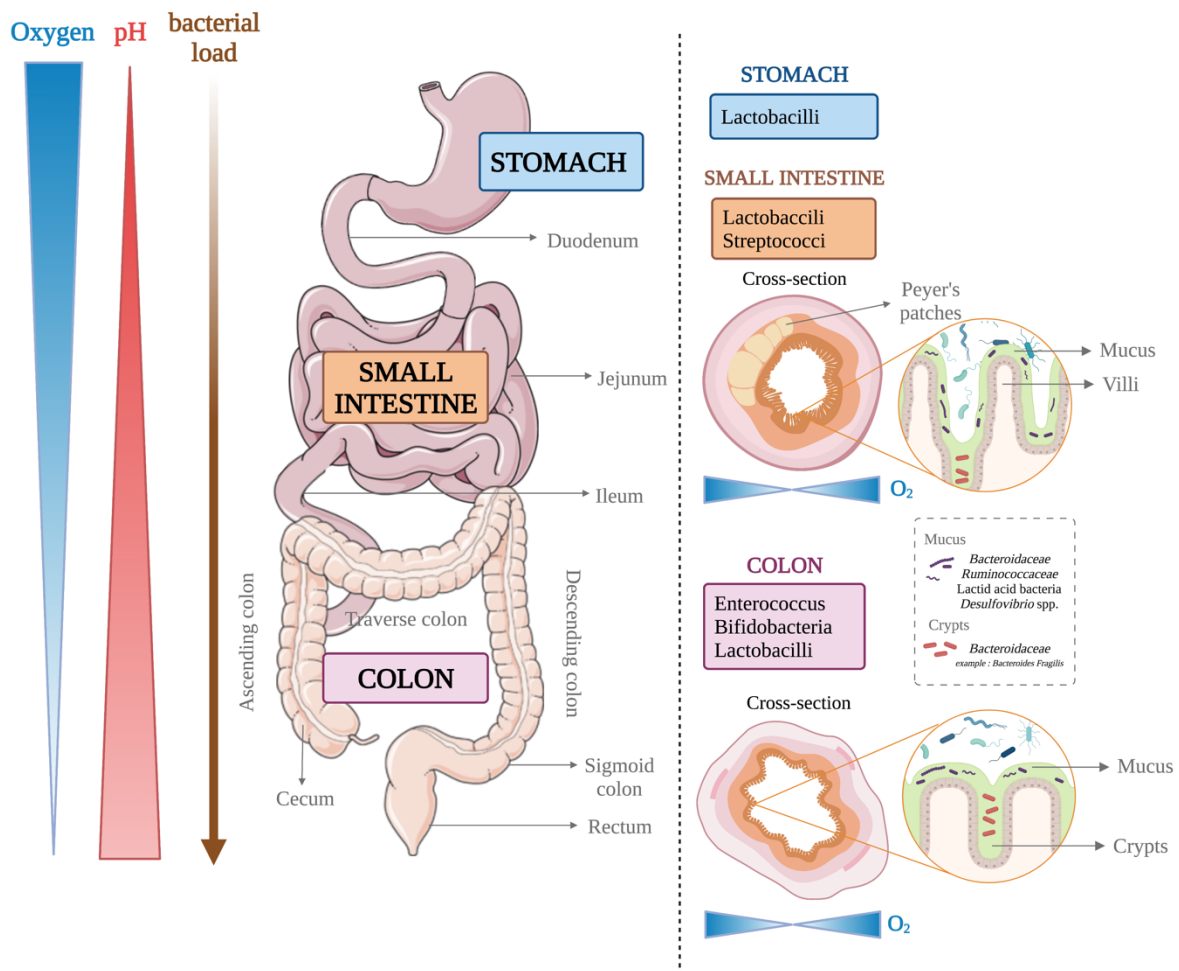


Figure 2 Longitudinal and axial bacterial distribution of the gut microbiota. Adapted from Pereira et al., 2017 ; Sartor, 2998 ; Donaldson et al., 2016. Figure made with biorender.com.

The colon The colon microbiota represents over 70% of all microbes found in the body with a microbial biomass of 10^{11} cells per gram of content. The human colon consists of the ascending, transverse, descending and sigmoid colon as well as the rectum [19]. An important role of the colon is the resorption of water and electrolytes while the microbiota contributes to the digestion of undigested dietary ingredients. Here the conditions are less severe for microbial proliferation : the pH is close to neutrality and there is a lower concentration of bile salts. The structure of the mucosa is also different in the large intestine (no Peyer's patches) which enables the host to tolerate such a high density of bacteria [18]. The long transit times (24-72h) in the large intestine can have a large impact on bacterial physiology and metabolism. *In vivo* and *in vitro* studies show that the transit time in the large intestine leads to breakdown of proteins, sugars, amino acids, but mostly fibers marking an increased contribution to the short-chain fatty acids (SCFAs) available in the colon. These SCFAs contribute to shape the gut environment, acting notably as a source of energy for the host [18].

1.2 Sectional distribution of the microbiota along the GIT

In addition to the gradient along the length of the GIT, there are also differences in the microbiota composition along the sectional axis. From the epithelium to the lumen, a gradient of oxygen, pH, carbohydrates, and antimicrobial peptides (AMPs) can be defined (Figure 2). Due to these different environmental conditions of the intestine, we can define specific bacterial populations belonging to different ecological niches from the lumen to the mucosal surface.

The mucosa in the intestinal tract is composed of a monolayer of epithelial cells that overlay the lamina propria (a layer of connective tissue) and the muscularis mucosae (a thin layer of muscle). The intestinal cells secrete a layer of mucus acting as a physical and biological barrier that protects the epithelium from the gut content. The mucus consists of high molecular weight glycoproteins that is a source of nutrients for some gut bacteria. Compared to the gut lumen, the mucus covering the gut mucosa harbors fewer bacteria with a low density ranging from 10^5 - 10^6 cells [20]. A key factor that determines which microorganisms can be found in the mucus layer is the ability to metabolize the glycans and peptide backbone of mucin glycoproteins. In the mucus, species such as *Bacteroides fragilis* and *Akkermansia muciniphila* can be found [21]. These bacteria develop a close relationship with the intestinal epithelium more than any other bacteria and affect host's health even though more knowledge is needed regarding its mechanisms. The thickness of the mucus layer varies in the GIT. It increases along the length of the small intestine remaining patchy while it is dense, thick, and continuous in the colon. In

the large intestine the mucosa is smooth and lacks villi structures but contains proliferative crypts (**Figure 2**). The colonic crypts correspond to invaginations of the epithelium creating a niche environment characterized by a higher concentration of oxygen as well as glycans compared to the lumen. These structures form a specific ecological niche where only mucin-degrading aerotolerant bacteria such as *Bacteroides fragilis* can be found at low density close to the mucosal surface [22]. Furthermore, the mucus contains AMPs and secretory immunoglobulins that shape the composition of the mucosa-associated microbiota [23]. Due to these different factors, bacteria belonging to *Bacteroides*, *Bifidobacterium*, *Streptococcus*, *Enterobacteriaceae*, *Enterococcus*, *Clostridium*, *Lactobacillus* and *Ruminococcus* are mostly found in the luminal microbiota while *Clostridium*, *Lactobacillus*, *Enterococcus*, *Escherichia*, and *Akkermansia* can be found in the intestine mucosa [24].

2. Structure and composition of the human gut microbiota

As described above, the microbiota is different between the distinct parts of the GIT. From now on, I will only describe and refer to the colonic microbiota, except when clearly indicated. It is also important to note that the majority of studies focusing on the characterization of the distal gut microbiota relies on stool specimens because it's easier to collect samples [25],[26]. Although easily acquired, fecal samples do not represent accurately the composition of the gut microbiota as well as the composition of the bacterial species present at mucosal level [27]. However, it is admitted that fecal microbiota and colonic microbiota have qualitative similarities even if differences have been identified [28].

2.1 Composition of the gut microbiota of the healthy adult

As previously mentioned, the colonic microbiota of the healthy adult is composed of 6 main phyla: Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, Euryarcheota, and Verrucomicrobia [27] (**Figure 3**). The phyla Firmicutes and Bacteroidetes are always highly represented in the healthy individual accounting for more than 90% of the bacterial population in the colon. The Bacteroidetes phylum is composed of anaerobic Gram-negative bacteria, the predominant genera being *Bacteroides* and *Prevotella*. The *Bacteroides* genus is one of the most predominant groups in the intestine representing about 30% of gut microorganisms, indicating that this genus has important functions [29],[30]. The Firmicutes are composed of obligate and facultative anaerobic Gram-positive bacteria. These bacteria can form endospores, a resistant form that helps survive adverse conditions in the gut. This phylum comprises species

belonging to the genera *Eubacterium*, *Clostridium*, *Ruminococcus*, *Roseburia*, *Lactobacillus* and *Butyrivibrio*.

Even though the taxa composition at the phylum level is similar between individuals, the species diversity found in each person is highly specific to the individual [31]. This individual specific population is acquired due to the diet of the individual and environmental factors [32]. Even though there is a fecal print specific to each individual, a core microbiota among all individuals can be defined. The core microbiota corresponds to the bacterial species shared by most healthy individuals, corresponding to approximately 57 species present in 90% of individuals [33]. Several factors indicate a healthy gut microbiota such as great diversity and richness but also the presence of microorganisms that are known to enhance metabolism and promote resilience to infection and inflammation [1],[27],[34]. Microorganisms that have a positive impact on health are *Bacteroides*, *Bifidobacterium*, *Clostridium* clusters XIVa and IVa (butyrate producers), *Eubacterium*, *Faecalibacterium*, *Lactobacillus* and *Roseburia* [27].

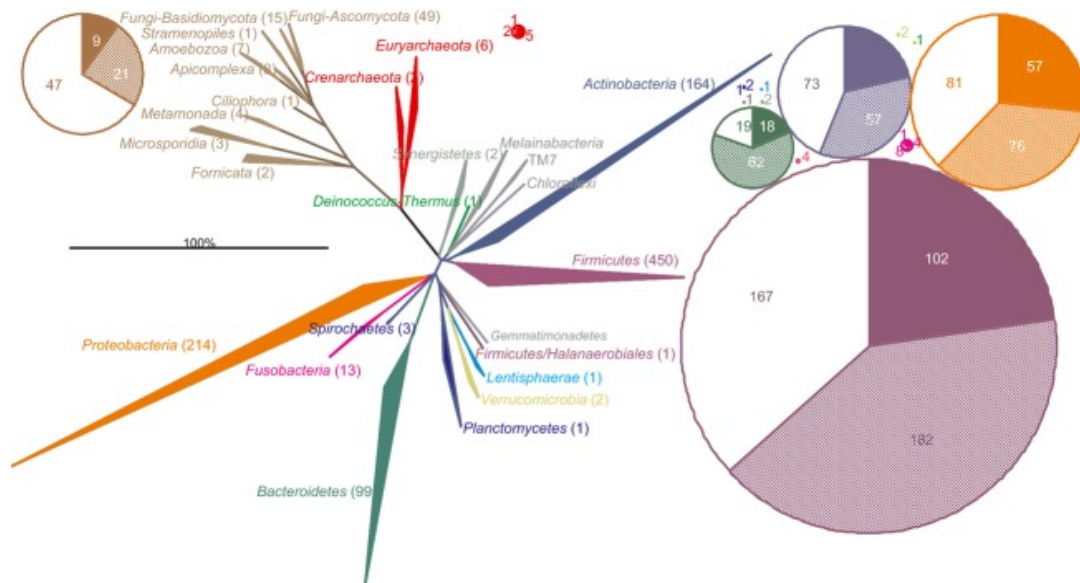


Figure 3 Phylogenetic tree of the human gastrointestinal microbiota. The numbers in parentheses indicate the number of cultured species given per phylum. The pie charts illustrate distribution between the number of species with full genome sequence genome (full sectors), the number of species with partial genome sequence (semi-full sectors) and number of species without any genome sequence (empty sectors) given for Archaea, Eukarya and per phylum for Bacteria. The color code of pies corresponds to the color code of the phylogenetic tree. From Rajilić-Stojanović and de Vos, 2014.

2.2 Factors affecting gut microbial composition

The gut microbiota of the healthy individual is mostly stable during adult life but this community is constantly exposed to perturbations. Perturbations of the gut microbiota disrupt the stability of the ecosystem and can possibly result in unhealthy states associated with disease [35]. There are different factors that impact the gut microbiota composition and diversity : host factors and environmental factors.

A) Host factors

Several host factors can impact the gut microbiota composition such as age [36], immune system dysfunction [37], and production of AMPs [38].

Age During lifetime, the gut microbiota is quite stable in the healthy adult however it varies mostly during the first years of life (**Figure 4**). The exact moment when colonization occurs remains controversial. Gosalbes *et al.* have successfully detected bacterial DNA in the meconium¹ of full-term healthy neonates. These authors showed that the meconium microbiota is a simple community mainly dominated by genera such as *Escherichia*, *Shigella*, *Enterococcus*, *Leuconostoc*, *Lactococcus* and *Streptococcus* [39]. Upon delivery, the neonate is exposed to a wide variety of microorganisms. Several factors influence the composition of the newborn's microbiota such as the kind of delivery during birth. When a child is born vaginally, the gut is mainly colonized by microorganisms from the genera *Lactobacillus* and *Prevotella* originated from the maternal vagina. On the other hand, when a child is born by cesarian, the gut microbiota resembles the human skin microbiota dominated by *Streptococcus*, *Corynebacterium*, and *Propionibacterium* [40],[41]. Differences of the initial community impact the progression of the composition and the immunological functions of the gut microbiota. For instances, culture-based studies have shown that the intestinal colonization by *Lactobacillus*, *Bifidobacterium*, and *Bacteroides* in infants born by cesarian is delayed [42]. Also, recent studies suggest that cesarian newborns may be more susceptible to allergies and asthma impacting the development of the gut immune system [43].

Another event that highly impacts the newborn's microbiota depends if the newborn is breastfed or formula fed. Breastfeeding contributes to a gut microbiota mainly dominated by *Bifidobacterium* and *Bacteroides* species that are essential for the metabolism of milk oligosaccharides [44],[45]. Whilst formula fed newborn's have a more diversified gut

¹ Meconium : the first feces of the newborn.

microbiota dominated by *Clostridium*, *Granulicatella*, *Citrobacter*, *Enterobacter* and *Bilophila* [46]. It has been demonstrated that breast milk has different advantages for health and the development of intestinal immune function of the newborn. It provides protective factors such as polymeric Immunoglobulin A, antibacterial peptides, and components of the innate immune response [47]. Compared with formulas, breast milk has greater impact in barrier integrity and mucosal defenses of the intestinal tract [48]. Indeed, it affects health promoting microorganisms such as *Lactobacillus*, *Bacteroides*, and *Clostridia* which active immunological functions such as mucus production, and tight junction expression providing mucosal barrier homeostasis [49]. After the introduction of solid food, the gut microbiota develops rapidly in response to different factors such as the type of nutrition, geographical localization, and lifestyle. Consequently, the microbiota is no longer represented by Bifidobacteria and Lactobacilli. It increases in richness and diversity to resemble to a more adult-like microbiota. When a child reaches 3 year old, the microbiota resembles 40 to 60% to the microbiota of the healthy adult (**Figure 4**) [36].

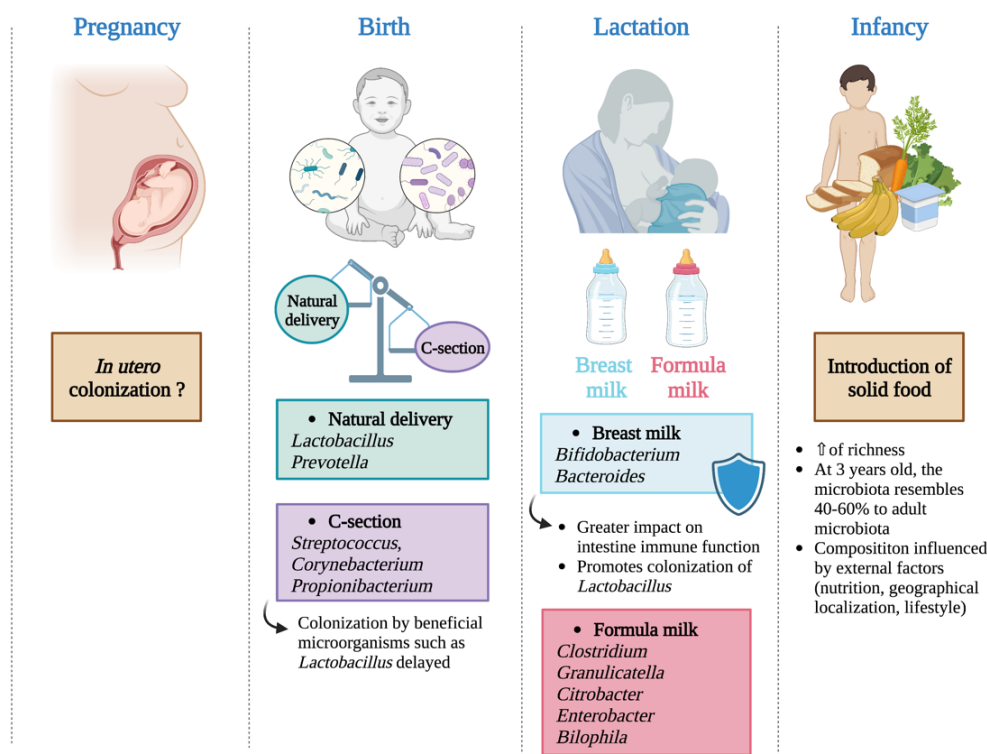


Figure 4 Evolution of the composition of the gut microbiota from the womb to infancy. Adapted from Kalbermatter et al., 2021. Figure made with biorender.com.

Immune system dysfunction The mammalian immune system corresponds to a complex network of innate and adaptive components in all tissues of the body, and plays a vital role in host defenses against various potentially harmful external agents and endogenous perturbations of homeostasis. The host and the gut microbiota co-evolved towards a balanced relationship and homeostasis [50]. Such intimate relationship requires proper functioning of the host's immunity to prevent commensals from over-exploitation of host resources while maintaining immune tolerance. In addition to the regulation of infection and commensal spread, microbiota-immune interactions are implicated in a variety of diseases including inflammatory bowel disease (IBD) [51] as well as extra-intestinal diseases such as rheumatoid arthritis [52]. The interactions between the gut microbiota and the host's immunity are complex, dynamic and context-dependent.

AMPs Among host factors, molecules such as AMPs, and immunoglobulin A are produced by intestinal cells which can encourage or inhibit growth of certain microbial species [38]. The epithelial cells of the host secrete a layer of mucus composed of a mucin glycoprotein that is actively secreted from goblet cells and assembled into a viscous gel like layer on the intestinal epithelia. The mucus has a key role in shaping gut microbiota and selecting microbial species beneficial for the host [53]. In the small intestine where mucus is less present, it's the AMPs that shape gut microbiota [54].

B) Environmental factors

Several environmental factors can affect the microbial communities, as illustrated in **Figure 5**, such as extreme dietary changes [55], antibiotics, lifestyle (smoking [56], exercise [57]), and physiological stress and anxiety [58]. Most of the time, the altered gut microbiota will go back to its initial state after exposure to a stress agent, and this is called resilience [59]. However, the gut microbiota may also not be able to return to its original state and this non-resilient microbiota will shift to an altered new state termed dysbiosis (**Figure 5**). Little is known about the different mechanisms involved in the phenomenon and the proposals in the literature remain theoretical. What distinguishes a resilient from a non-resilient microbiota is the diversity of its composition and the host immune status. For instances, in a human study, Raymond *et al.* showed that perturbations by antibiotics are weaker when microbiota diversity is higher [60]. On two different mouse studies, authors have shown that ablation of the bacterial sensor NOD2 leads to impaired recovery of the microbiota from antibiotic perturbation [61],[62]. These two

studies illustrate the importance of the stability of host-microbiota symbiosis to the ability to return quickly and fully to state of homeostasis.

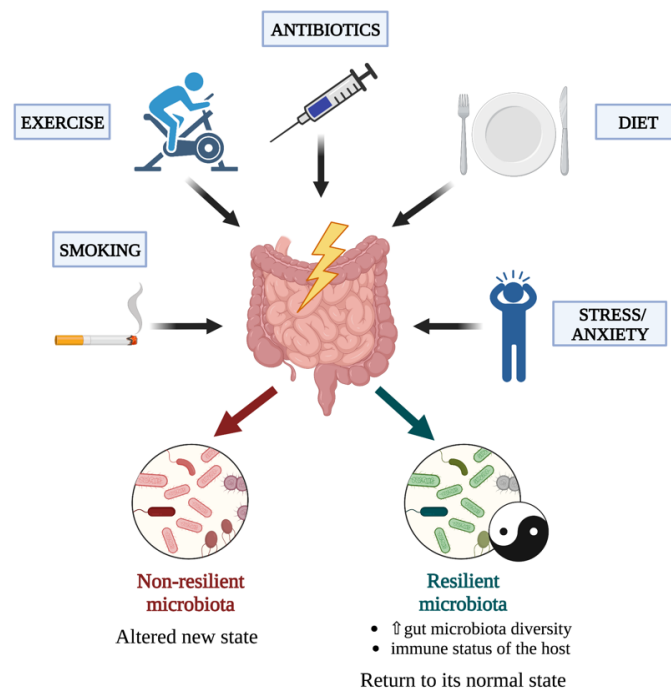


Figure 5 Different factors affecting the gut microbiota composition and how resilient and non-resilient microbiota reacts. Adapted from Hasans and Yang, 2019. Figure made with biorender.com.

The diet is one of the main factors that modulates gut microbiota either in a useful or harmful way. Diet and metabolites derived from microbial fermentation of nutrients are key modulators of the composition of the gut microbiota that influence directly different biological processes and host homeostasis [63]. The diet can influence the availability of nutrients present in the gut and, consequently, the increase or decrease of some species from the gut microbiota [64]. For instances, dietary fibers are important energy sources for the colon-residing microbiota. The consumption of a diet rich in fibers is known to promote fermentation by SCFAs-producing bacteria. Dietary fibers and SCFAs (mainly acetate, propionate and butyrate) stimulate mucus production and secretion keeping the integrity of the mucosal barrier function [65]. Both acetate and butyrate maintain a balance for mucus production and secretion. *Bacteroides thetaiotaomicron*, an acetate and propionate producer, promotes Goblet cell differentiation and expression of mucin-related genes [66]. A long-term consumption of plant-based diets have been associated with an increase in some bacterial taxa such as *Ruminococcus*, *Enterococcus rectale*, and *Roseburia* that are plant cell wall degraders. Furthermore, dietary fibers can also mechanically stimulate the intestinal epithelium to secrete mucus [67]. This type of diet has an anti-inflammatory and anti-pathogenic effect [68]. By contrast, a diet rich in

proteins and fats, has been correlated with an abundance of bile-tolerant species, such as *Bacteroides*, *Bilophila* and *Alistipes* and a suppression of Firmicutes [55],[69]. The consumption of a high fat/high sugar diet, usually common in Europe, increased susceptibility to infection and developing metabolic diseases [70]. Animal-based diets like the Western diet are usually energy-full characterized by a high intake of fat, and sugar and a low intake of fibers, fruits, and vegetables. This type of diet promotes higher proportions of *Bacteroides* species that are not able to digest fibers but use protein, fat, and simple sugars for their growth [71].

3. Function of the normal gut microbiota

Commensal bacteria co-habit with the enterocytes in a symbiotic relationship and bring several advantages to the host. The gut microbiota plays an essential role in the health through main functions including: supply of energy for the host, protection against external pathogens, maintenance of the epithelial structure and development and maturation of the immune system (Figure 6) [72].

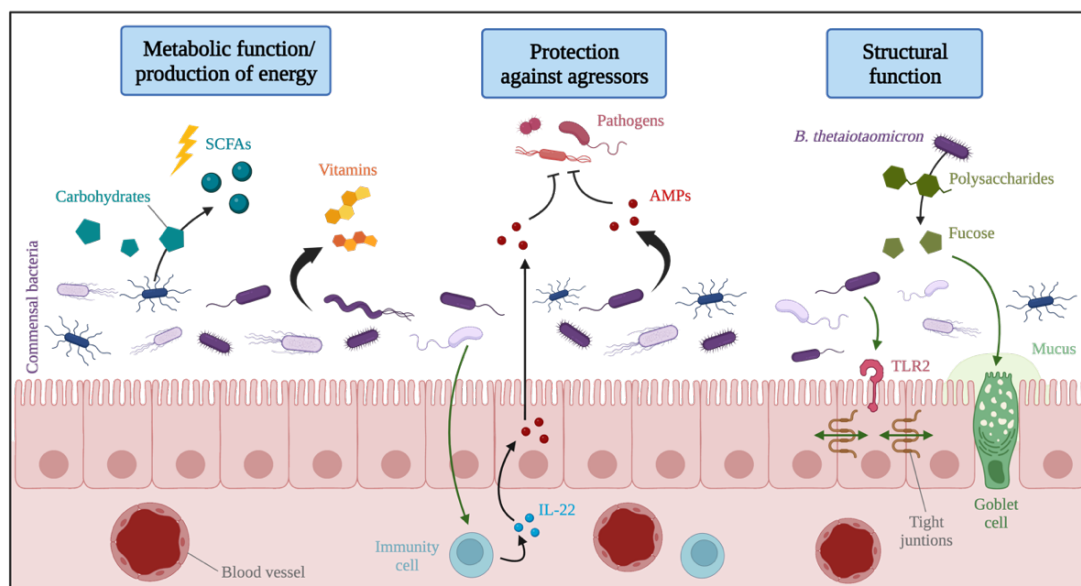


Figure 6 Schematic representation of the functions of the gut microbiota. The commensal bacteria have important metabolic functions such as fermentation of carbohydrates producing short chain fatty acids and vitamins useful for the host. Furthermore, commensal bacteria defend the host against pathogen infection by producing antimicrobial peptides or inducing the immune system. And lastly, commensal bacteria play a role on barrier fortification through induction of mucus production by the Goblet cells and stimulation of TLR2 leading to fortification of tight junctions. Adapted from O'Hara and Shanahan, 2006. Figure made with biorender.com.

3.1 Role of the gut microbiota in metabolism

The metabolic activity of the gut microbiota is a key function that produces different molecules with various host effects. Degradation and fermentation of dietary compounds by

commensal bacteria produces several nutrients and metabolites such as amino acids, vitamins, SCFAs that are mostly absorbed by the intestinal mucosa and benefits to the host [73]. For example, around 10% of the total energy from ingested food is produced thanks to the colonic microbiota.

Dietary carbohydrates Some carbohydrates cannot be digested and absorbed in the small intestine and arrive in the large intestine where they are fermented by the intestinal microbiota. These carbohydrates are mainly from plant sources and include plant cell wall polysaccharides (hemicelluloses, cellulose, pectin), resistant starch and some sugars. Fermentation of these substrates by microorganisms leads to the production of SCFAs, mainly acetate, propionate and butyrate. 98% of produced SCFAs are absorbed and 2% are excreted in faeces [74],[75]. Firmicutes and Bacteroidetes have an important role on the digestion of complex carbohydrates that the host digestive enzymes are not able to breakdown. Among butyrate producers, there is mainly species from the *Clostridium* clusters IV and XIVa such as *Roseburia*, *Eubacterium*, and *Faecalibacterium*. [76],[77],[78]. SCFAs have several health benefits for the host playing an important role in biological processes influencing intestinal epithelial cell proliferation and differentiation, and also modulation of the immune system [79]. In particular, butyrate and propionate regulate intestinal physiology and immune function whereas acetate participates in lipogenesis and glucogenesis [80].

Vitamins Vitamins are organic nutrients essential for the host, that act as coenzymes for many different enzymes or as antioxidants. The diet of the host provides vitamins which are absorbed in the small intestine but the gut microbiota also participates in the production of vitamins in the large intestine. Some species of *Bacteroides* can synthesize vitamins *de novo* [81],[82]. The gut microbiota is capable of producing vitamin K2, water-soluble B-vitamins such as folic acid, niacin, biotin, pantothenic acid, cobalamin, pyridoxine, riboflavin, and thiamine [83],[84].

Proteins and Amino acids Dietary protein are mainly absorbed in the small intestine. In the colon, the main source of protein comes from endogenous substrates such as mucins, enzymes, and desquamated cells. The colonic gut microbiota transforms these proteins and amino acids (AAs) sources into biogenic amines, immunomodulatory compounds and other signaling

molecules such as antimicrobial peptides thanks to decarboxylation [27]. Even though essential AAs are acquired through diet and absorption via amino acid transport proteins in the intestine, extraction, synthesis and absorption of AAs such as alanine, aspartate, glutamate, glycine and tryptophan depends of the commensal bacteria activity.

3.2 Structural role of the gut microbiota

The intestinal epithelium is composed of a single layer of cells which are tightly bound between them thanks to different complexes that regulate permeability. There are different types of structures : tight junctions, adherent junctions and desmosomes [85]. The gut microbiota plays an important role on the maintenance of these structural components of the gut mucosal barrier. For instances, tight junctions are maintained by the toll-like receptor 2 (TLR2) mediated signaling that is stimulated by the microbial cell wall peptidoglycan [86]. The mucus layer keeps luminal bacteria away from epithelial contact but also provides a direct source of nutriment for gut bacteria. For example, *Bacteroides thetaiotaomicron* can break down mucosal polysaccharides into fucose thanks to multiple fucosidases. The activity of this mucinolytic bacteria participate to the structuration of the intestinal epithelium because fucose is capable of inducing the secretion of mucin by Goblet cells [87]. It has also been reported that *B. thetaiotaomicron* induces expression of the small proline-rich protein 2A which is required for maintenance of desmosomes at the epithelial villus [88].

The importance of the gut microbiota in the structure of the GIT has mainly been demonstrated in germ free (GF) rodents [89]. The lack of microbiota on these rodents have a large impact in the host capacity to absorb intestinal nutrients. The development of the intestinal microvasculature is induced by the transcription factor angiogenin-3 which is activated by the gut microbiota. This is also supported by a significant reduction of villus capillary network in GF mice, which in turn can impair nutrient digestion and absorption [90]. Colonocytes can also use butyrate produced by the gut microbiota for energy helping to strengthen the colonic defense barrier by inducing mucin secretion and production of antimicrobial peptides. Moreover, this metabolite participates in cell cycle inhibition, induction of programmed cell death and cellular differentiation [78],[91].

3.3 Antimicrobial protection against pathogens

Another important role of the gut microbiota is resistance against pathogen colonization. There are different protection mechanisms by which the commensal microbiota acts against

colonization by pathogens : nutritional competition, production of antimicrobial compounds, and induction of the host immunological responses.

Nutritional competition One of the ways of protecting the host against bacterial infections is based on nutritional competitions that can occur between pathogens and the gut microbiota. The preferential consumption of nutrients by commensal bacteria which are also consumed by pathogens, is a way of outcompeting the pathogenic bacteria. For instances, commensal *E. coli* compete for the same nutrients such as amino acids, organic acids, and other nutrients as pathogenic enterohemorrhagic *E. coli* (EHEC) [92],[93]. By consuming the same nutrients, well established commensal bacteria cause starvation of the competing pathogenic bacteria struggling to find an uncontested niche. Maltby *et al.* (2013), demonstrated that in a mouse model of EHEC infection, gut colonization by the pathogen was impossible when mice were colonized with 3 commensal *E. coli* strains selected for their ability to consume a panel of carbohydrates used by EHEC [94].

The important role of the gut microbiota in protecting from bacterial infections is also supported by the frequent emergence of enteric pathogens following an oral antibiotherapy. Disruption of the gut microbiota by antibiotics facilitates the expansion of some gut pathogens such as *Salmonella enterica* and *Clostridium difficile*, which have developed strategies of catabolizing microbiota-liberated mucosal carbohydrates like sialic acid. Indeed, some commensal species secrete sialidases that can cleave terminal sialic acid residues from mucin glycans [95]. When liberated, sialic acid can be used by any nearby bacteria including those that do not produce sialidases on their own. Katharine *et al.*(2013) detected an increase in free sialic acid in the cecum of mice treated with antibiotics for 1 day. Concomitantly, mice were also more susceptible to infection with *Salmonella* and *Clostridium difficile* [96]. These findings suggest that antibiotic treatment eliminates sialic acid consuming commensal bacteria. Thus the disappearance of the commensal bacteria leads to an open niche for pathogens to use this mucosa-derived carbohydrate. In addition of competition for carbon and energy sources, commensal bacteria also compete with pathogens for trace metals. For example, probiotic strain *E. coli* Nissle 1917 protects mice against *Salmonella* infection depending on iron uptake [97].

Production of antimicrobial compounds The gut microbiota can directly inhibit growth of bacterial pathogens and restrict access to the epithelium by producing AMPs such as bacteriocins [98]. Bacteriocins are considered natural variants of clinical antibiotics having an activity against Gram-positive and/or Gram-negative pathogens. Some of the best characterized

bacteriocins are those produced by members of *Lactobacillales* order, for example Nisin produced by *Lactococcus lactis* that inhibits a broad range of Gram-positive pathogens [99]. Members of the *Enterobacteriaceae* family such as commensal *E. coli* Nissle 1917 can produce microcins, a class of short peptide bacteriocins that act against close related strains [100] [101]. Under inflamed conditions, this strain can compete against pathogenic *E. coli* and *Salmonella* in a manner dependant on microcin expansion [102]. A healthy gut microbiota is key to a correct production of AMPs. The gut species *B. thetaiotaomicron* and *Lactobacillus innocua* are the main species that drive this production [103],[104]. Antunes *et al.* (2014) demonstrated that the expression of type III secretion gene in *Salmonella*, that is essential for colonization, is reduced when the bacteria was grown in presence of organic compounds isolated from human faeces [105]. Furthermore, the key role of commensal bacteria on the production of molecules that can limit the expansion of enterohemorrhagic *E. coli* was highlighted in our lab [106],[107]. Contrary to the large intestine where the mucus play an important role in the physical protection against pathogens, antimicrobial compounds play a larger role in the small intestine where the mucus is discontinuous.

Induction of host immune response On the other hand, commensal bacteria can also indirectly participate to the production of AMPs by the host. The commensal gut microbiota can stimulate the innate and adaptive immune system of the host in order to alter its susceptibility to exogenous pathogens [108],[109]. Flagellin or lipopolysaccharide (LPS) of bacteria can induce the release of interleukine (IL)-22 by innate immune cells which in turn stimulate AMPs release from epithelial cells [110],[111]. This mechanism was explored by Willing *et al.* (2011). *In vivo* infection with *Citrobacter rodentium*, a natural mouse pathogen, in NIH Swiss mice causes colitis that resolves within 3 to 4 weeks. On the contrary, for some mice strains such as C3H/HeJ, this infection is fatal. One of the main differences between these mice strains is the commensal microbiota. Indeed, microbiota transplantation of C3H/HeJ to NIH Swiss mice significantly decreased the mortality of transplanted NIH Swiss mice following infection with *C. rodentium*. The authors analyzed the effect of microbiota transplantation on expression of cytokine IL-22 which is known to be stimulated by the commensal microbiota. Indeed, levels of IL-22 were higher in the gut of resistant NIH Swiss mice [112]. This study illustrates the induction of immune responses mediated by IL-22 being one of the mechanisms by which the gut microbiota can protect against diarrheal infection.

CHAPTER 2

Gut microbiota and diseases

Gut microbiota contributes to the health and the well-being of the individual. Any disruption in the composition of the gut microbiota, called dysbiosis, can have negative impacts on the health of the individual. Dysbiosis can be defined by typical features such as a decrease in microbial diversity and/or richness, an altered ratio Bacteroidetes / Firmicutes, a decrease in beneficial species (i.e. *Bifidobacterium*, *Faecalibacterium prauznitsii*...) and/or an increase of deleterious bacteria such as Proteobacteria. Indeed, Proteobacteria usually represent 1-5% of the healthy gut microbiota but this phylum can reach up to 35% in the gut of patients suffering from various diseases, especially inflammatory ones. Currently, dysbiosis has been linked with several disorders and this list continues to grow. Indeed, several studies aiming to study the composition of the gut microbiota of patients suffering from different diseases showed that this composition differs comparing to the healthy individual. Today, link between intestinal dysbiosis and diseases such as obesity, allergic disorders, diabetes, autism, Alzheimer, Parkinson, colorectal cancer and cardiovascular diseases has been shown.

Here, the relationship between gut microbiota and two examples of inflammatory diseases are described : Crohn's disease as an intestinal inflammatory disease and rheumatoid arthritis as an extraintestinal inflammatory disease.

1. Intestinal inflammatory diseases : example of Crohn's disease

1.1 *Etiology and pathophysiology of Crohn's disease*

IBD is a chronic and relapsing inflammatory disorder of the intestine with two subtypes : Crohn's disease (CD) and ulcerative colitis (UC). Common symptoms include diarrhea, abdominal pain, rectal bleeding, fever, weight loss, and fatigue [113]. IBD affects around 3.1 millions people in the USA [114]. The incidence of this disease has increased in the Western world and keeps rapidly increasing during the last 50 years [115],[116]. CD is a chronic, segmental, inflammatory disease that can affect all parts along the GIT and can cause lesions from mouth to anus accompanied in some cases with extraintestinal complications. The exact etiology is still unclear and probably results from the combination of multiples factors such as immunological disorders, altered gut microbiota, genetic factors and environmental factors. Up to date, treatments of CD focus mainly on mucosal healing and reduction of the symptoms.

Many medical options are available but still not resolutive, among them anti-inflammatory drugs, antibiotics, immunosuppressants, and targeted medical options directed towards tumor-necrosis (TNF) factor α , integrins (Vedolimumab) and interleukins IL-12 and IL-23. In some cases, surgery has a key role in treating complications such as stenosis, perforations, fistulas and abscesses. The medical management should consider several factors such as disease severity, subtype, behavior and location [117].

Host factors Genetic factors play an important role on susceptibility to CD and influence the composition and diversity of the gut microbiota. The first ever IBD susceptibility gene identified was NOD2 that codes for a protein belonging to the NOD-like receptors. Several studies involving NOD2-deficient mice showed an increase of the relative abundance of Bacteroidetes and the *Enterobacteriaceae* family in the gut [118],[119]. An example of a gene specifically linked to CD is *Nod2/CARD15* in the chromosome 16 [120]. In case of mutation of *Nod2/CARD15*, there is an important reaction of TLR2 (that is normally regulated by NOD2) to commensal antigens initiating the anti-inflammatory reaction in the gut. Besides, this mutation also induces a decrease of the secretion of antimicrobial peptides such as alpha defensins by Paneth cells. This mechanism was shown *in vitro* to be linked to the difficulty in eliminating *E. coli* [121]. Genetic factors were estimated to contribute to 20% of the risk of developing CD. Genome wide association studies (GWAS) have identified 30 different specific loci associated with CD [122],[123]. It has been demonstrated that genes involved in epithelial barrier function were more associated with UC whilst genes involved in cellular innate immunity were more associated with CD. The gene ATG16L1 was identified by GWAS as a risk gene for both UC and CD. This gene regulates the development of T cells and cell autophagy and the deletion of ATG16L1 in mouse models triggers spontaneous inflammation in the GIT [124]. Another host related factor associated with the development of CD is a default in production of mucus by Goblet cells. In patients suffering from CD, the gene *MUC1* has a decreased expression indicating that mucin secretion is insufficient, affecting the role of the intestinal mucosa as a barrier against aggressions [125]. Moreover, GWAS also showed that genes *MUC1*, *MUC19*, and *PTGER4* were associated with CD supporting this hypothesis [122]. The intestinal barrier is described as “leaky” for CD patients since tight-junction proteins expression is altered. This phenomenon gives easy access through the lamina propria of luminal content and antigens [126].

Environmental factors Regarding the environmental factors that influence the development of CD, there is a specific geographic distribution of this disease. High incidence areas have been identified such as northern Europe and North America [127]. Even in ethnic groups such as Asians and Hispanics that were less affected by CD, the incidence is raising due to the environmental factors [128]. Among these factors, the main one is industrialization that has a great impact on the lifestyle of an individual. Thanks to industrialization, people have a more sedentary lifestyle [129], are exposed to air pollution [130], consumption of a western diet [131], and smoking. Smoking is the most studied environmental factor impacting the clinical course of CD. Patients who smoke are more likely to develop this disease, increasing the risk of requiring immunosuppressive agents [132],[133],[134]. Undoubtedly, the western diet loaded with excessive amounts of sugars and polyunsaturated fats contribute to the dysbiosis observed in patients suffering from CD. This diet was shown to induce intestinal dysbiosis with reduced microbial diversity resulting in poor production of SCFAs that are known to have effects on maintaining the gut homeostasis. In addition, the Western diet incites the expansion and activity of colonic mucus-degrading bacteria resulting in barrier dysfunction [135]. These observations indicate that this type of diet increases susceptibility to develop CD. The environmental factors described above in genetically susceptible people will trigger dysbiosis, impacting the intestinal barrier and causing inflammation in the GIT.

1.2 Dysbiosis in Crohn's disease

Dysbiosis has been reported in IBD over the last 20 years [136]. The link between IBD and the gut microbiota dysbiosis was first demonstrated in GF IL 10^{-/-} that do not develop colitis unless they are colonized by enteric bacteria [137]. Even though CD and UC resemble in epidemiologic, immunological, and clinical features, the gut microbial community detected in patients is distinct. There are dramatic changes that have been demonstrated in CD patients including changes in the population levels, reduced species richness as well as modification of gut microbiota associated metabolites [138]. Using single strand confirmation polymorphism fingerprint, the diversity of the gut microbiota was shown to be decreased by 50% in CD [136]. Dysbiosis in CD patients is usually characterized by an increase in Proteobacteria while other commensal populations are reduced notably members of the phyla Firmicutes and Bacteroidetes [139],[140]. Bacteroidetes and Firmicutes represent only 20% in CD instead of 90% in healthy individuals. In addition, the Proteobacteria phylum raises up to 10% whereas it represents only ~1% in control cases (**Figure 7**) [141]. A reduction of the species *Faecalibacterium prausnitzii*

was demonstrated to be linked to increased risk of developing CD. It is now admitted that the microbiota associated with mucus is different from the microbiota of luminal contents in terms of abundance, diversity and metabolites [142]. In the case of CD, a reduction in Firmicutes and an increase in Proteobacteria such as *Escherichia coli*, and in *Veillonella*, *haemophilus*, and *Fusobacteria* were observed in ileal and rectal mucosal samples [143]. Another study showed that mucosa associated microbiota includes more *E. coli* and less *F. prausnitzii* in CD patients than in healthy individuals [144]. Patients have a greater number of mucosal surface-associated bacteria with the capacity of adhering and invading the epithelial cells compared to healthy individuals [145]. Therefore, mucosa-associated microbiota changes in CD are more marked than changes in the stool. A better access to the microbial community in the different intestinal ecological niches seems essential to better understand their ecological roles in CD patients.

Moreover, dysbiosis observed in CD patients is different according to the genetic mutations associated with the risk of developing the disease. Different changes are detected depending on the type of mutation carried by the patient. For instances, ATG16L1 homozygous² CD patients have an enrichment of *Fusobacteriaceae*, *Bacteroidaceae* and *Enterobacteriaceae* in the ileum [146].

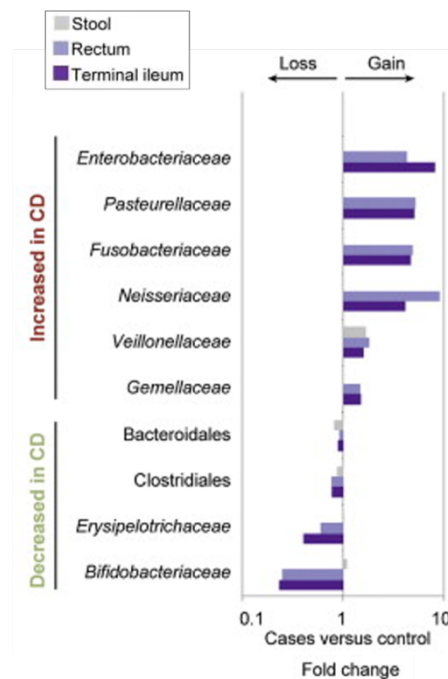


Figure 7 Differences in microbial abundance of bacterial families among CD patients. The fold change for each taxon was calculated by dividing the mean abundance in the patient cases by that of the controls. Several taxonomic biomarkers measured at both the ileal and rectal sites were found to be significantly correlated with disease phenotype. However, most of that microbial signal was lost in the stool samples. The taxa at the top are increased in disease state, whereas the taxa at the bottom follow an opposite trend. Adapted from Gevers et al., 2014.

² Homozygous : having two identical alleles of a particular gene or genes.

CD patients lacking FUT2 alleles (coding the enzyme Fucosyltransferase 2 that regulates intestinal cells-microbes interactions) have an abnormal mucosal barrier with decreased microbial diversity [146],[147]. Also, an abundance of *Anaerostipes*, *Coprococcus*, and *Lachnospiraceae* was found in CD patients which can be linked to the risk locus SLC39A8 a gene located on chromosome 4q24 that encodes for a manganese transporter [148].

1.3 *E. coli* associated with Crohn's disease

In CD patients, several studies have proven an increase in Proteobacteria, mainly *E. coli* associated to the intestinal mucosa [149]. Most of these *E. coli* strains can adhere and invade mucosal cells and were denominated AIEC for Adherent and Invasive *E. coli* [150]. The prevalence of AIEC within the gut mucosa of ileal CD patients is of 38% comparing to only 6% in healthy individuals [151]. It has been shown that these *E. coli* strains were abnormally prevalent in early and chronic ileal lesions indicating that they may have a potential role on the early pathology of CD [152]. Criteria that characterize AIEC are : the ability to adhere to and to invade differentiated intestinal epithelial cells and the ability to survive within macrophages. Most AIEC strains do not harbor any of the known virulence factors expressed by other *E.coli* pathotypes such as enteropathogenic *E. coli* (EPEC) or EHEC [152]. The absence of conventional pathogenicity genes suggests that AIEC should be more considered as a pathobiont rather than a pathogen *sensu stricto*. Moreover, the ability of AIEC to gain intracellular access relies on defects in mucosal immunity or barrier dysfunction as observed in CD patients [153],[154].

Reference *E. coli* AIEC strain LF82 is able to adhere and to invade intestinal epithelial cells via mechanisms involving actin polymerization. *In vitro* and *in vivo* studies described the different mechanisms used by AIEC to reach, adhere and invade intestinal epithelial cells. AIEC penetrates the mucus layer by promoting mucin degradation by a protease referred as VAT-AIEC. In a murine model, it has been reported that this protease contribute to gut colonization by enhancing the spread of bacteria and adhesion to epithelial cells [155]. The binding to the host cells depend on type 1 pili-adhesion. The role of this structure to the adhesion capacity of LF82 was confirmed by studies using non-invasive mutants of the *fim* operon encoding type 1 pili [156]. Polymorphisms in the *fimH* type 1 pilus subunit sequence have been reported in AIEC strains conferring higher adhesion ability [157],[158]. When interacting with the epithelial barrier, AIEC can form biofilms thanks to presence of the H1 flagellar antigen [159]. The next step is internalization into the intestinal mucosa. Indeed, AIEC are capable of invading

not only epithelial cells but also macrophages. The AIEC outer membrane vesicles interact with the endoplasmic reticulum stress response factor Gp96 and, in association with flagellin that target TLR5, elicit an epithelial IL-8 inflammatory response and promote internalization within intestinal epithelial cells [160]. Moreover, AIEC can invade and replicate within macrophages without inducing host-cell death [161]. Macrophages are phagocytic cells with a critical role in the intestine's innate immune defense [162]. Macrophages can recognize molecular patterns synthesized by pathogenic bacteria such as fimbriae, flagella, LPS and peptidoglycan. There are key genes such as *gipA* that are responsible for AIEC survival and replication within macrophages. Expression of this gene is induced in specific intracellular, gastrointestinal and phagolysosomal conditions. Deletion of this gene reduces the proinflammatory response induced by AIEC and also AIEC's ability to replicate within macrophages [163].

Other works demonstrated that AIEC can induce inflammatory responses by regulating the expression of inflammatory genes [164]. *In vitro*, the strain LF82 can induce the expression TNF- α and IL-8 transcripts on colon biopsies of patients with CD and affect the cell cycle progression on Caco-2 cell lines [165]. Altogether, these data suggests that the high level of AIEC population in the gut mucosa of CD patients might participate to the progression of disease severity in CD patients, notably through exacerbation of the inflammatory response [166]. However, research into the role of AIEC on promoting CD has been limited by the lack of suitable animal models of long-term AIEC colonization without the need for co-administration of colitogenic chemicals to transgenic mice [167].

2. Extraintestinal diseases : example of Rheumatoid Arthritis

2.1 Etiology and pathophysiology of Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory disease characterized by a constant immune response located at the joints leading to inflammation of synovial tissue and subsequent cartilage and bone erosion [168]. RA is linked to the production of self-antigens (i.e. anti-rheumatoid factor or anti-citrullinated protein antibodies) in joints causing the inability to recognize the hosts antigens. This process leads to overproduction of cytokines by the hosts immune system which drives persistent inflammation and joint and bone destruction. This disease affects mainly the joints of the hands, wrists, and other synovial joint but has a negative impact in the body as a whole. Symptoms can include morning stiffness of the affected joints, fatigue, fever, weight loss, joints that are tender and swollen forming rheumatoid nodules under the skin [169]. RA affects 0.5-1.1% of the worldwide population with a higher prevalence in

developed countries like Northern Europe and North America [170]. Hence, incidence and prevalence of RA vary across populations and is highly dependent of the choice of criteria for cohort studies.

RA is a multifactorial disease that has both autoimmune and inflammatory components and is subjected to diverse genetic, environmental, and epigenetic influences (**Figure 8**) [170],[171]. RA etiology includes genetic factors that play a significant role on development of the disease. However, the exposure to various environmental factors, even before the appearance of clinical signs of RA, impacts the severity and the clinical course of the disease. These factors have intricate interactions with the immune system of the individual entailing several inflammatory cascades. One of the most studied factors includes the overproduction and overexpression of TNF [172]. The TNF pathway drives both inflammation and destruction of the joints. Overreaction of T and B lymphocytes, synovial-like fibroblasts and macrophages against synovial proteins causes overproduction of TNF and many cytokines such as interleukin 6 leading to permanent inflammation and joint destruction [173].

Etiology of Rheumatoid Arthritis

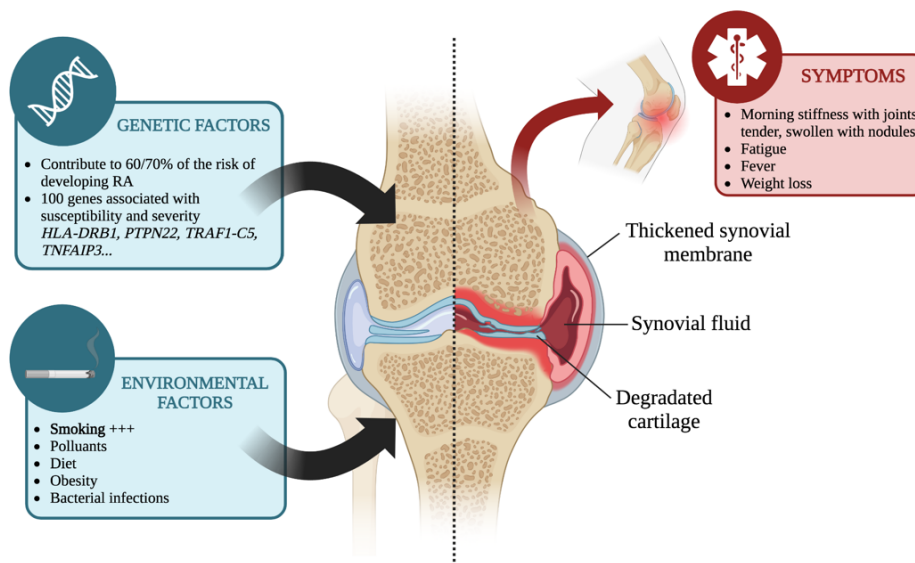


Figure 8 Different factors participate to the etiology of RA. Among host factors, genetic factors play an important role contributing to 60-70% of the risk of developing RA. One of the most studied genes associated with high risk and severity of this disease is HLA-DRB1 responsible for triggering T cells of the immune system. Several environmental factors participate to the early stages of disease, one of the most studied being cigarette smoke. RA is an autoimmune disease characterized by symptoms such as morning stiffness, joints tender, swollen and with nodules, fatigue, fever and weight loss. *Figure made with biorender.com.*

Diagnosis of RA is made based on clinical features according to established classification criteria [174]. The major role of treatment in RA is remission with no active joint inflammation and no erosive or functional deterioration. Depending on the intensity and type of treatment, 10-50% of patients with early RA achieve remission. The treatments of RA target mainly the major symptoms with analgesics and non-steroidal anti-inflammatory drugs to lessen pain and stiffness [175]. Disease-modifying antirheumatic drugs are conventionally used, and reduce joint swelling and pain, limit progressive joint damage, and improve function [176]. These drugs could however have some side - effects on the patient thought (nausea, hepatotoxicity, blood dyscrasias, and interstitial lung disease) [177]. Furthermore, biological agents (TNF inhibitors) can be used on non-responsive patients and supportive treatment (exercise, joint protection, food care, and psychological support) is crucial for the well-being of the patient [178],[179]. Prevention of RA seems possible by avoiding smoking, and by modification of the diet. However, at present these recommendations are not supported by strong evidence.

Host factors Genetic factors highly participate to RA etiology. These genetic factors are estimated to contribute to 60 to 70% of the risk of developing RA and several loci have already been described. About 100 genes can be associated with susceptibility, protection, severity, activity, and treatment response such as tyrosinephosphatase gene PTPN22, polymorphisms³ of TRAF1-C5, TNFAIP3 and others [180]. Among the known genetic factors, the presence of certain allele of Human leukocyte antigen (HLA) class II molecules is the strongest association with RA. One of the most well described in the literature is the gene *HLA-DRB1* that is likely to contribute to at least 30% of the total genetic component of this disease [181]. Associations of RA with HLA-DRB1 alleles have been observed in all racial and ethnic populations being the only genetic factor in common between several populations [182],[183]. HLA is responsible for triggering T cells of the immune system, to produce antibodies helping the immune system, to discriminate between host proteins and invaders proteins. The alleles associated to RA have a conserved motif of 5-amino acid sequence in positions 70–74 of the HLA–DRβ chain termed shared epitope⁴ (SE) [184]. There are SEs that confer a higher risk than other and influence on disease severity such as the presence of two SE alleles and in particular HLA-DRB1*0401/*0404 [185]. The SE alleles are linked to an increased susceptibility to RA in particular to anti-citrullinated protein antibody (ACPA). Citrullination is a posttranslational

³ Polymorphisms : the presence of genetic variation within a population.

⁴ Epitope : part of an antigen that is recognized by the immune system.

modification most frequently observed in RA that consists on the conversion of arginine onto citrulline by the enzyme peptidyl arginine deaminase. This modification is linked to self-antigen recognition in RA [186]. RA patients that carry these alleles have enhanced immune response to citrullinated proteins [187]. Indeed, citrulline may alter protein structure and generate new epitopes associated with the production of ACPAs [188].

Finally, higher prevalence of RA in females suggests a role for hormonal factors. The female-to-male ratio of this disease is 4:1 in patients younger than 50 years old [189],[190]. Several studies indicate that the risk of developing RA is not linked to the chromosome X even though the role of hormonal factors such as estrogen carries many controversies [191] The role of estrogens on developing RA is not clear because this hormone has been described having both a pro-inflammatory effect and inhibitory effect of the immune system [192]. Even though, several studies described the decline of estrogens (menopause, use of anti-estrogen agents) as a risk factor for RA disease.

In a large cohort study, it has been demonstrated that menopause has been associated with an increased risk of developing RA particularly for those who experienced it early before 44 years old [193],[194]. Furthermore, the use of anti-estrogen agents to treat breast cancer was shown to be associated with development of autoimmune disorders such as RA [195].

Environmental factors Environmental factors are also implicated in the development of RA, acting before the disease becomes clinically apparent. They include smoking, pollutants, diet, obesity, bacterial infections, and others. Smoking is the dominant environmental risk factor contributing to development of RA disease. Cross-sectional and longitudinal studies have shown that smoking increases susceptibility to RA and affects the clinical course of RA [196]. In fact, smoking multiplies the effect of the HLA-DRB1 SE alleles and is associated with the production of auto-antibodies such as ACPA and rheumatoid factor (RF) [197]. The diet is another environmental factor that influences the development of the disease emphasizing the role of the gut microbiota. Specific dietary choices increase disease risk, especially diets with high pro-inflammatory effects [198],[199]. Indeed, RA is higher in Western countries where the diet is characterized by high intake of red meat, saturated and trans fats, a low ratio of omega-3:omega-6 fatty acids and high consumption of refined carbohydrates [200],[201]. A high salt regime common in Western countries was also suggested as a risk to develop RA. High levels of salt in diet, in particular for smokers, will induce serum glucocorticoid kinase-1 expression, with an increased Th17 lymphocyte differentiation and enhanced autoimmunity [202].

2.2 *Dysbiosis in Rheumatoid Arthritis*

Several studies have reported differences of the oral [203], lung [204], and gut [205] microbiota between RA patients and healthy individuals.

Oral microbiota The potential role of the buccal microbiota on developing RA has been suspected for years. Periodontal diseases⁵ are more common and severe in RA patients than in healthy individuals. Moreover, DNA sequences of oral pathobionts were found on the synovial fluid of RA patients. Patients have also high quantities of antibodies in the synovial fluid directed against the main oral pathobionts *Porphyromonas gingivalis*, *Prevotella intermedia* and *Tanarella forsythia* [206]. The potential role of *P. gingivalis* as a possible triggering factor of RA was particularly investigated. Indeed, this bacterium has the capacity of producing peptidylarginine deaminase, an enzyme responsible for generating citrullinated proteins which are major targets of autoantibodies in RA [207]. Antibodies directed against these proteins are found in 80% of RA patients which indicates that it might be an important participant in auto-immunity mechanisms [208].

Lung microbiota. Smoking is a well-established risk factor of developing RA. Furthermore, pulmonary disease is often a common extra-articular manifestation and includes pleurisy, parenchymal nodules, interstitial involvement, and airway disease [209]. The lungs have their importance on developing RA and theories suggest that the lungs could be a site of autoantibody generation [204]. *Prevotella* and *P. gingivalis* have also been identified within the lung microbiome [210],[211]. Similar to *Streptococcus pneumoniae* and *Streptococcus pyogenes*, *P. gingivalis* is capable of causing respiratory infections due to the expression of α -enolase [211]. Indeed, infection by *P. gingivalis* is an additional factor that can promote the anti-citrullinated protein autoimmunity in RA. This bacterium produces its own citrullinated enolase that resembles in terms of amino acid sequence the human enolase. Therefore, an original immune response to *P. gingivalis* may redirect the immune response to human citrullinated proteins (as a consequence of molecular mimicry) [212].

Gut microbiota Despite the fact that RA is an extra-intestinal disease, several studies showed that patients display significant differences of the intestinal microbiota comparing to healthy individuals. These alterations of the gut microbiota composition are related to risk and severity

⁵ Periodontal diseases : infections and inflammation of the gums and bone that surround and support the teeth.

of the disease. Overall studies reported that the gut microbiota of RA patients is characterized by a decrease of *Bifidobacteria* [213],[214],[215], an increase of *Lactobacillus* [216] a bloom of Proteobacteria and a high abundance of *Prevotella copri* [217]. The imbalance of the gut microbiota induces important shifts on the bacterial population present in the gut and also the metabolic activities associated with. For instances Chen *et al.* (2016) tried to search for specific biomarkers of RA by sequencing 16S ribosomal DNA of fecal samples from RA patients and healthy individuals. This study showed a shift in a rare lineage of intestinal microbes such an increase of *Collinsella aerofaciens* and *Eggerthella lenta*. The increase of these rare lineages could trigger inflammatory conditions along with elevated levels of three metabolites : β -alanine, α -aminoadipic acid, and asparagine [213]. This study suggests that it could be important to regulate the imbalance observed in the gut microbiota of RA patients to improve inflammation and health.

However it appears that the bacterial groups detected on the different studies are variable and often depend on the cohort of patients used in the study. Different parameters such as disease state, age and treatment influence the composition on the bacterial species during dysbiosis and must be considered when trying to define a pattern of its composition (**Table 1**).

Studygroups	Sample Type	Technology Employed	Bacterial Taxa (↓low, ↑enriched)	Ref.
Early RA (51) vs. Fibromyalgia (50)	Stool	16S rRNA hybridization, and DNA-staining	↓ <i>Bifidobacteria</i> , ↓ <i>Bacteroides-Porphyromonas-Prevotella</i> , ↓ <i>Bacteroides fragile</i> , ↓ <i>Clostridium coccooides</i>	[24]
Early RA (15) vs. Healthy (15)	Stool	Quantitative real-time PCR	↑ <i>Lactobacillus</i>	[25]
New-Onset RA (44) vs. Healthy (28)	Stool	16S rRNA gene and WGS sequencing	↑ <i>Prevotella copri</i> , ↓ <i>Bacteroidetes</i>	[26]
RA (30) vs. Healthy (30)	Stool	16S rRNA gene and WGS sequencing	↑ <i>Enterococci</i> , ↑ <i>Clostridia</i> , ↑ <i>Colibacteria</i> , ↓ <i>Lactobacteria</i>	[27]
Treatment-naïve RA (94) vs. Healthy (97)	Stool, Dental, Saliva	Metagenomic shotgun sequencing	↑ <i>Lactobacillus salivarius</i> , ↑ <i>Gordonibacter pamelaeeae</i> , ↑ <i>Clostridium asparagiforme</i> , ... ↓ <i>Haemophilus</i> spp.	[28]

Table 1 Alterations of gut microbiota related with RA. Adapted from Wu *et al.*, 2016.

The link between the imbalance of the intestinal microbiota and RA can be due to different mechanisms that impact the immune system and its function. RA-associated antibodies might be detected in the blood long time before joint inflammation, supporting the hypothesis that the process leading to autoimmunity in the joints could originate from extra-articular mucosal sites such as the intestinal mucosa [218]. Several studies have explored the interaction of the lung

mucosa with the immune system of the host and the role of cigarette smoke. By contrast, few studies have concerned the potential role of the gut microbiota. Similarly to the lung mucosa, the intestinal mucosa contains immunological cells able to trigger an immune response activating T-cell mediated immunity against citrullinated autoantigens. The intestinal epithelium is a major producer of citrulline in the human body, this could explain why autoantibodies such as RF and ACPAs are detected well before the onset of RA in the absence of synovitis [219]. There are many possible roles for citrullinated microbial antigens and molecular mimicry, TLR signals, and other innate immune stimuli [220]. The bacterial community of the intestinal mucosa interacts with immune cells, resulting in inflammation and increased autoantigen production. The produced citrullinated antigens are presented by the ACPAs to activate the T and B cells, leading to the production of autoantibodies [218].

One of the most studied gut dysbiosis pattern in RA is the high abundance of *Prevotella copri* in fecal samples of RA patients [217]. *P. copri* is an obligate anaerobe, non-spore forming Gram-negative bacterium that was first isolated from human fecal samples in Japan [221]. Very recently, preclinical phase RA patients from European countries were shown to harbor a high abundance of *Prevotella* species, including *P. copri*, in the intestine, suggesting that dysbiosis precedes the development of arthritis [222]. Conversely, the relative abundance of *P. copri* in the gut was reduced in patients with chronic RA suggesting that this bacterium could have a role in the early stages of the disease. When *Prevotellaceae* enriched microbiome of RA patients is transferred to SKG mice (a model that spontaneously develop chronic autoimmune arthritis), it increases sensitivity to arthritis thanks to the activation of autoreactive T cells in the intestine [223]. Interestingly, the relative abundance of *P. copri* in the gut is associated with the absence of HLA-DRB1, meaning that the composition of the human gut microbiota could also be partially dependent on the host genome [223]. Overall, the immune responses of the host may be dependent on the presence of certain genera/species of the gut microbiota that are going to interact with the immune system and impact the integrality of the epithelial barrier. Mechanisms involving *P. copri* have been well studied because it seems this bacterium has an important role on the onset of RA. Moreover, oral administration of *P. copri* increased local inflammatory response in a colitis murine model meaning that this bacterium could alter intestinal permeability. Indeed one of the proposed mechanisms that link dysbiosis with the pathogenesis of RA is that increased colitis might lead to penetration of bacteria or its components through the intestinal barrier throughout the body [224]. Moreover *P. copri* has a pro-inflammatory potential impacting the inflammatory response of the host. It has been shown that it induces Th17 related cytokines such as IL-6 and IL-23 [225]. Altogether, the results

obtained support the hypothesis that *P. copri* thrives in a pro-inflammatory environment and may exacerbate inflammation for its own benefit [226][227].

In recent years, advances have been made to better understand the pathogenesis of RA and the link with the composition of the gut microbiota by the use of different animal models (Table 2). These studies need to be interpreted with caution though because there is no animal model that represents entirely RA developed in human. In the collagen-induced arthritis mouse model, which is the most commonly used animal model to study RA, the gut microbiota of collagen-treated mice presented a decreased microbial richness when compared to untreated mice.

Mice strain	Environmental condition	Mechanism of involvement of arthritis	Intestinal bacteria correlated with induction of arthritis	Ref.
SKG	GF, SPF: no arthritis conventional: arthritis	Production of auto-reactive T cells Activation of innate immunity by fungi	<i>Prevotella</i> -dominated microbiota	20, 22, 23
IL-1ra ^{-/-}	GF: no arthritis conventional: arthritis	Activation of TLR2 and TLR4 Th17 cells ↑ Treg cells ↓	<i>Lactobacillus Bifidus</i> <i>Helicobacter</i>	21, 24
K/BxN	GF: no arthritis SPF: arthritis	Production of GPI-antibody Th17 cell expansion in the intestine	SFB	12
CIA	ABX: reduced severity of arthritis SPF: arthritis	Production of anti-type II collagen antibody and serum inflammatory cytokines	–	27

GF germ-free, SPF specific pathogen free, GPI glucose-6-phosphate isomerase, SFB segmented filamentous bacteria, TLR Toll-like receptor, Treg cells regulatory T cells, CIA collagen-induced arthritis, ABX antibiotics, Ref references

Table 2 Murine models of RA known to be correlated with the gut microbiota. Adapted from Wu et al., 2016..

CHAPTER 3

Proteobacteria in the gut microbiota

The chapter 3 is devoted to the phylum Proteobacteria. As previously mentioned, members of the phylum Proteobacteria are systematically present in the gut of healthy humans but at a low abundance (< 5% of total microbiota abundance) [228]. However, compelling evidence shows that an imbalanced gut microbiota described in the context of a disease is often associated with a sustained increase in abundance of the phylum Proteobacteria [229],[230],[231],[232]. Microbial molecular analyses have notably linked a bloom of *Enterobacteriaceae*, the main representative family of Proteobacteria, with many conditions involving inflammation such as CD and RA described in Chapter 2 [233]. Environmental and nutritional changes in the gut associated with host pathology or dysbiosis could confer growth advantage to *Enterobacteriaceae*, a family known to have a high capacity of adaptation to changes. Thus, a better understanding of the mechanisms behind *Enterobacteriaceae* adaptation in the dysbiotic gut could help designing potential intervention strategies to revert dysbiosis and improve health.

1. Proteobacteria in the healthy human gut

The phylum of Proteobacteria has a large phylogenetic composition with 116 validated bacterial families. Members of this phylum have greatly variable morphology and versatile physiology which give a competitive advantage in surviving in various ecological niches. Proteobacteria can belong to six different bacterial classes: Alpha-, Beta-, Gamma-, Delta-, Epsilon-, and Zetaproteobacteria (**Figure 9**). Classes division is based on molecular relatedness, so there is no particular morphological or physiological trait that characterizes members of each class. In the literature, the family *Enterobacteriaceae* (Gamma-proteobacteria class) is considered the most representative of Proteobacteria in the GIT. *Enterobacteriaceae* is the most taxonomically diverse bacterial family recognized by the International Committee on Systematic Bacteriology [234],[235]. *Enterobacteriaceae* is constituted of gram-negative bacteria and mostly facultative anaerobes and *E. coli* is the most renowned representative of this family [236]. Moreover, *Escherichia* is highly prevalent in the human gut and is the only genus belonging to Proteobacteria that has been recovered among the 30 most abundant genera identified from 22 fecal samples of healthy individuals [237]. Besides *Enterobacteriaceae* and *E. coli*, information of the taxonomic composition of Proteobacteria members in healthy humans is often not available in the scientific literature, probably because Proteobacteria is a

low abundant phylum with no critical attributed function in the gut ecosystem. Nonetheless, the supplementary data of a study looking at the temporal variation in the human microbiome has given the bacterial abundance at different taxonomic levels from gut samples of 2 healthy individuals [238]. Similar to other studies, Proteobacteria account for 2,5 and 4,1% of total bacteria in the 2 samples. For each sample, these Proteobacteria are represented for 50 % by the family *Enterobacteriaceae* (mostly *Escherichia*) and for 25% each by families *Alcaligenaceae* (Beta-proteobacteria class) and *Desulfovibrionaceae* (Delta-proteobacteria class).

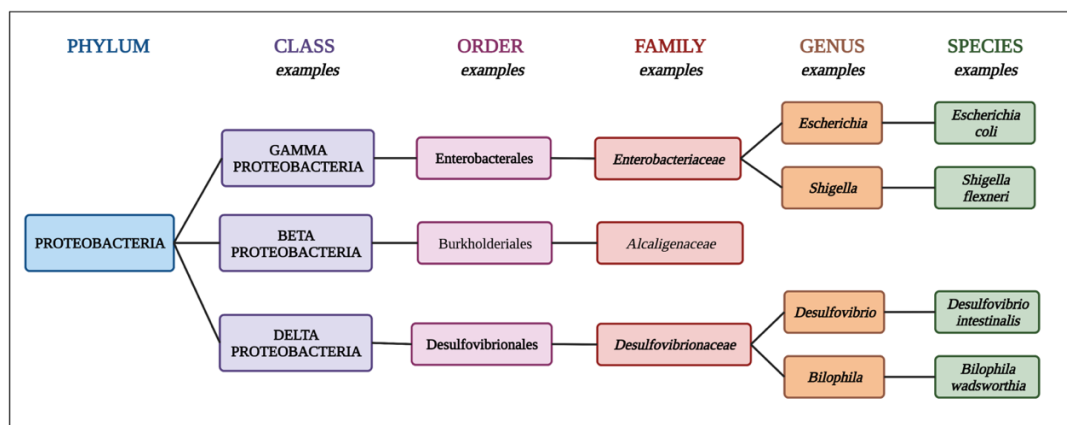


Figure 9 Organization of the Proteobacteria phylum with several examples of Class, Order, Family, Genus and Species. Figure made with biorender.com.

1.1 Focus on *E. coli* in the healthy human gut

Because *E. coli* is both the most abundant member of Proteobacteria in the gut as well as a reference model organism to study and understand biology, the majority of studies related to this phylum in the gut microbiota is dedicated to *E. coli*. Historically, Theodor Escherich, a German microbiologist and paediatrician studying the infant gut microbes and their role in digestion and disease, isolated in 1885 a fast-growing bacterium that was later named *E. coli* [239]. *E. coli* is best known as a ubiquitous Gram-negative, non-sporulating facultative anaerobe [240]. In addition to its use as a model organism, this bacterium has been mostly studied in a context of pathogenicity because it kills around two million of humans per year [241]. Nevertheless, *E. coli* is also and above all a widespread gut commensal of vertebrates [242].

Commensal *E.coli* strains can be mostly found in the large intestine, especially in the caecum and the colon. The main ecological niche of *E. coli* in the gut is the layer of mucus covering the epithelial cells of the colon. In the layer of mucus, *E. coli* grows in a complex, multi-species biofilm in which it competes for nutrients originating from the diet or from the mucus itself [243],[244]. Strains that are isolated from this part of the intestine grow on nutrients acquired from mucus, including at least seven mucus-derived sugars. Although the concentrations of these sugars in the intestine are low, *E. coli* maximizes its growth by using micro-aerobic and anaerobic respiration [245], maintaining population densities of 10^6 – 10^9 cells per gram of fecal matter [241].

A) Genetic and phylogenetic diversity of *E. coli*

A century later the discovery of *E. coli*, a strain collection called ECOR (*E. coli* reference) was established by Ochman and Selander based on multi-locus enzyme electrophoresis results. This collection comprises 72 isolates from healthy humans and 16 from other mammalian hosts and represents the genetic diversity of this species. This collection rapidly became a common substrate for various phenotypic, genetic and phylogenetic analyses all over the world. Regarding the evolution of *E. coli*, the population structure of this species is predominantly clonal, despite the occurrence of recombination events, and it allows the delineation of phylogenetic groups. The historical ECOR collection has been divided into six major phylogenetic lineages A, B1, B2, D, E and F [246] (**Figure 10**). Between these groups of strains, there are several differences in terms of phenotypic characteristics, including their ability to metabolize certain sugars, antibiotic resistance profiles and the distribution of virulence factors for pathogenic strains [247]. With the accumulation of whole genome sequence data over the years, phylogenetic analysis identified two new phylogroups, G and H, providing a new understanding of the evolutionary origin of *E. coli* [248].

The molecular basis underlying the phenotypic diversity of *E. coli* is based on a very dynamic genome structure leading to genome plasticity. The *E. coli* genome can be classified into two categories. Firstly, the core genome is defined as the shared genes present in all strains including essential housekeeping genes involved in replication, transcription and translation and it corresponds to the “backbone” that defines the strain. And secondly the flexible genome with genes called singletons that are only present in some strains. These are the genes that are responsible for the diverse phenotypes observed and the capability of adaptation to specific environment modifications [249].

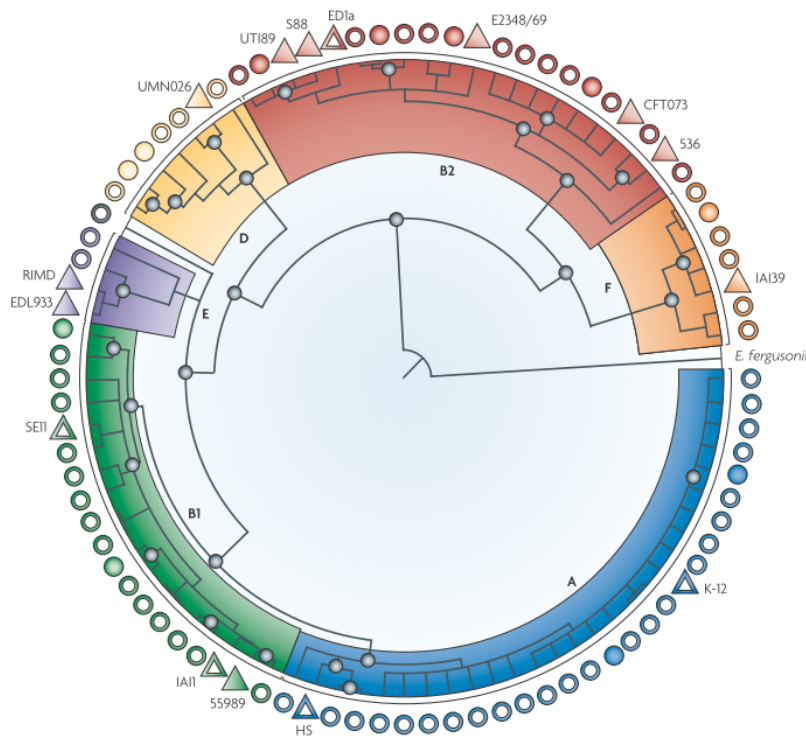


Figure 10 Phylogenetic history of *Escherichia coli*. Analysis based on the sequences of 8 housekeeping genes (4,095 nucleotides in total), 72 strains from the *E. coli* reference collection (ECOR) (outer circles) and 15 genome reference strains (outer triangles), rooted on *Escherichia fergusonii*. The commensal strains are indicated by an open symbol, whereas the pathogenic strains are represented by a full symbol. Colors indicate the 6 main phylogenetic groups. Blue dots on nodes indicate that the clade defined by the node is monophyletic, with more than 80% support. From Tenailon et al., 2010.

In a study from 2009, authors analysed the genomes of 20 different *E. coli* with an average gene number of 4,721 per genome and showed that the pan-genome (set of genes found in at least one strain of *E. coli*) reach more than 18,000 genes with only 2,000 genes being part of the core genome [250]. It is important to note that, the more *E. coli* genomes are sequenced and compared, the more the pan-genome continues to increase with the discovery of new singletons within species and the core genome reduces itself (Figure 11) [251].

The flexible genome can be acquired via horizontal gene transfer and includes prophages, transposable elements, and accessory genes [251],[252]. For instances, a study from 2002 showed that 24,5% of the genome of *E. coli* K-12 (the lab reference strain) was acquired from horizontal gene transfer [253]. These genetic features can help improve fitness in particular niches, increase metabolic flexibility and affect pathogenicity [254],[255].

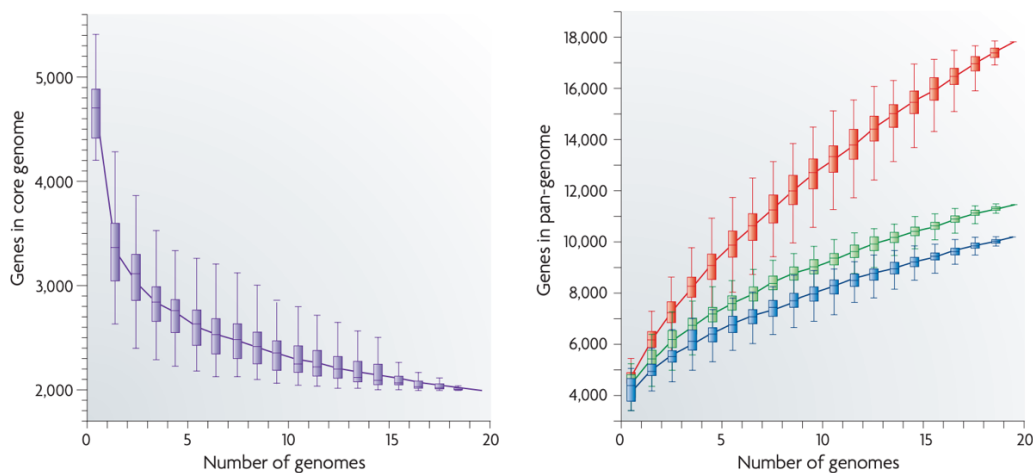


Figure 11 *E. coli* pan-genome and core genome evolution according to the number of sequenced genomes. Number of genes in common (left) and total number of non-orthologous genes (right) for a given number of genomes analysed for the different strains of *E. coli*. From Tenaillon et al., 2010.

B) *E. coli* residency in the GIT

Whereas *E. coli* is the best characterized organism on earth, little was known about the factors influencing the abundance and distribution of representatives of the species in the gut of mammals until several research groups began to ask important ecological questions about *E. coli*. In cross-sectional studies of human adults, *E. coli* is a member of the intestinal microbiome of over 90% of individuals [241]. The number of distinct *E. coli* strains per host was also assessed in individuals and the average number was found to be between 1.1 and 3.5 *E. coli* strains per individual [256],[257]. Several studies in the literature also investigated the prevalence of *E. coli* phylogenetic groups in the gut of the healthy individual. Results are variable between studies and appear to depend on different factors such as geographic location, diet, gut morphology (**Table 3**). On contrary to samples isolated in the 1980s, strains isolated from Europe (France and Sweden) in the 2000s and from USA, Japan and Australia belong mainly to the B2 group (43%), followed by the A phylogroup (24%), D (21%) and B1 (12%) groups [258],[259],[260],[261],[262].

Influence of the environment was exemplified in a study comparing gut *E. coli* community between subjects from metropolitan France, native of French Guyana or recently expatriated to Guyana. The *E. coli* community from subjects expatriated to Guyana was intermediate between the two other populations, highly suggesting that the composition of gut commensals changes dynamically in response to environment modifications [263]. Another concept that has recently emerged corresponds to the existence of both long-term residential strains and short-term transients of *E. coli* [264]. It is becoming clear that *E. coli* strains are not all equal with respect

to their ability to colonize and reside in the adult human gut. Given the high probability of exposure to *E. coli* from multiple environmental sources, humans almost certainly ingest many different types of *E. coli* throughout their adult life. Depending on their properties, ingested *E. coli* strains are able to establish in the gut and become resident for a long period or are quickly lost at the rate of gut transit. However, it appears that *E. coli* residency is a dynamic process and resident clones turn over at different rates in individuals. Several factors influence the *E. coli* residency in the gut such as diet, health, travel or exposure to antibiotics. However, the rules governing *E. coli* residency in the gut are far from being fully understood and require further research. Nonetheless, it has been shown that long-lived resident *E. coli* are phylogenetically distinct from short-lived transients and belong preferentially to phylogenetic groups A, B2 and F [264]. Gene association studies are warranted in this context to better define residency-associated traits.

Population	Phylogenetic group [†]			
	A	B1	D	B2
France (1980)	61	12.5	16	10.5
Croatia	35	32	14	19
Mali	24	58	16	2
Benin	50	32.5	0	17.5
Pakistan	47	18	23	12
French Guiana (native populations)	63.5	20.5	13	3
Bolivia (native populations)	77	10	8	5
Colombia	57	3.5	14.5	25
France (2000)	25.5	21	24	29.5
Sweden	29	11	14	46
USA	20.5	12.5	19	48
Japan	28	0	28	44
Australia	19.5	12.5	23	45

Table 3 Prevalence of the main *E. coli* groups in humans. Prevalence is given as a percentage of the four phylogenetic groups of *E. coli* in human fecal samples from the indicated populations. The phylogenetic groups were determined by triplex-PCR. *Adapted from Tenaillon et al., 2010.*

C) Functional roles of *E. coli* in the GIT

The ecological role of *E. coli* in the gut of healthy individuals, and Proteobacteria in its entirety, is poorly understood and requires further works. To date, the following roles can be attributed to *E. coli* in healthy individuals:

a) Oxygen consumption to maintain an efficient anaerobic environment in the gut

As a facultative anaerobe, *E. coli* depletes the oxygen diffusing from the mucosal surface and therefore participates to the establishment of a perfect environment for strict anaerobes which constitute almost all members of the gut microbiota. Several works suggest that *E. coli* develops a relationship with the anaerobic members of the gut microbiota. In the gut, the anaerobes degrade complex polysaccharides and release simple saccharides used by *E. coli* for growth. In return, *E. coli* helps create an anaerobic environment by scavenging oxygen [245]. This function is important in the adult but also in the neonate. Indeed, *E. coli* is one of the first bacteria to colonize the gut of neonates at birth. This newly established and rapidly growing *E. coli* population then changes the structure and function of the epithelial cells in ways that appear crucial for healthy microbiome development [265]. In particular, oxygen consumption by *E. coli* and other facultative anaerobes creates hospitable conditions for strict anaerobes to colonize and become dominant inhabitants of the gut [266].

b) Production of vitamins

In humans and animals, vitamin K, also referred as menaquinone, has multiple roles and one of them is to be a co-factor for many important enzymes. *E. coli* is known to produce menaquinone during anaerobic growth and uses it as an electron carrier in the respiratory chain [267]. With other members of the gut microbiota, *E. coli* modulate menaquinone concentrations in the human gut and contribute to supply the host for its vitamin K requirement [268]. *E. coli* is also able to synthesize vitamin B₁₂ through modifications of uptake corrinoids [269]. Vitamin B₁₂ (cyanocobalamin) is an unusual vitamin in that it is not made by plants and most vertebrates and instead is synthesized exclusively by bacteria and archaea. The essential nature of these pathways makes many eukaryotes including humans dependent on microbial corrinoids.

However, the direct contribution of gut microbial communities to host cobalamin levels is discussed. Indeed, cobalamin produced in the colon, where microbial numbers are highest, is probably not bioavailable for the host because the receptors necessary for absorbing the vitamin are mostly found in the small intestine. In humans, diet is instead more likely to constitute the primary source of cobalamin [270].

c) Protection against gut pathogen infections

The best characterized role of *E. coli* in gut homeostasis concerns its contribution to resistance against colonization by exogenous pathogens. Protection offered by the gut microbiota occurs through different mechanisms including secretion of antimicrobial products, nutrient competition, support of epithelial barrier integrity and immune activation [271]. In the case of commensal *E. coli*, the protective mechanisms are thought to rely mainly on bacteriocin synthesis and competition for nutrient sources that can limit the growth of enteric pathogens. As an example, a study from Hudault *et al.* showed that colonization of GF mice by an *E. coli* strain had a protecting effect against *Salmonella typhimurium* infection [272]. Competition for the same nutritional substrates between commensal and pathogenic *E. coli* strains has been also largely evaluated in the literature. Commensal *E. coli* consume sugars such as fucose, mannose, galactose, ribose, arabinose, and N-acetyl glucosamine in the mucus layer, sugars that are also used by pathogenic *E. coli* strain O157:H7 (EHEC) during infection. Several *in vivo* mouse studies demonstrated that commensal *E. coli* can prevent EHEC infection due to competition for nutrients such as carbohydrates, amino acids and organic acids [243],[93],[273],[94]. During my thesis, I co-authored a book chapter focusing of the relation between gut microbiota and intestinal infectious diseases. This manuscript is available in the [Annex](#).

It appears clear that enteric Proteobacteria have yet unidentified functions. Whereas commensals of this phylum are safe when they are in minor proportion, they can become colitogenic microbes triggering inflammatory responses under certain gut environments. A better understanding of the ecological roles of these bacteria is key to identifying the symbiotic and/or pathological relations between host and Proteobacteria members in the human gut.

2. Proteobacteria expansion in dysbiosis associated with human disease

2.1 Proteobacteria bloom : a marker of gut microbiota in disease

Changes in the composition of the gut-associated microbial community are observed in many human intestinal or extra-intestinal diseases, including CD and RA as described in chapter 2. Among the multiple taxonomic variations associated with gut dysbiosis, the most consistent and robust ecological pattern is an expansion of facultative anaerobic bacteria belonging to the phylum Proteobacteria (**Figure 12**). While information at a lower taxonomic level than the Proteobacteria phylum is not systematically described in the different works, the *Enterobacteriaceae* family appears to have a prominent place for explaining the expansion of Proteobacteria associated with disease. From these observations, several groups proposed that an enrichment of Proteobacteria in the gut is a marker for an unstable microbial community and a potential criterion for disease diagnostic [232], [274], [231].

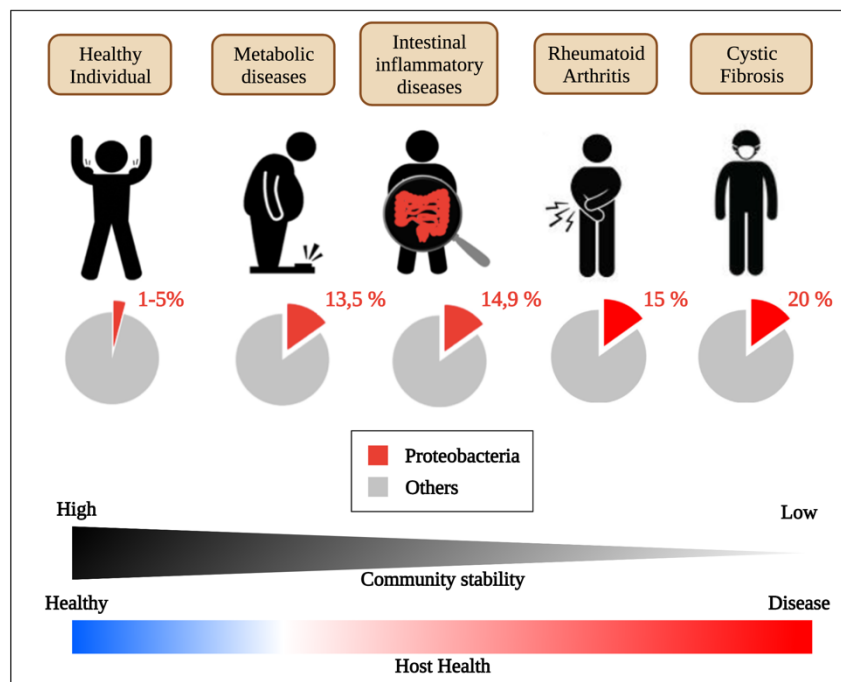


Figure 12 Expansion of Proteobacteria in gut microbiota under different host conditions. Adapted from Shin et al., 2015. Figure made with biorender.com.

Among works correlating disease state, gut dysbiosis and Proteobacteria expansion, gut inflammation at low or high levels appears to be associated with expansion of Proteobacteria in various disease contexts such as IBD, colorectal cancer, or metabolic syndrome [275], [276], [277], [278] (**Figure 13**).

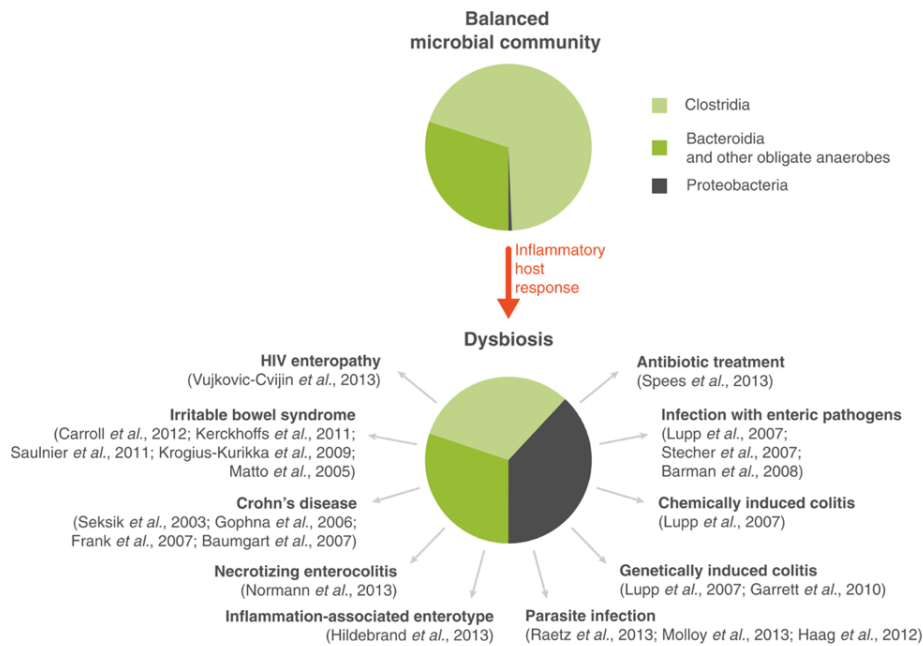


Figure 13 Changes in the gut microbiota associated with inflammation of the lower GI tract. Pie charts provide a schematic illustration of the balanced microbial community structure of the healthy lower GI tract (top) and the imbalanced microbial community structure (dysbiosis) associated with inflammation of the large bowel (bottom). Black arrows point to different studies suggesting that this association between the inflammatory host response and dysbiosis is a conserved ecological pattern observed in the lower GI tract. *From Winter and Bäumlner, 2014.*

Naturally, a legitimate question concerns the causal relationship between Proteobacteria bloom and disease state, asking what is the cause and what is the consequence. Works with mechanistic insights of the processes connecting gut dysbiosis and disease are now recent and will unravel the principles governing the balance or unbalance of microbial community structure in case of disease. For instances, a causal link between a strain of *Enterobacter cloacae* (*Enterobacteriaceae* family) and obesity has been deciphered in a work published by the group of L. Zhao [278]. During a clinical study, they found that the gut microbiota of a morbidly obese volunteer was composed at 35 % by bacteria of the genus *Enterobacter*. After a specific diet program, the weight loss of patient was concomitant with the disappearance of *Enterobacter* in the gut microbial community. Moreover, colonization of GF mice with the *Enterobacter cloacae* strain isolated from the obese human gut, leads to the development of obesity in animals fed on a high-fat diet. This work strongly suggests that specific bacterial species of the gut microbiota may causatively contribute to the development of disease in human. A better understanding of mechanisms governing disease state and gut dysbiosis would shed light on new approaches for manipulation of the gut microbiota to sustain, improve or restore balance and health to the individual.

2.2 Causes of *Proteobacteria* proliferation in dysbiotic conditions

Several mechanisms have been proposed in the literature that could explain blooms of *Enterobacteriaceae* in the dysbiotic gut which are resumed in **Figure 14**.

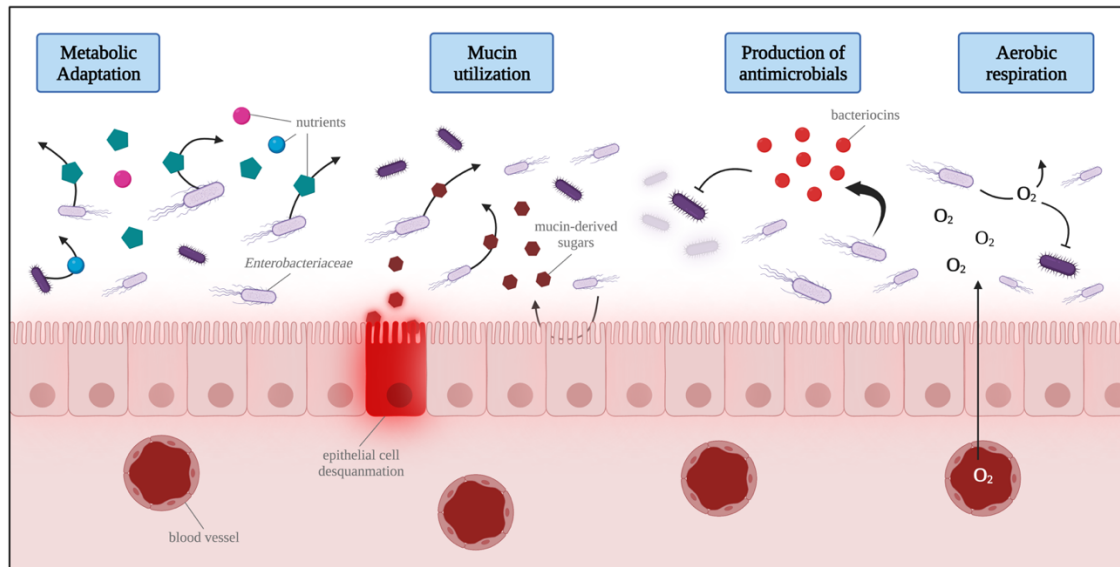


Figure 14 Mechanisms of expansion of *Enterobacteriaceae* in the dysbiotic gut. Figure made with *biorender.com*.

A) Increase of electron acceptors for respiration

Oxygen Emerging evidence supports that oxygen diffusion within the gut lumen can be a driver of gut dysbiosis. Indeed, an increase of oxygen level in the gut would destabilize the oxygen-sensitive strict anaerobes and promote the expansion of facultative anaerobes [279], [280]. For instances, antibiotic treatment or inflammation are conditions known to promote *Enterobacteriaceae* expansion. Consistent with this idea, metagenomic analysis in the murine model of colitis identified oxygen respiration as a dominant signature associated with commensal *E. coli* expansion in the gut [281],[282]. A streptomycin-mediated disruption of the gut microbiota in mice is also associated with an uncontrolled expansion of facultative anaerobic *Enterobacteriaceae* [283]. Interestingly, streptomycin treatment also leads to a depletion of bacteria belonging to the class Clostridia which are important producers of the short-chain fatty acid butyrate. In normal conditions, colonocytes are hypoxic because they consume oxygen through mitochondrial β -oxidation of microbiota-derived butyrate to CO₂, which represents their main pathway for producing energy [284]. Depletion of butyrate-producing bacteria by antibiotic treatment or other dysbiotic conditions would reduce luminal butyrate levels, resulting in a metabolic reorientation of intestinal epithelial cells towards anaerobic glycolysis, a process that does not consume oxygen. The consequent increase in the

amount of oxygen emanating from the colonic surface drives an expansion of facultative anaerobes such as *Enterobacteriaceae* in the gut.

Alternative electron acceptors for anaerobic respiration Intestinal inflammation also contributes to an increase availability of various electron acceptors that promote anaerobic respiration by *Enterobacteriaceae* [285] (Figure 15). The host inflammatory response produces reactive nitrogen species and reactive oxygen species that, upon diffusion into the gut lumen, will react with organic sulphides, tertiary amines or other molecules to form S-oxides, N-oxides and nitrate. *Enterobacteriaceae* can use these molecules as terminal electron acceptors for anaerobic respiration by expressing DMSO, TMAO and nitrate reductases, respectively [286] [287]. For instances, *E. coli* encodes two DMSO reductases (*dmsABC*, *ynfFGH*), three TMAO reductases (*torCAD*, *torYZ*, *yedYZ*) and three nitrate reductases (*narGHJI*, *narZYWV*, *napFDAGHBC*). Such alternative electron acceptors-rich environment therefore confers a growth advantage for *Enterobacteriaceae* in inflammatory conditions. Indeed, anaerobic respiration is considered a process more efficient for energy production than fermentation made by other members of the gut microbial community.

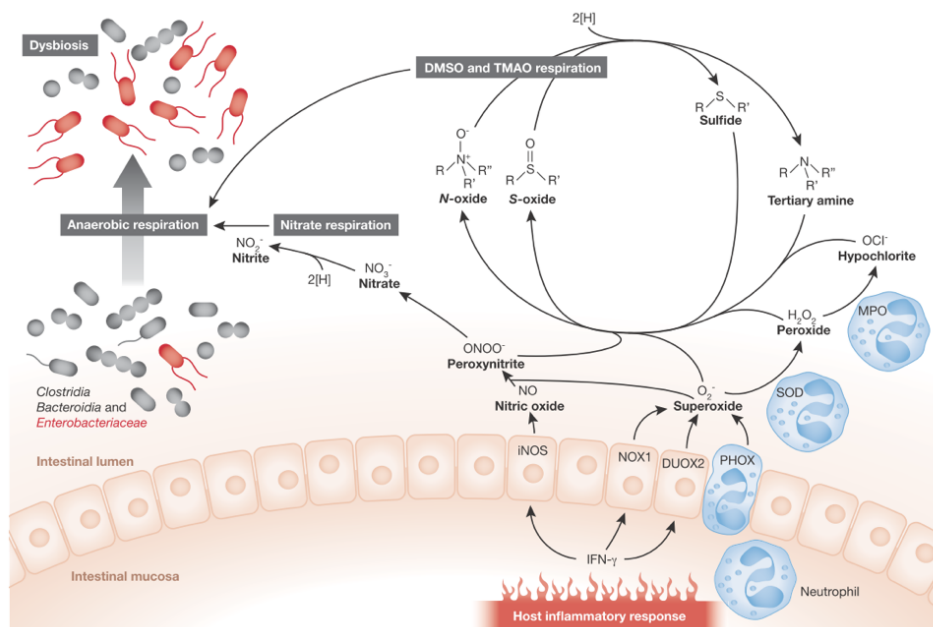


Figure 15 Phylum-level changes in the microbiota composition on intestinal inflammation. Inflammatory infiltrates are a source of or induce the expression of genes that generate antimicrobial radicals, such as superoxide, peroxide, hypochloride, nitric oxide and peroxynitrite. In the intestinal lumen, these radicals react to form harmless oxidation products that serve as electron acceptors. These support the growth of facultative anaerobic bacteria by anaerobic respiration. From Winter et al., 2013. Legend : DMSO, dimethyl S-oxide; DUOX2, dual function NAD(P)H oxidase 2; IFN- γ , interferon gamma; iNOS, inducible nitric oxide synthase; MPO, myeloperoxidase; NOX1, NADPH oxidase 1; PHOX, phagocyte NADPH oxidase; SOD, superoxide dismutase; TMAO, trimethylamine N-oxide

B) Metabolic adaptation to dysbiotic conditions

Dysbiosis often causes significant changes in the gut environment and may influence the diversity and concentrations of nutrients available in the lumen that the different members of the gut microbiota can utilize with variable efficacy. As mentioned previously, *Enterobacteriaceae* and particularly *E. coli* present a dynamic genome structure leading to a high level of genome plasticity. Moreover, *Enterobacteriaceae* have been shown to be the major source of variable genes in the gut microbiome between healthy individuals whereas they represent only a minor population, indicating that this family brings variability in gut microbial gene function [288]. Consistent with these observations, changes in the gut environment caused by disease-associated dysbiosis can lead to Proteobacteria expansion as a result of a better adaptability to metabolize potential arising substrates. Different studies have explored microbial metabolic adaptation and some nutrient sources available to *Enterobacteriaceae* during intestinal dysbiosis are listed in **Table 4**. Although some of these studies have been done with enteric pathogens, it is reasonable to believe that most traits can be transposed to commensal *Enterobacteriaceae*.

Nutrient	Source	Species	Reference(s)
Ethanolamine	Host	<i>S. Typhimurium</i> ,	[86]
Lactate	Host	<i>S. Typhimurium</i>	[87]
Glucarate/galactarate	Microbiota	<i>S. Typhimurium</i> , commensal <i>E. coli</i>	[88]
1,2-propanediol	Microbiota	<i>S. Typhimurium</i>	[90]
Succinate	Microbiota	<i>S. Typhimurium</i>	[91]
L-Serine	Diet	Adherent-invasive <i>E. coli</i> , <i>C. rodentium</i>	[92],[102]
Glycerol	Host	Commensal <i>E. coli</i>	Metamouros <i>et al.</i> , 2018

Table 4 Studies available reporting nutrient sources available to *Enterobacteriaceae* during intestinal dysbiosis. Adapted from Shelton and Byndloss, 2020. Reference regarding Glycerol added from Metamouros *et al.*, 2018.

Some examples illustrating metabolic adaptation and/or specificity involved in the expansion of *Enterobacteriaceae* in dysbiotic gut are presented below.

a) Glycerol in cystic fibrosis patient

Cystic fibrosis is a genetic disease affecting the normal functions of multiples organs including lung, pancreas and GIT with a pronounced inability of patients to digest and absorb proteins and lipids [289]. A gut dysbiosis characterized by a bloom of *Enterobacteriaceae* is associated with cystic fibrosis. Indeed, *E. coli* that normally represents less than 1% of the gut microbiota, can reach more than 90% in the fecal microbiota of young children with cystic fibrosis compared with age-matched healthy controls [290]. In 2018, Matamouros *et al.* isolated *E. coli* strains from fecal samples of young children suffering from this disease and showed an increased growth rate in the presence of glycerol as a sole carbon source [291]. Glycerol is a major component with fatty acids of intestinal fat present in high concentrations in the gut of patients as a result of fat malabsorption. In addition, cystic fibrosis and control *E. coli* isolates have differential gene expression when grown in presence of glycerol, resulting likely in a better growth rate of the patient strain. The authors suggest that *E. coli* expansion in the gut of cystic fibrosis children can be explained through adaptation and/or selection of *E. coli* strains possessing a high ability to metabolize glycerol [291].

b) Intestinal mucosa-derived substrates

In the case of diseases marked by an important gut inflammation, alteration of the gut environment can generate a unique set of carbon sources available to *Enterobacteriaceae* to outgrow other members of the gut microbiota. For instances, inflammation cause shedding of epithelial cells in the lumen and alterations of the mucus layer. Components derived from the epithelial cell membrane such as phosphatidylcholine and phosphatidylethanolamine increase in this context. Ethanolamine (EA) is derived from phosphatidylethanolamine and can be used as a source of carbon and/or nitrogen by a variety of species in the Firmicutes, Actinobacteria and Proteobacteria phyla, as well as by pathogenic species such as *Salmonella* and *Pseudomonas* [292]. Utilization of EA involves breaking ethanolamine into ammonia and acetaldehyde thanks to a specific ethanolamine ammonia lyase [282]. The ammonia resulting from this reaction, can be used as a cellular supply of reduced nitrogen, and the acetaldehyde is converted to acetyl-CoA that can be used in various bacterial metabolic cycles such as the tricarboxylic acid cycle, the glyoxylate cycle, or lipid biosynthesis [292].

The mucus layer of the gut epithelium is composed of mucin proteins that are glycosylated with five major sugars that can be potential nutrients for bacteria: N-acetyl-D-galactosamine, N-acetyl-D-glucosamine, N-acetylneuraminic acid (sialic acid), L-fucose, and D-galactose. In case of gut inflammation, there is an increase in production and secretion of mucins as a mechanism of maintaining the integrity of the mucus layer. In a study using a mouse model of induced gut inflammation, it has been reported that *E. coli* outgrowth in the inflamed gut consequently to an increase of commensal species expressing sialidases. Such increased activity leads to the release of high amount of free sialic acid which is efficiently consumed by *E. coli* [293]. In another study examining CD patients, it was detected an increase of glycosidase encoding genes in the microbiome of inflamed gut samples, consistent with an enhanced mucin-degrading activity. Mucin degrading-bacteria such as *Mucispirillum schaedleri*, *Akkermansia muciniphila* and *Bacteroides acidifaciens*, expand and mediate release of less complex sugars (lactose, melibiose, raffinose, and galactinol) from mucins [282]. These simple sugars and metabolites accumulate and may lead on a decrease of commensal bacteria from Bacteroidia and Clostridia classes in the inflamed gut but also confer a growth advantage to *Enterobacteriaceae* as well as to pathogens such as *S. Typhimurium* and *Clostridium difficile* [282],[96].

c) Diet-derived substrates

A recent work by Kitamoto *et al.* shows that the inflamed gut also alters the amino acid availability in the gut lumen [294]. In the dysbiotic gut of IBD patients, AIEC reprogram their metabolic pathways to use L-serine in order to gain fitness advantage over resident microbiota. L-serine is acquired from diet and can be used as a source of energy after conversion into pyruvate which is a substrate necessary for gluconeogenesis and the tricarboxylic acid cycle [295]. While this metabolic pathway has a minor role on AIEC fitness in the healthy gut, this amino acid is a key resource for AIEC expansion during IBD intestinal dysbiosis. This study has also successfully demonstrated that restriction of dietary amino acid, particularly L-serine, intake prevents the expansion of *Enterobacteriaceae* in the inflamed gut. This illustrates the importance of understanding the nutrients preferentially used by *Enterobacteriaceae* to thrive during dysbiosis in order to provide novel approaches to restore balance of the gut microbiota.

RESULTS

PART 1

Identification of metabolic properties potentially involved in the expansion of *Enterobacteriaceae* in the gut of patients suffering from CD or RA

As described in the bibliographical introduction of my manuscript, a bloom of *Enterobacteriaceae* has been proposed as a signature of dysbiosis in several inflammatory diseases such as CD and RA. Identifying the mechanisms by which *Enterobacteriaceae* thrive in the dysbiotic gut could open new doors on therapeutic approaches and on ways to modulate the gut microbiota composition and restore gut homeostasis and health. Among the mechanisms proposed in the literature to explain Proteobacteria expansion during dysbiosis, nutritional adaptation of bacteria to changes in the gut environment can lead to Proteobacteria expansion as a result of a better adaptability to metabolize potential arising substrates.

The first aim of my PhD project was to compare the metabolic capacity of *Enterobacteriaceae* sampled either from healthy individuals or from patients with a pathology associated with an *Enterobacteriaceae* expansion. The differences identified, could open potential investigations on the contribution of such particular metabolic abilities to explain the expansion of *Enterobacteriaceae* in the dysbiotic gut. For this purpose, collections of *Enterobacteriaceae* strains isolated from fecal samples of healthy individuals (HI), of RA patients or of CD patients were constituted. For HI and RA collections, fecal samples were spread on Mac Conkey plates, a medium selective for *Enterobacteriaceae*. Interestingly and accordingly to the literature, we observed that the level of *Enterobacteriaceae* per gram of faeces was higher in the group of RA patients than in the group of HI (**Figure 16**), indicating that differences in *Enterobacteriaceae* populations were observed between the two groups despite their small size.

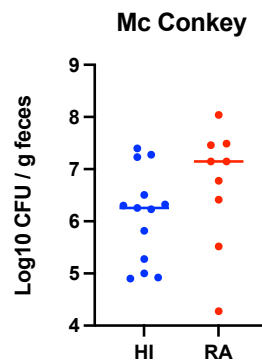


Figure 16 *Enterobacteriaceae* isolated on McConkey medium represented by CFU/g feces. Fecal samples were spread on Mac Conkey plates, a medium selective for *Enterobacteriaceae*. Each point represents a fecal sample from an individual.

From Mac Conkey plates, approximately 10 colonies per individual were subjected to RAPD-PCR in order to exclude clonal isolates from a same individual. Taxonomic identification was then determined by amplification and sequencing of the rDNA 16S. Most *Enterobacteriaceae* selected on Mac Conkey plates were of the genus *Escherichia* (>75%). Because the aim of our study is to compare metabolic capacities between strain collections and because these properties can largely vary among bacterial genus, we decided to conserve only *Escherichia* strains in these collections. Finally, the HI and RA collections are constituted respectively of 16 strains collected from 15 different individuals and of 14 strains collected from 9 different patients. The CD collection of strains was provided by the laboratory M2iSH (Université Clermont Auvergne, France). For this group, each *Escherichia* strain was sampled from feces of different CD patients. As described in the chapter 2 of my literature review, the prevalence of a specific *E. coli* type called AIEC is higher in CD patients than in HI. Therefore, we first analyzed the prevalence of AIEC in our HI and CD collection of strains. In that aim, invasion tests were performed on epithelial cells using AIEC prototype strain LF82 as a positive control and *E. coli* K12 C600 strain as a negative control. Results represented in **Figure 17**, show that 3 HI strains but no CD strain present a low degree of invasive properties (C2, N3 V3) within epithelial cells. These data strongly suggests that our collection of CD strains does not include AIEC.

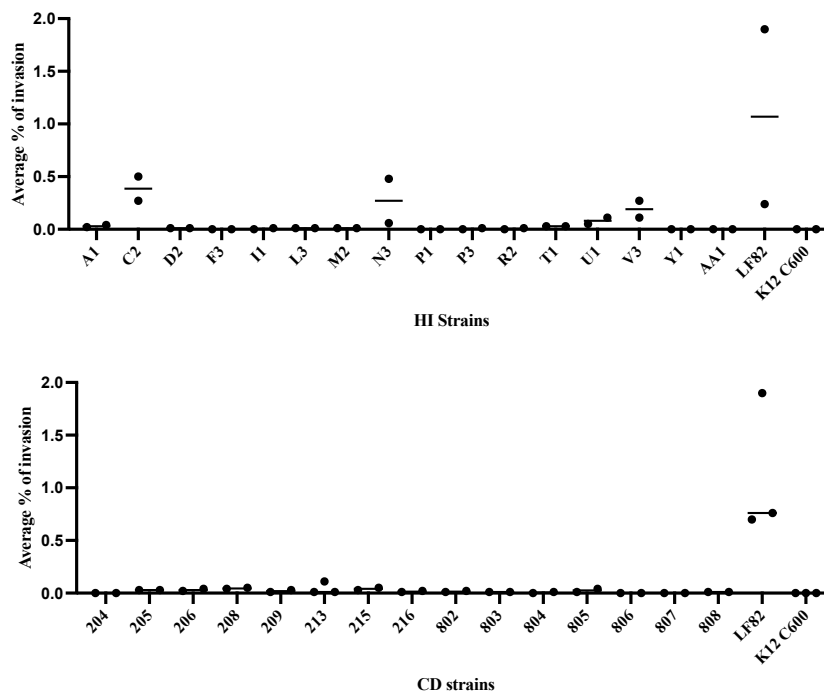


Figure 17 Invasion tests on epithelial cells using our HI and CD collection of strains. Each point represents a replicate and bars represent the mean. Strain *E. coli* LF82 was used as a positive control and *E. coli* K12 C600 as a negative control.

With the aim to compare metabolic activities of *Enterobacteriaceae* isolated from HI, RA or CD patients, we decided to screen our 3 groups of strains by using the OmniLog technique that provides a semi-high throughput assay for characterization and monitoring of microbial phenotypes. Two plates, PM01 and PM02, are available to test the utilization of substrates as a carbon source. Each well includes a different molecule as a unique carbon source, allowing to evaluate the ability of each strain to use a total of 190 different compounds (<https://www.biolog.com/wp-content/uploads/2020/04/00A-042-Rev-C-Phenotype-MicroArrays-1-10-Plate-Maps.pdf>). The OmniLog system is composed of a plate incubator equipped with a camera quantifying cell respiration (NADH production) as a universal reporter. If the inoculated strain is capable of using the substrate, the cell respire actively, reducing a tetrazolium dye and forming a strong color that is recorded by the OmniLog system. The OmniLog approach was therefore used to screen HI, RA and CD collections in order to find potential differential metabolic properties. Following this screening, D-serine was identified as better metabolized by the CD strains while sucrose by the HI strains. Additional works have been undertaken to characterize and explore the potential role of D-serine and sucrose in the fitness of strains from both groups. This work awaits upcoming results, but our findings are presented herein as the core parts of a manuscript to be submitted.

MANUSCRIPT 1

Identification of metabolic properties potentially involved in the expansion of *Enterobacteriaceae* in the gut of patients suffering from inflammatory diseases

Maria Ines Moreira de Gouveia¹, Annie Garrivier¹, Julien Daniel¹, Christophe Del'homme¹
Nicolas Barnich², Annick Bernalier-Donadille¹, Gregory Jubelin^{1*}

¹ *Université Clermont Auvergne, INRAE, MEDIS UMR454, F-63000 Clermont-Ferrand, France*

² *Université Clermont Auvergne, Inserm U1071, USC-INRAE 2018, M2iSH, Clermont-Ferrand, France.*

* corresponding author: Gregory Jubelin, INRAE Auvergne Rhône-Alpes, UMR454 MEDIS, site de Theix, 63122 Saint-Genès-Champanelle, France. Email: gregory.jubelin@inrae.fr

ABSTRACT

Gut dysbiosis associated with several intestinal or extra-intestinal diseases is often characterized by an expansion of *Enterobacteriaceae* that usually represent only 1 % of the gut microbiota in the healthy human. This proliferation can be explained in part by emergence and proliferation of strains fitting with the modified gut environment induced by the disease. In this study, we performed a screening to identify substrates differentially catabolized between *Enterobacteriaceae* commensal strains isolated either from healthy individuals (HI) or from patients suffering from Crohn's disease (CD) or rheumatoid arthritis (RA), two diseases associated with gut dysbiosis. Among the different substrates identified, we observed an opposite ability of HI and CD strains to metabolize D-serine and sucrose, a characteristic arising from genetic rearrangements in the *argW* locus of *Escherichia coli* chromosome. The D-serine⁺ phenotype of CD strains was shown to be dependent of the presence of *dsdCXA* genes and to confer a nutritional advantage over HI strains. Our study provides new clues on potential substrates that could promote the expansion of *Enterobacteriaceae* in dysbiotic conditions.

KEYWORDS :

gut dysbiosis, Crohn's disease, Rheumatoid Arthritis, *Enterobacteriaceae*, nutritional competition, D-serine

INTRODUCTION

During lifetime, the gut microbiota is relatively stable in the healthy individual (HI) and plays an essential role in human health. When in balance, the gut microbiota provides energy for the host and plays a structural, protective, and immunological role [1]. Disruption of the gut microbiota homeostasis, termed dysbiosis, can negatively impact health of the individual. Gut microbiota imbalance can be defined by typical features such as a decrease in microbial diversity and/or richness, a decrease in beneficial species and/or an increase of deleterious bacteria such as Proteobacteria. Indeed, Proteobacteria usually represent 1-5% of the healthy gut microbiota but the proportion of this phylum can significantly increase in dysbiosis associated with various intestinal or extra-intestinal diseases [2]. The *Enterobacteriaceae* family, especially *Escherichia coli* species, appears to have a prominent place for explaining the expansion of Proteobacteria associated with disease. An enrichment of Proteobacteria in the gut has been even proposed to be a marker for an unstable microbial community and a potential criterion for disease diagnostic [3]–[5]. Several hypotheses have been proposed to explain Proteobacteria expansion in the dysbiotic gut. Among them, adaptation or selection of strains fitting with the modified gut environment associated with disease can occur as in cystic fibrosis patient for which fat malabsorption in the gut leads to expansion of *E. coli* adapted to high level of intestinal glycerol [6]. In addition, *Enterobacteriaceae* displays a high level of genome plasticity and has been shown to be the major source of variable genes in the gut microbiome between healthy individuals whereas they represent only a minor population, indicating that this family brings variability in gut microbial gene function [7]. Consistent with these observations, changes in the gut environment caused by disease-associated dysbiosis can lead to Proteobacteria expansion as a result of a better adaptability to metabolize potential arising substrates. The aim of our study was to explore this hypothesis in the context of two inflammatory diseases associated with Proteobacteria expansion: Crohn's disease (CD) and

rheumatoid arthritis (RA) [8]–[10]. CD is a chronic, segmental, inflammatory disease that can affect all parts along the GIT that can cause lesions from mouth to anus [11]. RA is a chronic autoimmune inflammatory extra-intestinal disease characterized by a constant immune response located at the joints leading to inflammation [12]. In this work, we performed a screening of *E. coli* strains isolated from HI or from CD or RA patients for their capacity to metabolize an array of diverse carbon substrates. Several molecules including carbohydrates and amino acids were found to be differentially utilized between the strain groups. In particular, the ability to catabolize D-serine and sucrose was significantly distinct between HI and CD strains. Interestingly, the two substrates required distinct sets of genes that are both localized in the highly variable *argW* locus of the *E. coli* genome [13]. The D-serine phenotype of CD strains was shown to be correlated with the presence of *dsdCXA* genes and to confer a nutritional advantage over strains isolated from healthy individuals in specific conditions.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Enterobacteriaceae from the collection were isolated from fecal samples of either healthy individuals (HI), individuals with rheumatoid arthritis (RA) or individuals with Crohn's disease (CD) (CD strains were provided by M2iSH, Clermont-Ferrand, France). Strains used in this study are listed in [Table S1](#). For HI and RA individuals, stool samples were diluted and spread on MacConkey and Hektoen agar plates. A RAPD-PCR [14] was performed on several colonies of each sample in order to detect and exclude clonal isolates from our collection. Taxonomic identification of each strain was determined by amplification and sequencing of the rDNA 16S and only *Escherichia* strains were conserved in the collection. *E. coli* strains were cultured in LB or in a modified M9 minimal salt medium containing NH₄Cl (18.5 mM), Na₂HPO₄ (42 mM), KH₂PO₄ (22 mM), NaCl (8.5 mM), MgSO₄ (2 mM) and CaCl₂ (0.1 mM). Carbon source

added to the M9 medium was either glucose 0.02% or 0.2%, sucrose 0.2%, D-serine 0.2% or a mix of glucose 0.02% + D-serine 0.2%. For growth in M9 media, bacterial strains were first isolated on LB plates and LB precultures inoculated from a single colony were performed for 8 h at 37 °C with aeration. The precultures were then diluted 100-fold in M9 + glucose and grown overnight. Precultures were washed once with M9 and used to inoculate cultures in M9 media with different carbon sources. Cultures were incubated with shaking at 37 °C and growth was evaluated by measuring OD_{600nm}. When required, antibiotics were used at the following concentrations: kanamycin (25 µg.ml⁻¹); ampicillin (100 µg.ml⁻¹).

Detection of *dsd* and *csc* operons

Detection of *dsd* and *csc* operons was performed with a triplex PCR using genomic DNA extracted from indicated strains and primers *dsdA_R*, *cscR_F* and *dsdX_F* ([Table S2](#)). The amplification was performed with Taq DNA polymerase under the following cycling conditions: one cycle of initial denaturation at 94°C for 2 min, 35 cycles of denaturation at 94°C for 1 min, annealing for 30 sec at 57°C, and extension at 72°C for 1 min, followed by a final extension at 72°C for 5 min. The amplified products were separated on agarose gel stained with ethidium bromide. For sequenced strains, *in silico* triplex PCR were performed with the same 3 primers and results are presented as a virtual agarose gel (SnapGene® software, Insightful Science).

Phenotype Microarray experiments

Phenotype MicroArray technique (Biolog®) provides a semi-high throughput assay for characterization and monitoring of microbial phenotypes. In our study, plates PM01 and PM02 were used to test 190 substrates for carbon utilization. Strains were isolated on LB plates from the -80°C stock and a single colony was then re-streaked on LB plates one day before

microarray experiment. For each strain, a pool of colonies was scraped with a cotton swab to inoculate IF-0 1X mixed with Dye A until 85% of transmittance was reached (Turbidimeter Biolog®). Each PM plate was inoculated with 100 µL per well of inoculated media and incubated within the OmniLog system for 24 h at 37°C with output reads recorded every 15 min. The raw kinetic values were exported to the Data Analysis Software (version 1.7) and area under the curve (AUC) values were used for subsequent analyses.

Selection of strains representative of HI, CD and RA groups

Selection of representative strains for each group was done by making a score table as follow: first, the percentage of strains catabolizing a substrate was determined for each strain group and substrate was defined to be used by the group is the percentage was above 66%. In contrast, if less than 33% of strains of a group were unable to metabolize a substrate, it was considered that the group is unable to use the substrate. Substrates presenting percentages between 33% and 66% of utilization for a group were considered heterogeneous and were discarded to make the score table. For each strain, a score was calculated by giving 1 point per substrate when the strain phenotype was in agreement to the phenotype of the group. Strains presenting the highest score were considered the most representative of a group.

Construction of *dsdCXA* mutant

Replacement of the locus *dsd* by the gene conferring resistance to kanamycin were obtained in strain 204 by using the one-step PCR-based method of Datsenko and Wanner [15]. Designed primers used to construct the mutants are presented in Table S2. Mutations were confirmed by PCR and sequencing using specific primers for *yfdC* and *emrY* genes flanking the *dsd* locus (Table S2).

In vitro competition experiments

Competitions assays between the CD 204 strain and the HI T1, R2 or A1 strains (the 3 HI strains are naturally resistant to ampicillin) were performed *in vitro* in M9 supplemented with glucose 0.02% or a mix of glucose 0.02% and D-serine 0.2%. Briefly, strains were cultured overnight in LB, washed once with M9 and used to inoculate M9 media at a ratio CD:HI of 1:1. Cultures were incubated at 37 °C for 24 h and samples were diluted in PBS before plating on LB and LB + Amp plates to count CD + HI strains and HI strain alone, respectively. The CD population was obtained by subtracting CFU from LB + Amp plates to CFU from LB plates. Results are expressed as competitive indices corresponding to the ratio between CD strain CFU and HI strain CFU (CD / HI).

In vivo competition experiments

Germ-free (GF) Fisher 344 albinos rats were bred at the INRAE facilities of Clermont-Ferrand-Theix, and housed in polycarbonate cages on wood chips in an environmentally controlled room (21 °C) with a 12-h light-dark cycle. GF rats were kept in positive-pressure sterile isolators with free access to irradiated standard rodent diet (UAR, Villemoisson, France) and sterilized drinking water. At 8 weeks of age, each GF rat was co-inoculated intragastrically with A1 and 204 strains at day 0 with 1mL of PBS containing each strain at a ratio 1:1. For one group of rats, drinking water was supplemented with 0,2 % of D-serine, starting from one day prior to bacterial inoculation and during the 15 days of the experiment. At indicated days, fecal samples were collected, homogenized in PBS and diluted before plating on LB or LB + Amp plates to count the total bacteria or strain A1, respectively. The 204 population was obtained by subtracting CFU from LB + Amp plates to CFU from LB plates.

RESULTS AND DISCUSSION

Phenotype microarrays identified carbon substrates differentially used between HI and CD or RA strains

Screening with OmniLog was performed using 16 HI strains, 14 RA strains and 15 CD strains and plates PM01 and PM02 were used to evaluate the utilization of 190 different substrates as a carbon source. To get a qualitative overview of potential differences between strain groups, the OmniLog Data Analysis Software generates graphics with an overlay of mean growth curves between 2 strain groups (**Figure 1**). Areas of overlap are colored yellow whereas differences are highlighted as patches of red when the substrate is better used by HI strains or as patches of green if the substrate is better used by RA or CD strains. *E. coli* strains were unable to use several substrates as a unique carbon source, especially in plate PM02 which includes polymers, amino acids and L-sugars. For other substrates, a similar metabolization pattern was observed in most cases but some differences were also observed between the HI strain group and RA or CD strain groups. In a second step, we performed a quantitative analysis of growth variation between strain groups by extracting values of the area under the curve (AUC). AUC values were then compared between strain groups for individual substrate (**Figure 2**). Few significant differences were detected between the HI strain group and the RA strain group. *E. coli* strains isolated from RA patients present significantly higher AUC for α -ketobutyric acid and D-malic acid and significantly lower AUC for D-tagatose and L-sorbose when comparing to HI group ($P \leq 0.01$) (**Figure 2A**). In contrast, more substrates were found to be differentially consumed between the HI strain group and the CD strain group. Indeed, D-serine and D-Arabinose were both identified as compounds that were significantly better catabolized by the CD group than by HI strains ($P \leq 0.001$ and $P \leq 0.0001$ respectively) (**Figure 2B**). The AUC for 8 other substrates such as sucrose, D-raffinose or D-xylose were significantly higher for the group HI comparing to CD strains.

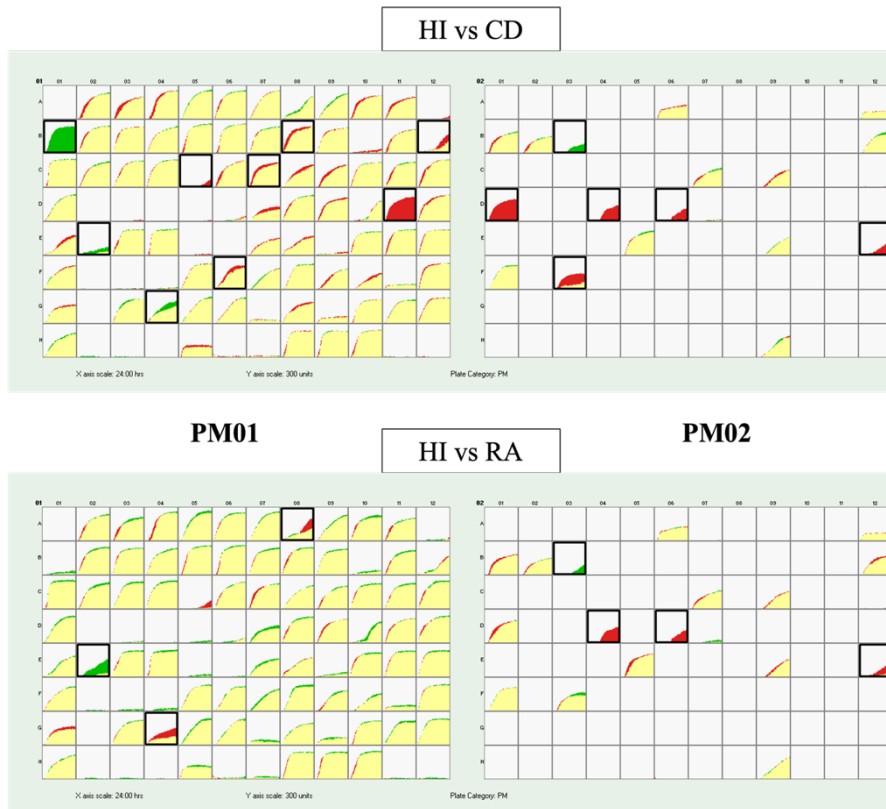


Figure 1 Growth comparisons between HI, RA and CD strain groups for 190 carbon substrates. Plates PM01 and PM02 (plate composition available at <https://www.biolog.com/wp-content/uploads/2020/04/00A-042-Rev-C-Phenotype-MicroArrays-1-10-Plate-Maps.pdf>) were individually inoculated with 16 HI strains, 14 RA strains or 15 CD strains. After a 24 h period incubation, means of AUC were represented for each well as follow: yellow colored AUC indicate similar growth between HI and CD or RA strain groups; green colored AUC indicate a higher growth for CD or RA strain group than HI strain group; red colored AUC indicate a higher growth for HI strain group than for CD or RA strain group.

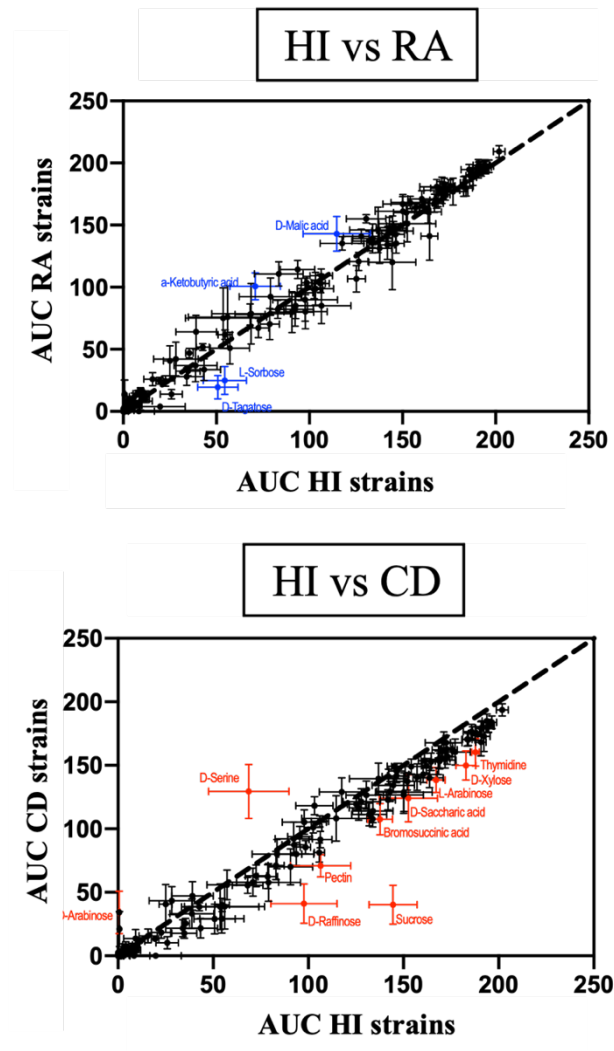


Figure 2 AUC comparison between HI strain group and CD or RA strain groups. The AUC calculated by the OmniLog Data Analysis Software were plotted on a 2-axis graph in which x-axis corresponds to AUC from HI strain group and y-axis correspond to RA (**A**) or CD (**B**) strain groups. Data are represented by black dots and associated error-bars (standard deviation). Significant differences ($P < 0.05$) are indicated with blue symbols in panel A (HI vs RA) or with red symbols in panel B (HI vs CD).

Utilization of D-Serine is differential between HI and CD strains

An interesting point arising from the phenotype microarray study concerns the metabolic phenotypes of D-serine and sucrose that were detected to be inversely correlated between HI and CD strain groups (**Figure 2B**). *dsdCXA* and *cscRAKB* clusters, involved in the catabolism of D-serine and sucrose respectively, are both located in the chromosomal *argW* locus. This region has been shown to be hypervariable and many strains present a substitution of *dsdCXA* genes by the sucrose utilization genes, *cscRAKB* [13], [16]. We therefore decided to focus our attention on the metabolic capacity of HI and CD strains for these 2 substrates. First, individual growth curves from the Omnilog screening were examined for both substrates (**Figure 3**). Only 37% of HI strains (6 of 16) grew with D-serine as the unique carbon source whereas more than 73% of CD strains (11 of 15) were able to consume it. Conversely, 94% of HI strains (15 of 16) consumed sucrose whereas only 40% of CD strains (6 of 15) harbored this phenotype. The Omnilog data obtained for these two substrates were verified and confirmed by growth assays in minimal media M9 containing D-serine and sucrose as a sole carbon source (data not shown). The capacity of *E. coli* to metabolize D- and L-serine has already been explored in the literature even though more studies focused on the L-enantiomer. Indeed, it has been shown that dietary L-serine confers a competitive fitness advantage to *Enterobacteriaceae* in the inflamed gut [17]. In this recent study, authors showed that *E. coli* strain LF82 shifts its metabolism to catabolize L-serine in order to maximize its growth potential only in case of gut inflammation. Less information are available on the catabolism of D-serine in the gut even if D-serine is present in the gut at μM concentrations [18]. Interestingly, D-serine has an important role as a signaling molecule in intestinal pathogenic *E. coli* such as enterohaemorrhagic *E. coli* (EHEC).

Indeed, it has been shown that D-serine is used by EHEC to control the expression of virulence gene in order to express colonization factors only in the presence of a favorable environment [19].

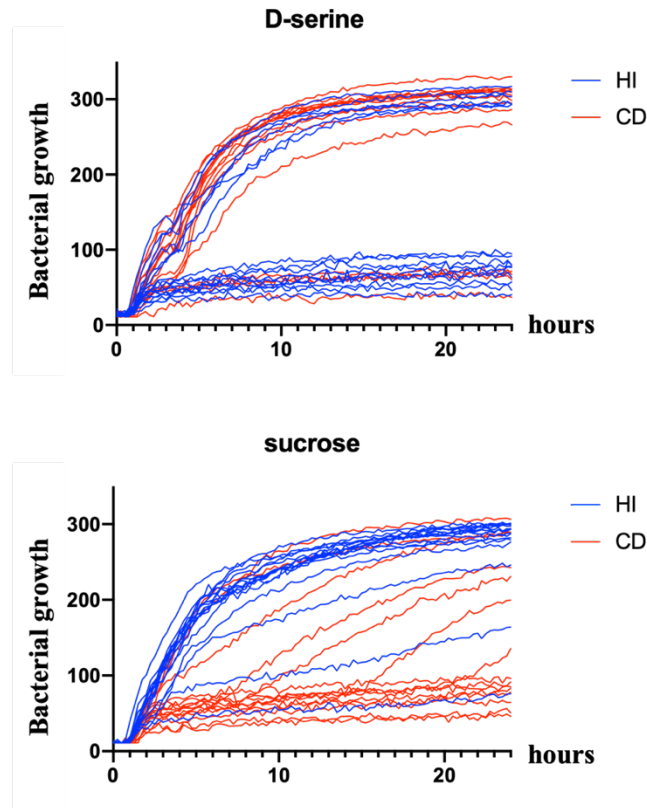


Figure 3 Representation of growth curves for individual HI and CD strains in presence of D-serine or sucrose as a carbon source. Growth curves are represented as a quantification of tetrazolium dye reduction (y) in function of time (x) for D-serine (A) and sucrose (B). Growth curves of HI strains are in blue and growth curves of CD strains are in red.

Link between phenotype and the presence of *dsd* and *csc* locus

As mentioned before, the *argW* locus is a hotspot site for integration of mobile genetic elements and *E. coli* strains can either possess the *dsd* locus responsible for the utilization of D-serine or the *csc* locus responsible for the catabolism of sucrose. In order to correlate the metabolic phenotypes with the *argW* locus organization, we investigated the presence of *dsd* and *csc* locus in HI and CD strains by a triplex PCR using primers annealing *dsdX*, *dsdA* or *cscR*. While a PCR product of 730 bp is indicative of the presence of the *dsd* locus, a PCR product of 550 bp reveals the presence of the *csc* locus (Figure 4A). As shown in Figure 4B, the ability to use D-

serine by HI and CD strains was systematically associated with the presence of the *dsd* locus without exception. On the contrary, there were some exceptions for the sucrose phenotype. Indeed, strains 807 and 808 cannot utilize sucrose whereas the triplex PCR suggests the presence of the *csc* cluster. However, we cannot exclude these strains present defective *csc* genes. Additionally, some strains (205, V3, D2, F3, M2) present a *dsd* locus in the *argW* region but were able to consume both D-serine and sucrose. This could be explained by the fact that sucrose catabolism is not strictly dependent of the presence of the *csc* locus. Indeed, sucrose can also be used by enteric bacteria thanks to the presence of the *scr* gene cluster. The *scr* genes encode a sucrose phosphotransferase system and can be found either on plasmids [20] or on the chromosome [21], including within mobile elements such as transposons [22].

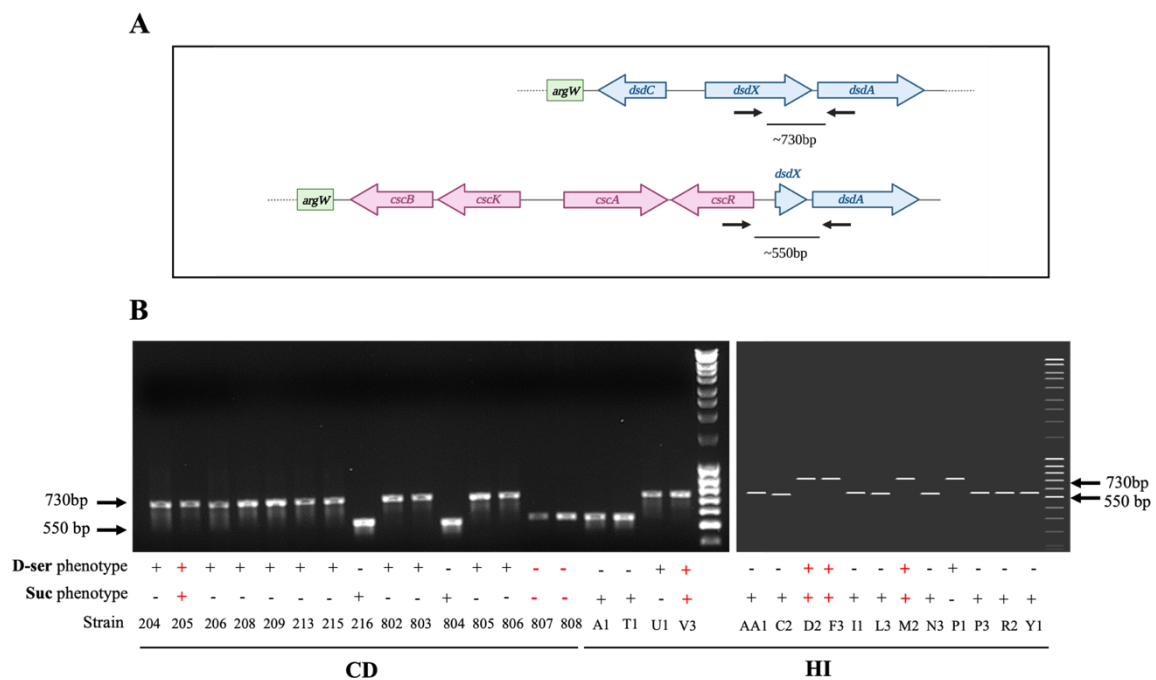


Figure 4 Correlation between D-serine and sucrose utilization phenotypes and presence of *dsd* and *csc* loci in HI and CD strains (A) Schematic representations of the *argW* locus in *Escherichia coli*. Strains present either the *dsd* locus (blue arrows) responsible for the utilization of D-serine or the *csc* locus (pink arrows) responsible for sucrose utilization. A triplex PCR using primers (thin black arrows) annealing sequences in *dsdX*, *dsdA* or *cscR* genes was used to detect *argW* configuration in indicated strains (B) Electrophoresis gel of the triplex PCR (left panel) and in silico gel for strains with available genome (right panel). A PCR fragment of 730 bp indicates the presence of the *dsd* locus and a PCR fragment of 550 bp indicates the presence of the *csc* locus. D-serine and sucrose utilization phenotypes are indicated with + or - symbols.

Selection of strains representative of HI and CD strain groups

In order to evaluate the influence of D-serine phenotype on the fitness of HI and CD strains, it is necessary to select the best representatives of the HI and CD group of strains. To this end, we compared the substrate phenotype for individual strains with the mean phenotype of the strain group by making a score table in which 1 point was given to a strain when its phenotype was the same than the phenotype of its corresponding strain group (HI or CD) (see M&M for details). Points from the different substrates were cumulated for each strain and the strains presenting the highest scores were chosen to represent a group (**Figure 5**). The strains A1, R2 and T1 were selected as representatives of the HI group and the strains 204 and 208 as representatives of the CD group. As expected, the HI representative strains have sucrose⁺ and D-serine⁻ phenotypes and CD representative strains have sucrose⁻ and D-serine⁺ phenotypes.

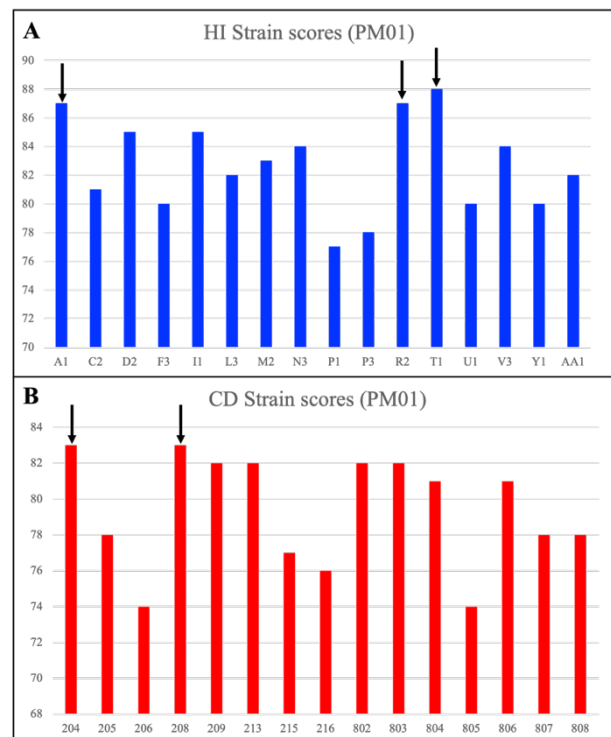


Figure 5 Determination of the best representative strains for HI and CD groups. A score was attributed to each strain based on adequacy between substrate utilization of individual strains and average utilization by their corresponding strain group (see M&M for details). Black arrows indicate representative strains selected to represent the HI strain group (A) and the CD strain group (B).

Utilization of D-serine gives an advantage to strain 204 over HI strains in vitro

We next evaluated if D-serine utilization by the CD representative strains 204 and 208 depends on the presence of the *dsd* cluster by constructing $\Delta dsdCXA$ mutants and testing their ability to grow on minimal media supplemented with D-serine as the sole carbon source. For unknown reasons, the mutagenesis was successful only for the 204 strain. Whereas the mutant grew as well as the WT strain in minimal media supplemented with glucose, it was not able to grow in M9 supplemented with 0.2% of D-serine (**Figure 6**), demonstrating that D-serine catabolism of the 204 strain depends on the presence of the *dsd* locus. In order to determine if utilization of D-serine gives a competitive advantage to CD strains over HI strains, competition assays were performed *in vitro* on minimal media with 0.02% of glucose and supplemented or not with 0.2% of D-serine. Different combinations were carried out with the CD representative strain 204 and A1, R2 and T1 as representative strains of the HI group (**Figure 7**). First, the CD strain 204 displayed a better fitness than HI strains in M9 + Glucose with competitive indices ranging from 5 to 44. Addition of D-serine in the medium strongly improved the fitness of 204 over HI strains with competitive indices that reached 110 to 3700. These data demonstrate that the CD strain 204 has a competitive advantage over HI strains A1, R2 and T1 in our experimental conditions, and that catabolism of D-serine contributes to such efficient fitness.

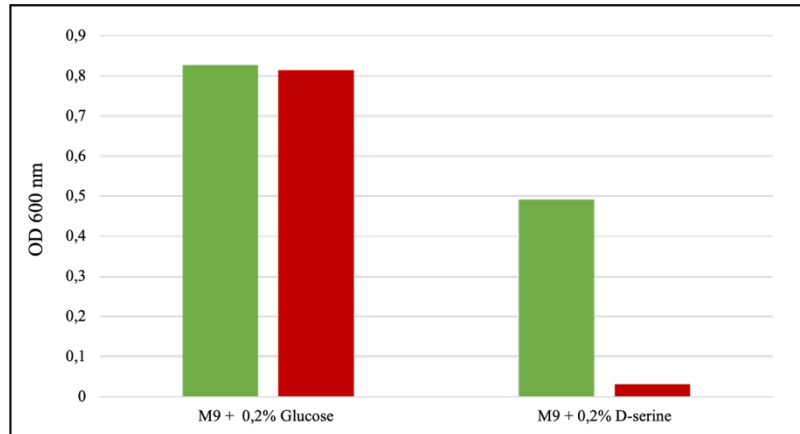


Figure 6 Growth of the WT and Δdsd mutant 204 strains. Growth was performed in M9 minimal media supplemented with 0.2% of glucose or 0.2% of D-serine as the unique carbon source. The optical density at 600 nm (OD_{600nm}) was measured after 30 h of culture for the WT strain (green bars) and the Δdsd mutant (red bars).

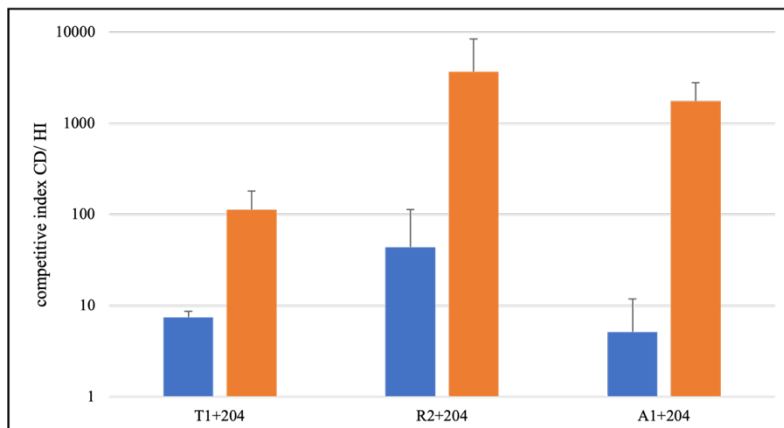


Figure 7 In vitro competition assays between HI and CD strains. Competitions were carried out in M9 minimal media supplemented with 0.02% of glucose (blue bars) or 0.02% of glucose and 0.2% of D-serine (orange bars). Three different combinations were tested between the CD representative strain 204 and HI representative strains T1, R2, and A1. Results are presented as the mean and standard deviation of competitive indices (ratio CD strain / HI strain) from 3 independent replicates.

To test if utilization of D-serine also confers a competitive advantage to CD strains over HI strains *in vivo*, competitions were carried out in germ free (GF) rats. An equal mix of CD strain 204 and HI strain A1 was given to GF rats by oral gavage and one group of inoculated rats was treated with drinking water supplemented with D-serine. Bacterial numerations from fecal samples of rats from both groups were performed during 15 days post-inoculation (**Figure 8**). No significant differences were observed between the groups, indicating that D-serine supplementation does not affect the fitness of both strains in the gut of gnotobiotic rats. More experiments should be scheduled to evaluate the role of D-serine catabolism in the fitness of CD strains *in vivo*. First, we need to validate that D-serine supplementation in drinking water artificially increases D-serine concentration in the gut through D-serine quantification assays. Competition assays in conventional rodents would be also of interest if we consider that nutritional competition between strains occurred more intensively in presence of a complex gut microbiota. It is also important to note that L-serine was shown to give a competitive advantage to *E. coli* only in case of gut inflammation [17]. Considering that the gut is inflamed in CD patients, it would be interesting as well to explore if D-serine catabolism could provide an advantage in inflammatory conditions.

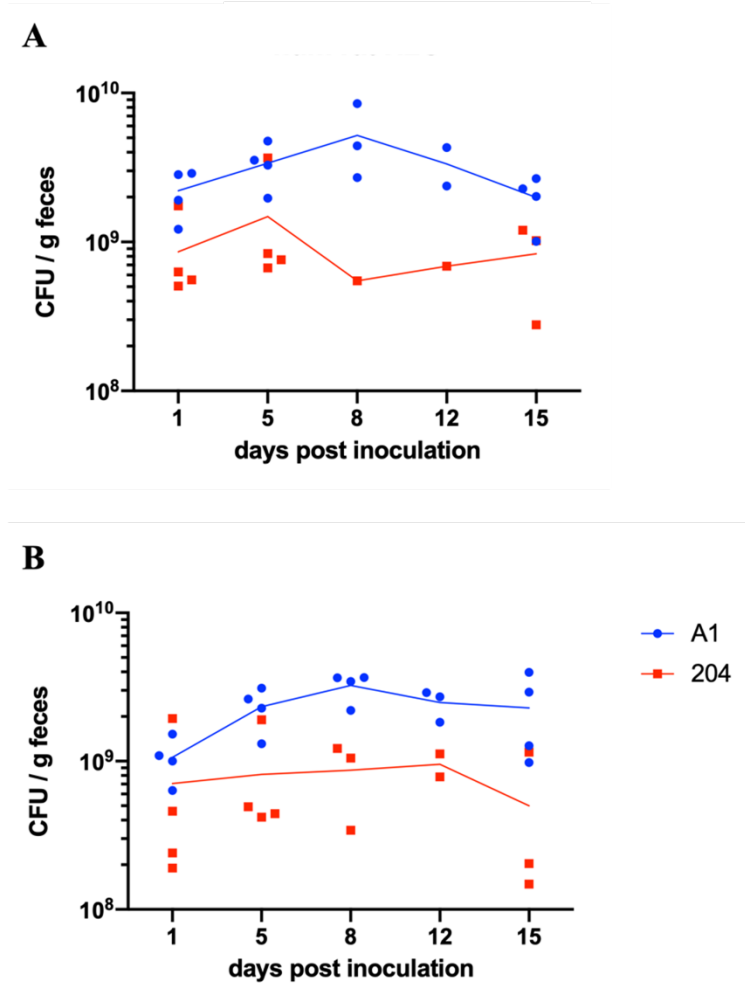


Figure 8 *In vivo* competition between HI strain A1 and CD strain 204 in the rat gut. Germ-free rats were co-inoculated with an equal mixture of A1 and 204 strains. Animals were given drinking water (A) or drinking water supplemented with D-serine 0.2% (B). At different time points, feces were sampled and spotted on LB or LB + Amp to count respectively the total bacteria and A1 strain alone. The 204 CFU were obtained by subtracting the A1 CFU from the total CFU. Each dot represents one rat, and curves represent mean value CFU/g of feces.

REFERENCES

- [1] G. De Palma et al., “Transplantation of fecal microbiota from patients with irritable bowel syndrome alters gut function and behavior in recipient mice,” *Sci. Transl. Med.*, vol. 9, no. 379, Mar. 2017, doi: 10.1126/SCITRANSLMED.AAF6397.
- [2] A. K. Degruittola, D. Low, A. Mizoguchi, and E. Mizoguchi, “Current Understanding of Dysbiosis in Disease in Human and Animal Models,” *Inflamm. Bowel Dis.*, vol. 22, no. 5, pp. 1137–1150, May 2016, doi: 10.1097/MIB.0000000000000750.
- [3] N.-R. Shin, T. W. Whon, and J.-W. Bae, “Proteobacteria: microbial signature of dysbiosis in gut microbiota,” *Trends Biotechnol.*, vol. 33, no. 9, pp. 496–503, Sep. 2015, doi: 10.1016/j.tibtech.2015.06.011.
- [4] G. Rizzatti, L. R. Lopetuso, G. Gibiino, C. Binda, and A. Gasbarrini, “Proteobacteria: A common factor in human diseases,” *Biomed Res. Int.*, vol. 2017, 2017, doi: 10.1155/2017/9351507.
- [5] S. E. Winter and A. J. Bäumlér, “Why related bacterial species bloom simultaneously in the gut: Principles underlying the ‘Like will to like’ concept,” *Cell. Microbiol.*, vol. 16, no. 2, p. 179, Feb. 2014, doi: 10.1111/CMI.12245.
- [6] M. Wouthuyzen-Bakker, F. A. J. A. Bodewes, and H. J. Verkade, “Persistent fat malabsorption in cystic fibrosis; lessons from patients and mice,” *J. Cyst. Fibros.*, vol. 10, no. 3, pp. 150–158, May 2011, doi: 10.1016/J.JCF.2011.03.008.
- [7] P. H. Bradley and K. S. Pollard, “Proteobacteria explain significant functional variability in the human gut microbiome,” *Microbiome*, vol. 5, no. 1, 2017, doi: 10.1186/S40168-017-0244-Z.
- [8] M. I. Gul’neva and S. M. Noskov, “[Colonic microbial biocenosis in rheumatoid arthritis],” *Klin. Med. (Mosk.)*, vol. 89, no. 4, pp. 45–48, 2011, Accessed: Oct. 20, 2022. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/21932563/>.

- [9] X. Wu et al., “Molecular Insight into Gut Microbiota and Rheumatoid Arthritis,” *Int. J. Mol. Sci.*, vol. 17, no. 3, Mar. 2016, doi: 10.3390/IJMS17030431.
- [10] M. Baumgart et al., “Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn’s disease involving the ileum,” *ISME J.*, vol. 1, no. 5, pp. 403–418, Sep. 2007, doi: 10.1038/ISMEJ.2007.52.
- [11] D. C. Baumgart and W. J. Sandborn, “Inflammatory bowel disease: clinical aspects and established and evolving therapies,” *Lancet (London, England)*, vol. 369, no. 9573, pp. 1641–1657, May 2007, doi: 10.1016/S0140-6736(07)60751-X.
- [12] F. H. Epstein and E. D. Harris, “Rheumatoid arthritis. Pathophysiology and implications for therapy,” *N. Engl. J. Med.*, vol. 322, no. 18, pp. 1277–1289, May 1990, doi: 10.1056/NEJM199005033221805.
- [13] K. Jahreis et al., “Adaptation of Sucrose Metabolism in the *Escherichia coli* Wild-Type Strain EC3132†,” *J. Bacteriol.*, vol. 184, no. 19, p. 5307, Oct. 2002, doi: 10.1128/JB.184.19.5307-5316.2002.
- [14] K. L. Nielsen, P. A. Godfrey, M. Stegger, P. S. Andersen, M. Feldgarden, and N. Frimodt-Møller, “Selection of unique *Escherichia coli* clones by random amplified polymorphic DNA (RAPD): Evaluation by whole genome sequencing,” *J. Microbiol. Methods*, vol. 103, pp. 101–103, 2014, doi: 10.1016/J.MIMET.2014.05.018.
- [15] K. A. Datsenko and B. L. Wanner, “One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 97, no. 12, p. 6640, Jun. 2000, doi: 10.1073/PNAS.120163297.
- [16] M. RL and W. RA, “The *Escherichia coli* *argW*-*dsdCXA* genetic island is highly variable, and *E. coli* K1 strains commonly possess two copies of *dsdCXA*,” *J. Clin. Microbiol.*, vol. 44, no. 11, pp. 4038–4048, Nov. 2006, doi: 10.1128/JCM.01172-06.

- [17] S. Kitamoto et al., “Dietary l-serine confers a competitive fitness advantage to Enterobacteriaceae in the inflamed gut,” *Nat. Microbiol.*, vol. 5, no. 1, pp. 116–125, Jan. 2020, doi: 10.1038/s41564-019-0591-6.
- [18] N. O’Boyle, J. P. R. Connolly, N. P. Tucker, and A. J. Roe, “Genomic plasticity of pathogenic *Escherichia coli* mediates d-serine tolerance via multiple adaptive mechanisms,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 36, p. 22484, Sep. 2020, doi: 10.1073/PNAS.2004977117.
- [19] C. JP and R. AJ, “When and where? Pathogenic *Escherichia coli* differentially sense host D-serine using a universal transporter system to monitor their environment,” *Microb. cell (Graz, Austria)*, vol. 3, no. 4, pp. 181–184, Apr. 2016, doi: 10.15698/MIC2016.04.494.
- [20] S. K, E. R, A. J, S. R, and L. JW, “Plasmid-mediated sucrose metabolism in *Escherichia coli* K12: mapping of the scr genes of pUR400,” *Mol. Microbiol.*, vol. 2, no. 1, pp. 1–8, 1988, doi: 10.1111/J.1365-2958.1988.TB00001.X.
- [21] W. F. Fricke et al., “Insights into the environmental resistance gene pool from the genome sequence of the multidrug-resistant environmental isolate *Escherichia coli* SMS-3-5,” *J. Bacteriol.*, vol. 190, no. 20, pp. 6779–6794, Oct. 2008, doi: 10.1128/JB.00661-08.
- [22] B. Hochhut, K. Jahreis, J. W. Lengeler, and K. Schmid, “CTnscr94, a conjugative transposon found in enterobacteria,” *J. Bacteriol.*, vol. 179, no. 7, pp. 2097–2102, 1997, doi: 10.1128/JB.179.7.2097-2102.1997.

SUPPLEMENTARY DATA

Table S1. Bacterial strains and plasmids used in this study

Strains or plasmids	Description	References
<i>Bacterial strains</i>		
A1, C2, D2, F3, I1, L3, M2, N3, P1, P3, R2, T1, U1, V3, Y1, AA1	<i>E. coli</i> strains isolated from fecal samples of healthy individuals	PMID: 35988812
204, 205, 206, 208, 209, 213, 215, 216, 802, 803, 804, 805, 806, 807, 808	<i>E. coli</i> commensal strain isolated from patients suffering with Crohn's disease	This study
W1, W4, W5, X4, AQ1, AQ3, AR1, AU1, AS2, AV1, AW1, AW2, AW3, AX1	<i>E. coli</i> commensal strain isolated from patients suffering with rheumatoid arthritis	This study
204 Δ <i>dsd</i>	204 strain with a deletion of genes <i>dsdC</i> , <i>dsdX</i> and <i>dsdA</i> , Kan ^R	This study
<i>Plasmids</i>		
pKD4	<i>kan</i> cassette template, Amp ^R , Kan ^R	10829079

Table S2. Primers used in this study

Purpose	Name	Sequence 5'-3'	References
16S	8F	AGAGTTTGATCCTGGCTCAG	2798131 Lane, D.J. (1991) 16S/23S rRNA Sequencing. In: Stackebrandt, E. and Goodfellow, M., Eds., Nucleic Acid Techniques in Bacterial Systematic, John Wiley and Sons, New York, 115-175.
	1492R	CGGTTACCTTGTACGACTT	
RAPD	1283	GCGATCCCCA	24912108
Detection of <i>dsd</i> and <i>csc</i> locus	<i>dsdA</i> _R		this study
	<i>cscR</i> _F		this study
	<i>dsdX</i> _F		this study
204 Δ <i>dsd</i> mutant	<i>dsdmut</i> _F	CTGCCTGCCCCGTAATGTGGTGATGTAATAATGCTGATGGCATTTCAC CGGTGTAGGCTGGAGCTGCTTC	this study
	<i>dsdmut</i> _R	CGGCACCATTCCACCTCCCCTCGCCACACCAGATGAGTGGCATTACG CACATATGAATATCCTCCTTAGTTCC	this study
	<i>yfdC</i> _F	ACGACATGAGCAACAAGCGC	this study
	<i>emrY</i> _R	GTTTACAGCGAAAGGCGTTGGG	this study

PART 2

The unexplored potential of Ethanolamine utilization by commensal *E. coli* strains

In the first part of my work, I evaluated the use of 190 different substrates as a carbon source to highlight potential molecules that could be used by *Enterobacteriaceae* as a competitive advantage in the dysbiotic gut. These 190 substrates, corresponding mainly to usual carbohydrates and amino acids, are only a small fraction of molecules that could be encountered in the GIT. Based on the literature and works from our laboratory, a compound that was not represented in the OmniLog plates but has a major interest in a gut environment is ethanolamine (EA). This small molecule ($\text{H}_2\text{N-CH}_2\text{-CH}_2\text{-OH}$), composed of a primary amine and a primary alcohol, is a constituent of phosphatidylethanolamine, a major component of every cell membrane lipid [296]. Thanks to the turnover of bacterial and epithelial cells, EA is naturally present in the GIT.

Interestingly, numerous studies determined that EA is a preferential substrate for several intestinal pathogens such as pathogenic *E. coli*, *Listeria monocytogenes* or *Salmonella enterica* serovar Typhimurium [292]. These pathogens take nutritional advantage of EA in the gut to proliferate and/or as a signalling molecule to control expression of virulence genes [297]. Although EA has been most studied for its key role during host-gut pathogens interactions, scarce studies suggest that EA consumption is not specifically associated to virulence. For example, a recent study has demonstrated that two human commensal *E. coli* strains, HS and Nissle, are able to consume EA as a nitrogen source during *in vitro* growth, and even more efficiently than a strain of enterohemorrhagic *E. coli* (EHEC) O157:H7 (Figure 18) [298].

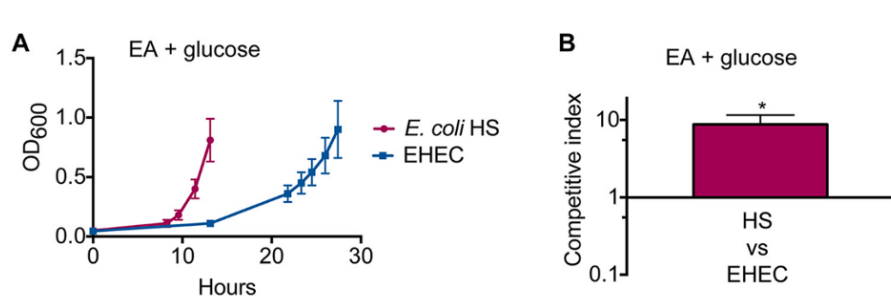


Figure 18 HS outcompetes EHEC specifically during growth on EA. (A) Growth curve of *E. coli* HS and EHEC in minimal medium with EA and glucose. (B) Competition assay between *E. coli* HS and EHEC in minimal medium with EA and glucose. Adapted from Rowley *et al.*, 2018.

In this second part of my PhD work, I have decided to focus on EA catabolism. First, I evaluated the ability of the 3 groups of *E. coli* isolates (HI, CD and RA groups) used in part 1 to consume EA. Growth assays were performed in minimal medium supplemented with EA as a sole nitrogen or carbon source (Figure 19).

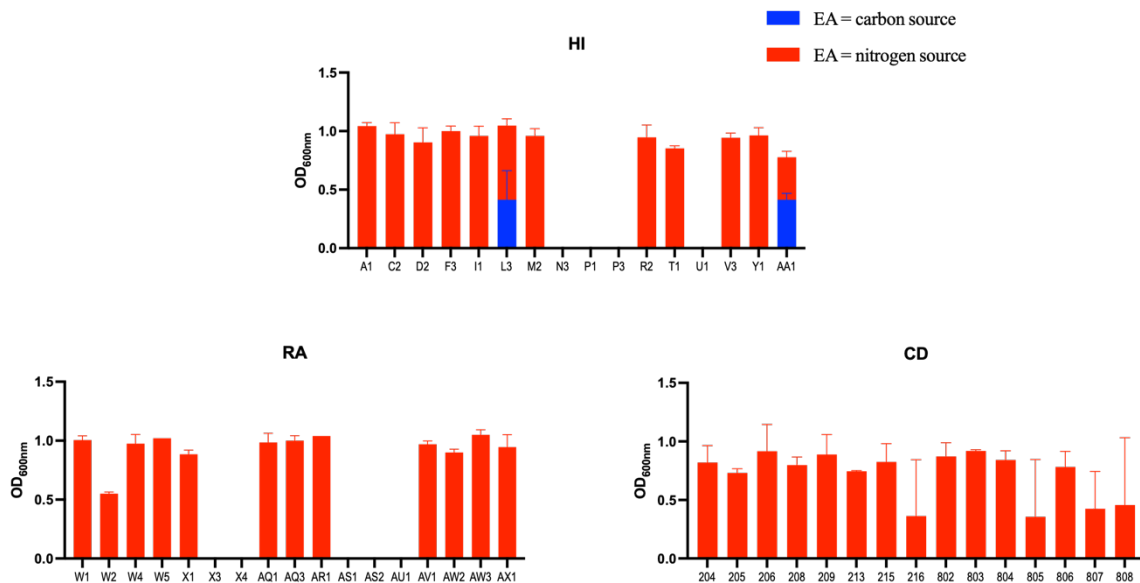


Figure 19 EA utilization profiles of commensal *E. coli* strains from HI, CD and RA collection. *In vitro* growth was evaluated in M9 minimal medium supplemented with EA as a sole nitrogen or carbon source. Results are presented as mean and standard deviations from two to four independent experiments. OD600nm was measured after 24 h of incubation for cultures in M9 supplemented with glucose and EA (red) and after 72 h for cultures in M9 supplemented with NH₄Cl and EA (blue).

We observed that most strains consume EA as a nitrogen source whatever the strain groups, indicating that EA catabolism is indeed prevalent in commensal *E. coli*. Additionally, two strains from the HI collection (L3 and AA1) were also able to use EA as a sole carbon source. This phenotype was absent among our CD and RA collection of strains. Some strains of HI and RA groups were also not able at all to use EA as a nutrient, demonstrating variability in EA utilization among commensal *E. coli*.

Taken together, these preliminary results led me to investigate the variability of EA metabolism among commensal *E. coli*. Because the highest variability was observed in the HI group and because CD and RA strains were isolated from patients with a possible bias through strain selection/adaptation in the dysbiotic gut environment, we chose to pursue further investigations only with HI strains. For this purpose, the HI strain collection was expanded to a total of 40 isolates. We first investigated the phenotypic diversity of EA utilization among our collection of HI strains. Then, the main goal was to explain the observed phenotypic diversity

in order to identify the reason that differentiates the profiles of EA utilization. This work led to a publication (**manuscript N°2**), which was accepted for publication in the peer reviewed journal Research in Microbiology on August 2022. In parallel to this work, we also developed a methodological study to design a new reporter system to monitor bacterial gene expression in complex conditions such as the gut ecosystem. Taking advantage of the well-known regulation processes involved in the transcription of EA catabolism associated genes in *E. coli*, we designed a new plasmid-based tool ensuring both the identification of a strain of interest in complex environments and the monitoring of gene expression through the combination of two distinct fluorescent proteins as reporter genes. The tool was evaluated for its ability to report the sensing of ethanolamine concentration by *E. coli in vitro* but also *in vivo* in the mouse gut. This work led to a publication (**manuscript N°3**) which has been submitted to the journal Applied Microbiology and Biotechnology.

Before presenting the manuscripts, I introduce here some background information related to the EA catabolic pathway and associated genes. The ability to catabolize EA is encoded by the ethanolamine utilization (*eut*) genes organized in an operon. Tsoy *et al* (2009) focused on the comparative genomics of EA utilization and discovered that almost 100 fully sequenced bacterial genomes contain *eut* operons [299]. These authors also determined the content and organizational differences among of bacteria that contain these genes. Among the different taxonomic groups, they showed that Actinobacteria and most Proteobacteria have short *eut* operons. In contrast, members of the proteobacterial family *Enterobacteriaceae*, which includes *S. typhimurium* and *E. coli*, have long *eut* operons, as do members of the phylum Firmicutes. In *E. coli*, this locus is composed of 17 *eut* genes coding not only for key enzymes of the catabolism of EA but also proteins that form the microcompartments where EA catabolism take place (**Figure 20**) [300]. The *eut* locus encodes the transcription factor EutR that senses EA and the cofactor vitamin B₁₂ to directly activate *eut* transcription [301],[302]. EA can diffuse freely across membranes but when pH is low, EutH facilitates diffusion across the cell envelope [303]. The central genes in the transformation of EA are *eutB* and *eutC*, the protein products of which form the ethanolamine ammonia lyase. The ethanolamine ammonia lyase EutBC breaks down EA into acetaldehyde and ammonia, a valuable nitrogen source of energy. Acetaldehyde is then converted into ethanol by EutG or into acetyl-CoA by EutE, which can be used in numerous cell processes (Krebs cycle, glyoxylate bypass, lipid biosynthesis, or other processes). Acetyl-CoA can also be converted into acetate (**Figure 20**) [304].

The *eut* operon also encodes structural proteins (EutS, EutM, EutK, EutL, and EutN) that self-assemble into bacterial microcompartments (BMCs) forming a selectively permeable, icosahedral protein shell-like structure [305]. All shell proteins except EutK form hexagonal units like the previously elucidated shell proteins of the carboxysome⁶ [306]. It is important to note that *eut* loci do not systematically encode the BMC structural proteins but for strains that do, these structures bring different advantages to EA catabolism. Firstly, the toxic intermediate acetaldehyde is kept away from the cytoplasm. Secondly, acetaldehyde is a volatile compound that is retained in BMCs rather than lost in the cytoplasm. Finally, acetaldehyde and other intermediates are concentrated inside these structures, increasing the efficiency of the downstream reactions [307],[308].

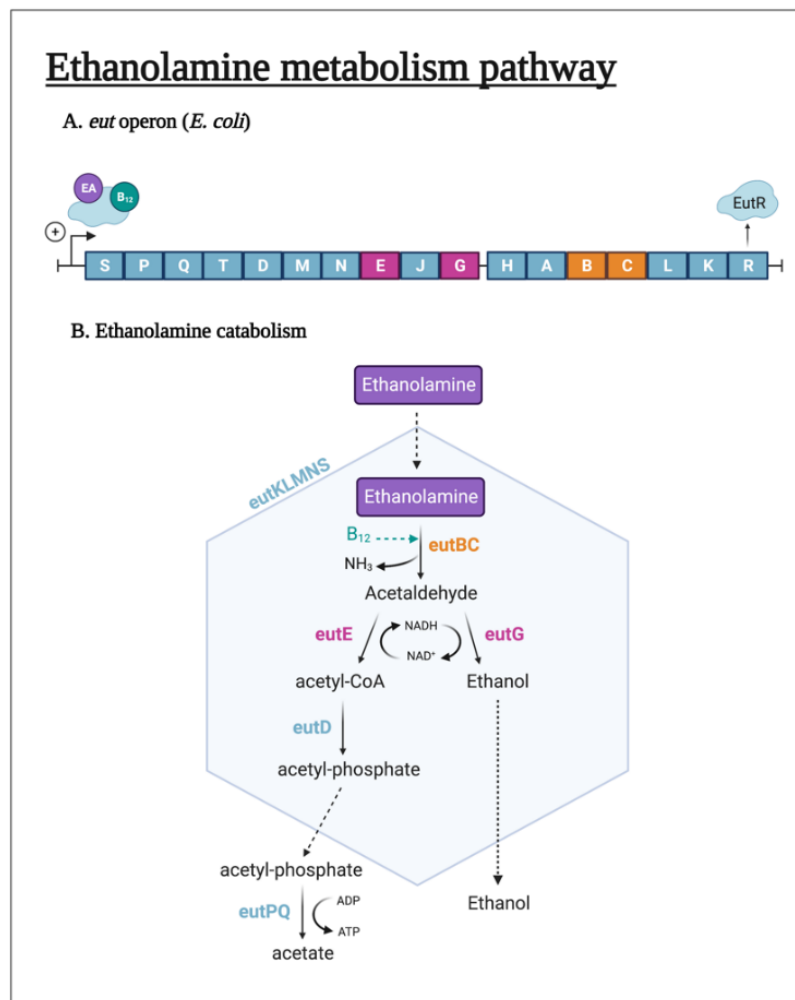


Figure 20 EA catabolism and organisation of the *eut* operon in *E. coli*. Schematic representation of (A) the organisation of the *eut* operon in *E. coli* and (B) the ethanolamine catabolism in the microcompartment structure. Figure made with biorender.com.

⁶ Carboxysome : bacterial organelle containing enzymes involved in carbon fixation.

MANUSCRIPT 2

Diversity of ethanolamine utilization by human commensal *Escherichia coli*

Maria Ines Moreira de Gouveia^a, Julien Daniel^a, Annie Garrivier^a, Annick Bernalier-Donadille^a, Gregory Jubelin^{a*}

^a Université Clermont Auvergne, INRAE, MEDIS UMR454, F-63000 Clermont-Ferrand, France

maria-ines.moreira-de-gouveia@inrae.fr
julien.daniel@inrae.fr
annie.garrivier@inrae.fr
annick.bernalier@inrae.fr
gregory.jubelin@inrae.fr *Correspondence and reprints

ABSTRACT

Ethanolamine (EA) is a substrate naturally present in the human gut and its catabolism by bacteria relies on the presence of *eut* genes encoding specific metabolic enzymes and accessory proteins. To date, EA utilization has been mostly investigated in gut bacterial pathogens. The aim of this study was to evaluate the ability of human gut commensal *Escherichia coli* isolates to utilize EA as a nitrogen and/or carbon sources. Although the capacity to consume EA is heterogeneous between the 40 strains of our collection, we determined that most of them could degrade EA to generate ammonia, a useful nitrogen resource for growth. Three isolates were also able to exploit EA as a carbon source. We also revealed that the inability of some strains to catabolize EA is explained either by mutations in the *eut* locus or by a defect in gene transcription. Finally, we demonstrated the importance of EA utilization for an optimal fitness of commensal *E. coli in vivo*. Our study provides new insights on the diversity of commensal *E. coli* strains to utilize EA as a nutrient in the gut and opens the way for new research in the field of interactions between host, gut microbiota and pathogens.

KEYWORDS :

Ethanolamine; *eut* operon; commensal *E. coli*; human gut; nutritional competition

INTRODUCTION

Ethanolamine (EA) is a small organic molecule found in most biological membranes since it is a precursor of the common phospholipid phosphatidylethanolamine. Even if its abundance has never been precisely quantified in the gastro-intestinal tract of the healthy human, free EA is present in the gut lumen, deriving either from the diet or from the constant renewal of intestinal epithelial cells [1]. Indeed, enterocyte turnover leads to the release of numerous compounds in the gut lumen that can be used as nutrient sources by intestinal bacteria. Several works have demonstrated that gut pathogens such as *Salmonella enterica*, enterohaemorrhagic *E. coli* (EHEC) or *Listeria monocytogenes*, consume EA and gain a competitive advantage for colonization over the commensal microbiota. EA also constitutes a signal molecule which modulates the expression of virulence factors and affects the pathogen's behavior for cell adhesion or macrophage survival for example [2,3]. EA catabolism requires the presence of the ethanolamine utilization (*eut*) genes which have been described in several bacterial families [4]. *E. coli* and other bacteria belonging to the *Enterobacteriaceae* family hold a long *eut* operon with 17 genes [5]. The *eut* operon includes genes encoding key enzymes involved in EA degradation but also genes encoding structural proteins required for the formation of metabolosome, a structural microcompartment where EA catabolism takes place within the bacterial cell. The first step of the EA metabolization pathway is the release of NH₃ from EA by the ethanolamine ammonia lyase EutBC, generating an extremely valuable nitrogen source. In addition to NH₃, EutBC produces acetaldehyde which is next converted by the dehydrogenase EutE into acetyl-CoA. This molecule can be employed in multiple metabolic reactions including TCA cycle or lipid biosynthesis, and therefore enables the assimilation of EA carbon atoms by bacteria. The last gene of the *eut* operon is *eutR* that encodes a transcriptional regulator which activates the transcription of all *eut* genes in presence of both EA and adenosylcobalamin [4].

Although EA has been most studied for its key role during host-gut pathogens interactions, this resource can also be metabolized by commensal strains as initially demonstrated several decades ago [6]. A comparative genomic study analyzed the *eut* genes in 100 sequenced bacterial genomes and almost all of the species containing long *eut* operons are facultative anaerobes and live in the gut as commensals or pathogens. Furthermore, the study shows no particular link between the presence of the *eut* operon and pathogenicity of bacteria, suggesting that EA consumption is not specifically associated to virulence [5]. In a recent study, two human commensal *E. coli* strains, HS and Nissle, have been shown to consume EA as a nitrogen source during in vitro growth. In addition, both commensal strains were able to outcompete a pathogenic *E. coli* strain when EA was provided as a sole nitrogen source [7]. The aim of our study was to investigate whether EA catabolism is rare or common in human commensal *E. coli*. Among the 40 tested strains, 32 can use EA as a nitrogen source and 3 were able to grow with EA provided as a unique carbon source. In addition, we determined that EA utilization deficiency observed for 5 commensal *E. coli* strains was caused either by DNA insertion within the *eut* operon or by a defect in the transcriptional regulation process. Finally, we demonstrated that inactivation of the key *eutB* gene decreases the fitness of commensal *E. coli* in the mouse gut. Taken together, our data reveal that EA catabolism is not specific to bacterial pathogens and occurs commonly in commensal *E. coli* despite a high degree of variability between isolates.

MATERIALS AND METHODS

Bacterial strains and plasmids

Strains and plasmids used in this study are listed in [Table S1](#). Our collection of commensal *E. coli* is composed of 40 strains that were isolated from 23 healthy individuals who gave their written informed consent to the study (protocol ID : 216-A00356-45, approved by committee for the protection of persons (CPP) Sud Est VI France). Stool samples were serially 10-fold diluted to 10⁻¹⁰ and spread on MacConkey and Hektoen agar plates. A RAPD-PCR [8] was performed on a minimum of 10 colonies for each sample in order to detect and exclude clonal isolates from our collection. Taxonomic identification of each strain was determined by amplification and sequencing of the rDNA 16S and only *Escherichia* strains were conserved in the collection. Overall, one to three different *E. coli* strains were isolated from each healthy individual.

Culture media and growth conditions

E. coli strains were cultured in LB or in a modified M9 minimal salt medium containing Na₂HPO₄ (42 mM), KH₂PO₄ (22 mM), NaCl (8.5 mM), MgSO₄ (2 mM) and CaCl₂ (0.1 mM). Complete M9 medium was obtained by addition of glucose (0.2%) and NH₄Cl (18.5 mM) as carbon and nitrogen source, respectively, unless otherwise indicated. Utilization of EA as a unique nitrogen source in M9 medium was investigated by addition of EA hydrochloride (30 mM), vitamin B₁₂ (150 nM) and glucose (0.2 %). Utilization of EA as a unique carbon source was investigated by addition of EA hydrochloride (30 mM), vitamin B₁₂ (150 nM) and NH₄Cl (18.5 mM). Bacterial strains were first isolated on LB plates and LB precultures inoculated from a single colony were performed for 8 h at 37 °C with aeration. The precultures were then diluted 100-fold in complete M9 and grown overnight. Precultures were washed once with M9 and used to inoculate cultures in M9 media. Cultures were incubated with shaking at 37 °C and

growth was evaluated by measuring OD_{600nm} at different time points. When required, antibiotics were used at the following concentrations: kanamycin (25 µg.ml⁻¹); chloramphenicol (25 µg.ml⁻¹); streptomycin (50 µg.ml⁻¹); gentamicin (15 µg.ml⁻¹).

Construction of *eut* mutants

Replacement of *eutB*, *eutE* or *eutG* genes by the gene conferring resistance to kanamycin in strain L3 were obtained by using the one-step PCR-based method of Datsenko and Wanner [9]. Designed primers used to construct the mutants are presented in [Table S2](#). Mutations were confirmed by PCR and sequencing using specific primers for the kanamycin gene and the gene flanking the mutated gene.

DNA extraction and genome sequencing

Genomic DNA was extracted using the DNeasy Bood Tissue kit (Qiagen) and then sequenced by HELIXIO company (Clermont-Ferrand, France). Briefly, DNA-seq libraries were prepared using NEBNext Ultra™ II FS DNA Library Prep Kit for Illumina (New England Biolabs). Quality of libraries was assessed using the Agilent Bioanalyzer. DNA-seq was performed using the NextSeq500 system (Illumina) according to the manufacturer's instructions. Reads were checked for quality using FastQC and then aligned along the genome of *E. coli* HS as a reference. Genome sequences have been deposited in the NCBI database and are available under the accession number PRJNA862305.

RNA extraction and RT-qPCR

Bacterial cultures were grown at 37 °C for 8 h and 10 ml were centrifuged (12,000 rpm for 2 minutes) to collect bacteria. RNA was extracted using standard TRIzol/Chloroform procedure. After treating each RNA sample with TURBO™ DNase (Invitrogen), 1 µg of RNA was reverse

transcribed using random primers and Superscript II reverse transcriptase (Invitrogen). Real-time PCR runs were carried out using the QuantiFast SYBR Green PCR kit (Qiagen) following the manufacturer's instructions and using 2 μ L of cDNA and primers specific to *eutS*, *eutB*, *eutE*, *eutG* or *rpoA* as the housekeeping gene. Standard curves for each gene were obtained by using genomic DNA of *E. coli* HS and results are presented as expression levels in absolute numbers of gene copies per sample.

Metabolite quantification

From bacterial cultures, samples were centrifuged and 0.2 μ m-filtrated supernatants were used for metabolite quantification by 1D 1H NMR (MetaToul, Toulouse, France). Briefly, a mix of 540 μ L of sample + 60 μ l of TSP-d4 (Sodium-2,2,3,3-trimethylsilylpropionate D4) at 10.724mM was done in D2O and acquisition of NMR spectra was performed on an Avance III 800MHz Bruker equipped with a CQPCI 5mm cryoprobe (1H, 13C, 31P, 15N). Phase and baseline corrections were performed using automatic tools from TopSpin 3.6.0 software before the manual integration of specific signals belonging to exo-metabolites present in the samples. Absolute quantitation of metabolites of interest was performed using TSP-d4 as a reference external standard.

Mouse experiments

C3H/HeOuj with specific-pathogen-free (SPF) status were purchased from Charles River laboratories. Five-to six-week-old mice were housed in cages containing no more than five animals, maintained under a 12 h light/dark cycle at a temperature of 21 ± 2 °C, and fed with standard diet and water ad libitum. The experiments performed herein were reviewed and approved by the Auvergne Committee for Animal Experimentation C2EA and received the agreement number #31333-2021042809157789 v2. For co-inoculation experiments, drinking

water containing 5g/L streptomycin sulfate was given throughout the experiment, starting from one day prior to bacterial inoculation to selectively remove facultative anaerobic bacteria and open the niche for *E. coli*. For mouse inoculation, WT, Δ *eutB* and Δ *eutE* L3 strains were transformed with a derivative of pMS vector (Life technologies) conferring resistance to streptomycin in order to allow their monitoring from fecal samples using selective plates. Mice were intragastrically co-inoculated at day 0 with 100 μ l of PBS containing 10^7 each of the WT and the *eut* L3 strains (*eutB* or *eutE*). The bacteria were enumerated in feces at day 7 post-inoculation. Fecal samples dilutions were plated on LB + Sm or LB + Sm + Km agar plates to count the WT + Δ *eut* strains and Δ *eut* alone, respectively. Competitive indices were calculated daily by dividing the output ratio (WT/mutant) by the corresponding input ratio (WT/mutant from mouse inoculum).

Statistical analyses

Statistical analyses were performed using the GraphPad Prism software version 8 (San Diego, USA). Each dataset was analyzed by the ROUT method to exclude outliers with a Q value of 1%. Statistical tests used for each data analysis are indicated in figure legends. $P < 0.05$ was considered significant.

RESULTS

Most commensal E.coli strains use EA as a nitrogen and/or carbon sources in vitro

For this study, 40 commensal *E. coli* strains were isolated from feces of healthy human individuals. The use of EA as a nitrogen (N) and/or as a carbon (C) source by our strain collection was evaluated by performing *in vitro* growth in minimal media (Fig. 1). The strain HS was used as a control since it has been previously described to use EA only as a N source [7]. Three groups of strains were identified as follows: 5 (12.5%) did not use EA either as a C or N sources (noted hereafter EA⁻ strains), 32 (80%) use EA only as a N source, similarly to HS (EA^{+(N)} strains) and 3 (7.5%) use EA both as a C and N sources (EA^{+(N/C)} strains). It is noteworthy that the utilization of EA as a carbon source requires a longer incubation time (48 to 72h) than utilization of EA as nitrogen source. To confirm EA utilization as a carbon source by strains AA1, B4 and L3, EA was quantified by NMR after bacterial culture in minimal media (Fig. S1). As expected, the concentration of EA decreased after the growth of the 3 strains and led to the accumulation of acetate and ethanol, two products of EA catabolism [4].

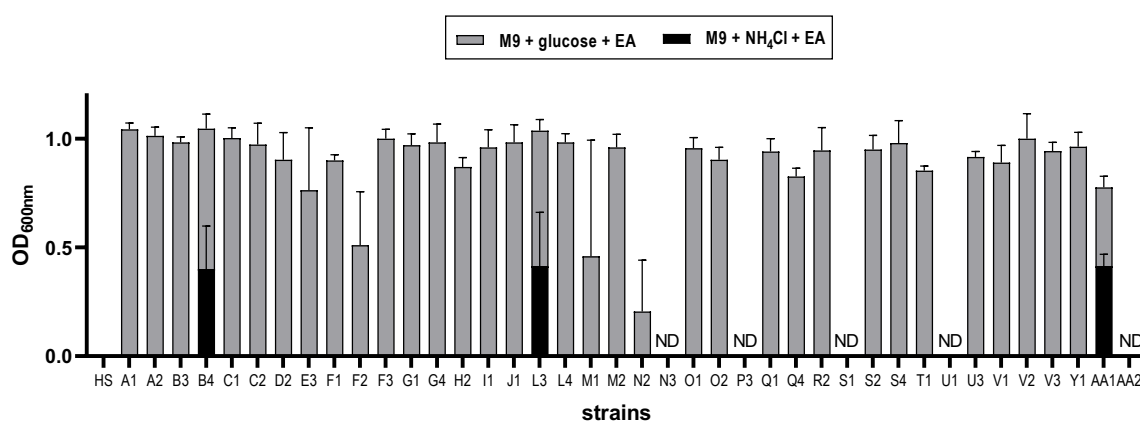


Figure 1 EA utilization profiles of commensal *E. coli* strains. In vitro growth of 40 commensal *E. coli* strains was evaluated in M9 minimal medium supplemented with EA as a sole nitrogen or carbon source. Results are presented as mean and standard deviations from three independent experiments. OD_{600nm} was measured after 24 h of incubation for cultures in M9 supplemented with glucose and EA (grey bars) and after 72 h for cultures in M9 supplemented with NH₄Cl and EA (black bars). ND: not detected

EA catabolism in *E. coli* is allowed by the expression of the *eut* operon and the different enzymatic steps are summarized in [fig S2](#). To confirm the involvement of key Eut enzymes in the use of EA, we arbitrarily selected one of the EA^{+(N/C)} isolate (strain L3) to perform mutagenesis. Genes *eutB*, *eutE* or *eutG* were inactivated in L3 and growth ability of each mutant was compared to that of the WT strain ([Fig. 2](#)). As expected, the *eutB* mutant did not grow in minimal media supplemented with EA as a N or C source, confirming that the release of ammonia from EA is a critical first step in EA catabolism pathway. Next, acetaldehyde is transformed either into acetyl-CoA by EutE or into ethanol by EutG. In contrast to the *eutG* mutant, the *eutE* mutant was unable to grow using EA as a carbon source, suggesting that conversion to acetyl-CoA is essential to exploit carbons from EA as an energy source. Altogether, these results demonstrate that *eut*-dependent exploitation of EA is heterogeneous in commensal *E. coli* and can provide nitrogen and/or carbon resources for growth.

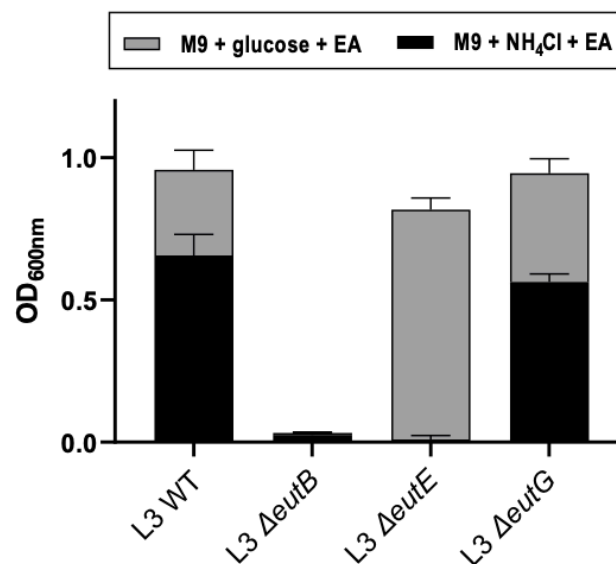


Figure 2 Importance of EutB, EutE and EutG enzymes for EA metabolism in EA^{+(N/C)} strain L3. Genes *eutB*, *eutE* or *eutG* were deleted and replaced by the gene conferring resistance to kanamycin. The different mutant's ability to use EA as a nitrogen or carbon source was evaluated in M9 medium. Results are presented as mean and standard deviations from three independent experiments. Growth was assessed by measuring OD_{600nm} after 24 h of incubation for cultures in M9 supplemented with glucose and EA (grey bars) and after 72 h for cultures in M9 supplemented with NH₄Cl and EA (black bars).

Most of EA⁻ strains present an altered *eut* locus

To explain the different phenotypes of EA utilization, we first sequenced the genome and analysed the *eut* cluster of the 5 EA⁻ strains, 7 of the 32 EA^{+(N)} strains and the 3 EA^{+(N/C)} strains (**Fig 3**). The *eut* operon consists of 17 genes from *eutS* to *eutR* and we used the HS one as a reference. Among the 5 strains unable to consume EA, 3 carry a DNA insertion within the *eut* operon. Strains N3 and U1 have a large insertion of respectively 10 or 12 kb, between *eutA* and *eutB*, corresponding to putative phage regions. The strain S1 has also a small insertion of 800 bp between *eutG* and *eutH*. An ORF analysis also showed that strain AA2 has premature stop codons within *eutH* and *eutR*, leading to presumably inactive truncated proteins. From this group of strains, only P3 presents an *eut* operon with an unaltered structure. In contrast, all EA^{+(N)} and EA^{+(N/C)} strains possess an intact *eut* cluster (**Fig 3**). We next investigated the overall identity of Eut proteins between strains. Using the protein sequences of strain HS as a reference, we observed that the percentages of protein identity between strains range from 92.7 to 100 % with similar variations between strains except for the EA⁻ strain S1, which present a higher level of punctual substitutions for most genes (**Fig S3**). Of note, the sequence of proteins involved in microcompartment structure or key enzymes such as the ammonia lyase EutBC are highly conserved between strains. We also noticed a 13 amino acid deletion in EutH (encoding EA transporter) in strain C2. Overall, these genetic analyses potentially explain why the EA⁻ strains of our collection cannot catabolize EA, except for the strain P3. However, it does not allow us to explain the differential ability of *E. coli* strains to catabolize EA as a N source only or as both N and C sources.

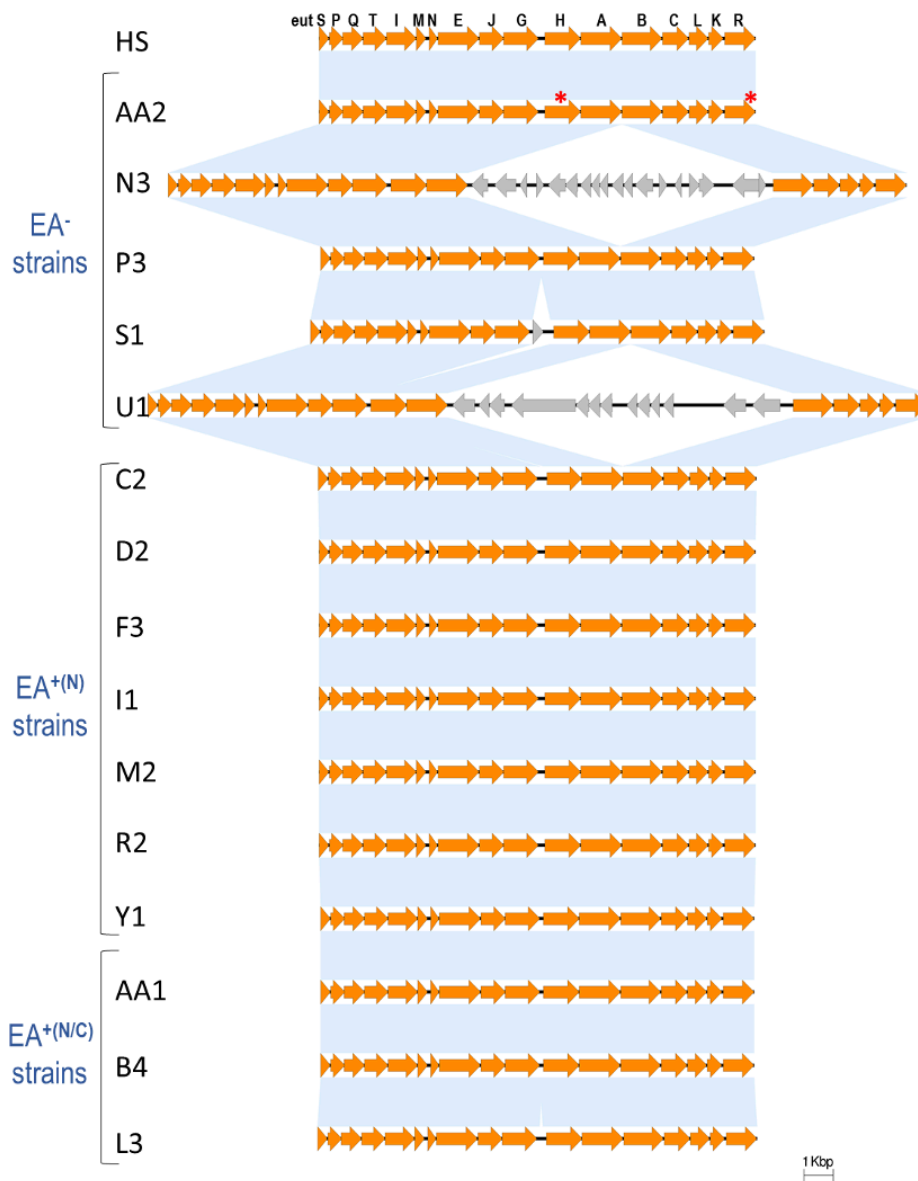


Figure 3 Schematic representation of the *eut* operon of *E. coli* strains according to their EA utilization phenotypes. The genome of 15 strains of different EA profiles was sequenced and the strain HS was used as a reference genome. The orange arrows represent *eut* genes with the same organization as the reference strain whilst grey arrows represent insertions of unrelated genes. The blue regions indicate high sequence conservation between strains in contrast to the white regions indicating low sequence conservation. The red stars indicate the presence of a premature STOP codon within the gene.

*EA-dependant upregulation of *eut* genes is inefficient in the EA⁻ strain P3*

We next evaluated the expression level of *eut* genes during growth in complete M9 medium (with NH₄Cl and glucose) supplemented or not with EA and vitamin B₁₂. Selected genes for qRT-PCR analysis were *eutS*, the first gene of the *eut* cluster, *eutB* encoding a subunit of the ammonia lyase, *eutE* and *eutG* encoding enzymes transforming acetaldehyde to acetyl-CoA and ethanol, respectively. As expected, the expression of *eut* genes was strongly induced in the presence of EA and vitamin B₁₂ for EA^{+(N)} strains (C2, F3) and EA^{+(N/C)} strains (L3, AA1) with fold changes ranging from 200 to 1024 for *eutS*, 42 to 416 for *eutB*, 347 to 2869 for *eutE* and 31 to 2555 for *eutG* (Fig 4). However, no particular differences were observed for the expression of *eutE* and *eutG* between EA^{+(N)} and EA^{+(N/C)} strains that could correlate with their EA utilization phenotype. In contrast, a very low level of expression was observed for all genes of the EA⁻ strain P3 whatever the presence or absence of EA. This data revealed that there is a transcription defect of the *eut* operon in this strain, explaining why this isolate is unable to use EA as a nutrient.

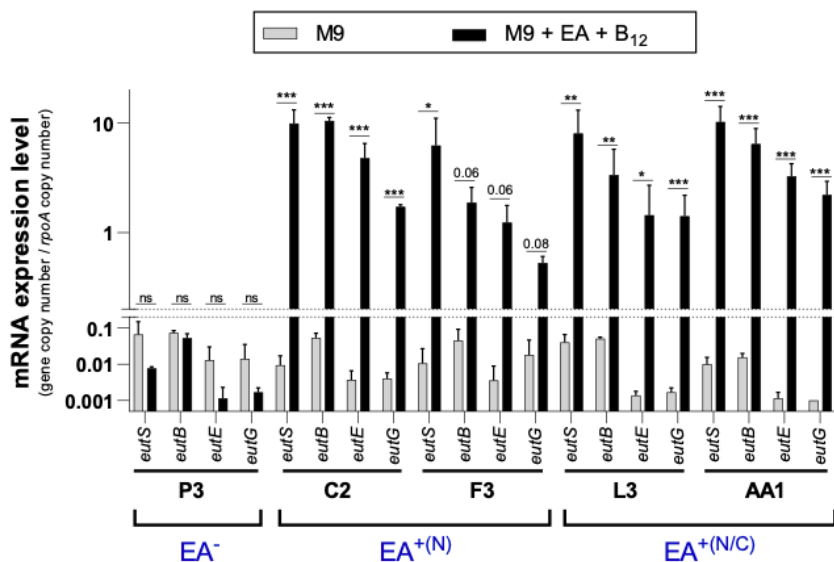


Figure 4 Expression levels of selected *eut* genes in presence of EA and vitamin B₁₂. Strains with different EA metabolization profiles were selected to assess expression levels of *eutS*, *eutB*, *eutE* and *eutG* genes. RNA was extracted after 8 h of growth in complete M9 medium (grey bars) or in complete M9 supplemented with EA and vitamin B₁₂ (black bars). Gene expression is presented as the ratio between indicated gene copy number and housekeeping gene *rpoA* copy number. Results are presented as mean and standard deviations from at least three independent experiments. An ordinary two-way ANOVA test was applied to compare gene expression level between uninduced and EA induced conditions and between strains. ns: non-significant; *p < 0.05; **p < 0.01; ***p < 0.001.

EA catabolism in the mouse gut

To further study EA catabolism in more physiological conditions, we carried out competition assays in the mouse gut. Mice were co-inoculated with the L3 WT strain and $\Delta eutB$ or $\Delta eutE$ mutant to evaluate the importance of these two genes in the fitness of an EA^{+(N/C)} strain *in vivo*. The bacteria were enumerated in feces using selective plates and a competitive index WT/mutant was calculated for each mouse (Fig 5). At day 7 post-inoculation, the competitive index of mice co-inoculated with WT and $\Delta eutB$ strains was superior to 1 and reached a mean value of 36, indicating that the mutant was outnumbered by the WT strain. In contrast, the competitive index calculated from mice co-inoculated with WT and $\Delta eutE$ strains stayed close to 1, showing that both strains have an equivalent fitness in the mouse gut. These data reveal that the consumption of EA by L3 occurs *in vivo* and is beneficial for the fitness of commensal *E. coli* in the gastro-intestinal of mice. Such advantage can be explained in our experimental conditions by the release of NH₃ by the ammonia lyase EutB but not by the production of acetylCoA by EutE even though L3 can use EA as a C source in our *in vitro* assays.

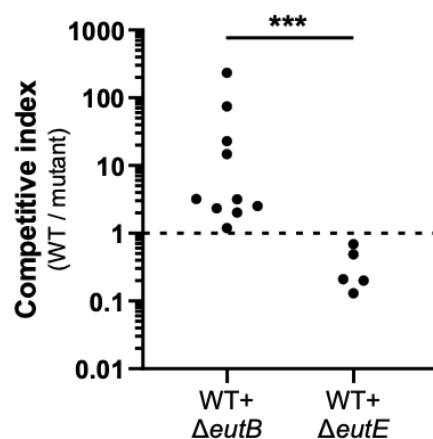


Figure 5 Role of *eutB* and *eutE* genes in the fitness of the EA^{+(N/C)} strain L3 in the mouse gut. Mice were co-inoculated with an equal mixture of L3 WT and $\Delta eutB$ or $\Delta eutE$ mutant. Seven days post-inoculation, feces were sampled and spotted on LB + Sm plates and LB + Sm + Kan plates to count, respectively, the WT + Δeut mutant and Δeut mutant alone. The WT population was obtained by subtracting the Δeut mutant colony-forming unit (CFU) from the total CFU. Competitive indices (ratio WT/mutant) were calculated for each animal and are represented by a dot. An unpaired Mann-Whitney test was performed to compare both groups. *** $p < 0.001$.

DISCUSSION

EA is a compound naturally present in the gastro-intestinal tract of mammals. Since the discovery of the *eut* operon in the 80s, the utilization of EA has been mostly investigated in bacterial pathogens such as *S. enterica* serovar Typhimurium, EHEC or *L. monocytogenes* and EA was defined as a potential virulence determinant [10]. However, scarce old and recent studies [6,7] have shown the capacity of some human commensal *E. coli* strains to consume EA as a nutrient in contrast to the K12 strain MG1655 widely used as the reference *E. coli* strain in the past [11]. The aim of this study was to evaluate at a large scale the potential of commensal *E. coli* strains to consume EA as an energy source. Among the 40 strains isolated from fecal samples of healthy individuals, 3 different profiles of EA utilization were observed: strains unable to use EA; strains able to use EA only as a N source, which are the most prevalent in our collection, and strains able to exploit both N and C atoms of EA. These data reveal the ability of most commensal *E. coli* strains to catabolize EA but also describe a degree of variability for this metabolism pathway between natural isolates. Of note, 3 strains of our collection were shown to use EA as both the nitrogen and carbon source. The EA concentration used in our work (30 mM) was lower than the one used in a previous study (82 mM) [12]. However, this difference cannot explain the inability of some of our strains to use EA as a C source since EA utilisation as a C source was also variable between strains with high EA levels [12,13]. In order to explain such variability, we sequenced the genome of several *E. coli* isolates belonging to the 3 described profiles. All EA⁻ strains except one have altered *eut* operons, which could explain their inability to use EA. Like for MG1655 and other *E. coli* strains [11,14], strains N3 and U1 present a large insertion of genes between *eutA* and *eutB*. DNA inserts from these strains contain distinct phage related genes, confirming this location within the *eut* operon is a hotspot for phage insertions.

The EA⁻ strain S1 also present a DNA insertion within the *eut* operon. However, this insertion is of small size (0.85 kb instead of ~11kb for N3 and U1) with only one putative ORF (no similarity with characterized proteins) and occurred between gene *eutG* and *eutH*. Among the other EA⁻ strains, AA2 has SNPs introducing premature stop codons in *eutH* and *eutR*, leading to a truncated transcriptional activator EutR and therefore a defect in EA catabolism. The last EA⁻ strain, P3, shows an intact *eut* operon structure but presents a gene expression defect in presence of EA as determined by mRNA quantification (**Fig 4**). It is noteworthy that the promoter sequence upstream of *eutS* in P3 did not present particular variations when compared to other strains (**Fig S4**), suggesting that the expression defect observed for this strain cannot be explained by an altered promoter. Although the EA⁻ phenotype was deciphered, our investigation did not allow us to determine the reason why 3 strains of our collection were also able to exploit EA as a C source *in vitro*. One hypothesis could rely on the specific location where EA catabolism takes place within the bacterial cell. Indeed, EA is transformed within a microcompartment formed by several Eut proteins called EA metabolosome [4]. Similar to organelles found in cyanobacteria that concentrate CO₂ for fixation [15], EA metabolosome improves the efficiency of enzymatic reactions by concentrating enzymes and substrates. Moreover, this structure prevents the dispersion of products especially acetaldehyde, a volatile intermediate known to be toxic for the cell [4,16]. A study in *S. typhimurium* has shown that deletion of genes *eutLKMN*, encoding shell proteins of the microcompartment, prevents the strain to metabolize EA as a C source [17]. We could hypothesize that EA^{+(N)} strains have an EA metabolosome less efficient to perform its functions than EA^{+(C/N)} strains despite high identity levels observed between shell proteins of EA^{+(N)} and EA^{+(C/N)} strains (**Fig S3**). Among the beneficial role of EA for bacterial pathogens, its utilization as a valuable source of nitrogen and/or carbon source was demonstrated to enhance host colonization for both intestinal and extra-intestinal pathogens [2,13,14,18,19].

Using the EA^{+(C/N)} strain L3, we determined in this study that commensal *E. coli* also gains a competitive advantage to consume EA for an efficient fitness in the mouse gut as demonstrated by co-inoculation experiments between the WT and the $\Delta eutB$ mutant (**Fig 5**). In contrast, the WT strain did not outcompete the $\Delta eutE$ mutant. These data suggest that, in our experimental conditions, the ecological advantage comes from the release of NH₃ by the ammonia lyase EutBC and not from the production of Acetyl-coA by the acetaldehyde dehydrogenase EutE. Despite the ecological advantage conferred by the consumption of EA in the gut, 5 out of the 40 strains of our collection are deficient in EA utilization. This conflicting finding might be explained by metabolic properties allowing the strains to utilise other limiting nutrients to efficiently colonize the gut [20], or either by the fact these strains were just in transit at the time of sampling. Recent studies revealed that EA also constitutes a signaling molecule in the gut that modulates the expression of virulence genes [2,21–23]. The case of EHEC virulence has been intensively investigated these last years. In addition to *eut* genes, EA promotes expression of genes encoding Shiga toxins and type 3 secretion system, which are the two critical virulence factors in EHEC [23,24]. Interestingly, influence of EA on the expression of virulence-associated genes in EHEC is not systematically dependent on EutR and can occur at micromolar concentrations that are insufficient to promote growth. Authors from these works suggest that another yet unidentified EA receptor exists in EHEC and coordinates EA-dependant gene expression in concert with EutR [23]. As exemplified with EHEC, EA may constitute a molecule of high importance for many gut pathogens to successfully colonize suitable niches and manage gene expression circuits in order to establish infection. From this knowledge, we can speculate that commensal strains highly proficient in the utilization of EA could be good competitors against intestinal pathogens such as EHEC and could interfere with EA signaling coordinating their virulence program.

In addition to EA⁺ commensal *E. coli* strains such as those described in this study, other strains from different genus represented in the human gut microbiota such as *Enterococcus*, *Clostridium*, *Streptococcus*, *Lactobacillus*, *Blautia* or *Fusobacterium* have been described to catabolize EA experimentally and/or to contain *eut* genes [5,25]. The beneficial role of efficient EA⁺ commensal strains in limiting bacterial infection might be evaluated in future investigations.

ACKNOWLEDGMENTS

This work was supported by funding from INRAE. Maria-Ines Moreira de Gouveia was a PhD Research Fellow funded by the French ministry of Education and Research.

REFERENCES

- [1] Patel D, Witt SN. Ethanolamine and Phosphatidylethanolamine: Partners in Health and Disease. *Oxid Med Cell Longev* 2017;2017. <https://doi.org/10.1155/2017/4829180>.
- [2] Anderson CJ, Clark DE, Adli M, Kendall MM. Ethanolamine Signaling Promotes Salmonella Niche Recognition and Adaptation during Infection. *PLoS Pathog* 2015;11. <https://doi.org/10.1371/journal.ppat.1005278>.
- [3] Luzader DH, Clark DE, Gonyar LA, Kendall MM. EutR is a direct regulator of genes that contribute to metabolism and virulence in enterohemorrhagic *Escherichia coli* O157: H7. *J Bacteriol* 2013;195:4947–53. <https://doi.org/10.1128/JB.00937-13>.
- [4] Kaval KG, Garsin DA. Ethanolamine Utilization in Bacteria. *MBio* 2018;9. <https://doi.org/10.1128/mBio.00066-18>.
- [5] Tsoy O, Ravcheev D, Mushegian A. Comparative genomics of ethanolamine utilization. *J Bacteriol* 2009;191. <https://doi.org/10.1128/JB.00838-09>.
- [6] Blackwell CM, Scarlett FA, Turner JM. Ethanolamine catabolism by bacteria, including *Escherichia coli*. *Biochem Soc Trans* 1976;4. <https://doi.org/10.1042/bst0040495>.
- [7] Rowley CA, Anderson CJ, Kendall MM. Ethanolamine influences human commensal *Escherichia coli* growth, gene expression, and competition with enterohemorrhagic *E. coli* O157:H7. *MBio* 2018;9. <https://doi.org/10.1128/mBio.01429-18>.
- [8] Nielsen KL, Godfrey PA, Stegger M, Andersen PS, Feldgarden M, Frimodt-Møller N. Selection of unique *Escherichia coli* clones by random amplified polymorphic DNA (RAPD): Evaluation by whole genome sequencing. *J Microbiol Methods* 2014;103. <https://doi.org/10.1016/j.mimet.2014.05.018>.
- [9] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 2000;97:6640–5. <https://doi.org/10.1073/pnas.120163297>.

- [10] Garsin DA. Ethanolamine: A signal to commence a host-associated lifestyle? *MBio* 2012;3. <https://doi.org/10.1128/mBio.00172-12>.
- [11] Soupene E, Van Heeswijk WC, Plumbridge J, Stewart V, Bertenthal D, Lee H, et al. Physiological studies of *Escherichia coli* strain MG1655: Growth defects and apparent cross-regulation of gene expression. *J Bacteriol* 2003;185. <https://doi.org/10.1128/JB.185.18.5611-5626.2003>.
- [12] Jones PW, Turner JM. Interrelationships between the enzymes of ethanolamine metabolism in *Escherichia coli*. *J Gen Microbiol* 1984;130. <https://doi.org/10.1099/00221287-130-2-299>.
- [13] Bertin Y, Girardeau JP, Chaucheyras-Durand F, Lyan B, Pujos-Guillot E, Harel J, et al. Enterohaemorrhagic *Escherichia coli* gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. *Environ Microbiol* 2011;13:365–77. <https://doi.org/10.1111/j.1462-2920.2010.02334.x>.
- [14] Dadswell K, Creagh S, McCullagh E, Liang M, Brown IR, Warren MJ, et al. Bacterial microcompartment-mediated ethanolamine metabolism in *Escherichia coli* urinary tract infection. *Infect Immun* 2019;87. <https://doi.org/10.1128/IAI.00211-19>.
- [15] Yeates TO, Kerfeld CA, Heinhorst S, Cannon GC, Shively JM. Protein-based organelles in bacteria: Carboxysomes and related microcompartments. *Nat Rev Microbiol* 2008;6. <https://doi.org/10.1038/nrmicro1913>.
- [16] Penrod JT, Roth JR. Conserving a volatile metabolite: A role for carboxysome-like organelles in *Salmonella enterica*. *J Bacteriol* 2006;188. <https://doi.org/10.1128/JB.188.8.2865-2874.2006>.
- [17] Brinsmade SR, Paldon T, Escalante-Semerena JC. Minimal functions and physiological conditions required for growth of *Salmonella enterica* on ethanolamine in the absence of the metabolosome. *J Bacteriol* 2005;187. <https://doi.org/10.1128/JB.187.23.8039-8046.2005>.

- [18] Rowley CA, Sauder AB, Kendall MM. The ethanolamine-sensing transcription factor EutR promotes virulence and transmission during *Citrobacter rodentium* intestinal infection. *Infect Immun* 2020;88. <https://doi.org/10.1128/IAI.00137-20>.
- [19] Mellin JR, Koutero M, Dar D, Nahori MA, Sorek R, Cossart P. Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA. *Science* (80-) 2014;345. <https://doi.org/10.1126/science.1255083>.
- [20] Conway T, Cohen PS. Commensal and pathogenic *Escherichia coli* metabolism in the gut. *Microbiol Spectr* 2015;3. <https://doi.org/10.1128/microbiolspec.mbp-0006-2014>.
- [21] Ormsby MJ, Logan M, Johnson SA, McIntosh A, Fallata G, Papadopoulou R, et al. Inflammation associated ethanolamine facilitates infection by Crohn's disease-linked adherent-invasive *Escherichia coli*. *EBioMedicine* 2019;43:325–32. <https://doi.org/10.1016/j.ebiom.2019.03.071>.
- [22] Anderson CJ, Kendall MM. Location, location, location. *Salmonella* senses ethanolamine to gauge distinct host environments and coordinate gene expression. *Microb Cell* 2016;3:89–91. <https://doi.org/10.15698/mic2016.02.479>.
- [23] Kendall MM, Gruber CC, Parker CT, Sperandio V. Ethanolamine controls expression of genes encoding components involved in interkingdom signaling and virulence in enterohemorrhagic *Escherichia coli* O157:H7. *MBio* 2012;3. <https://doi.org/10.1128/mBio.00050-12>.
- [24] Gonyar LA, Kendall MM. Ethanolamine and choline promote expression of putative and characterized fimbriae in enterohemorrhagic *Escherichia coli* O157: H7. *Infect Immun* 2014;82:193–201. <https://doi.org/10.1128/IAI.00980-13>.
- [25] Kaval KG, Singh K V., Cruz MR, DebRoy S, Winkler WC, Murray BE, et al. Loss of ethanolamine utilization in *Enterococcus faecalis* increases gastrointestinal tract colonization. *MBio* 2018;9. <https://doi.org/10.1128/mBio.00790-18>.

- [26] Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, et al. The pangenome structure of *Escherichia coli*: Comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J Bacteriol* 2008;190. <https://doi.org/10.1128/JB.00619-08>.
- [27] Doublet B, Douard G, Targant H, Meunier D, Madec JY, Cloeckaert A. Antibiotic marker modifications of λ Red and FLP helper plasmids, pKD46 and pCP20, for inactivation of chromosomal genes using PCR products in multidrug-resistant strains. *J Microbiol Methods* 2008;75. <https://doi.org/10.1016/j.mimet.2008.06.010>.
- [28] Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC. Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res* 1989;17. <https://doi.org/10.1093/nar/17.19.7843>.
- [29] Lane DJ. 16S/23S rRNA Sequencing. *Nucleic Acid Tech Bact Syst* 1991.
- [30] Branchu P, Matrat S, Vareille M, Garrivier A, Durand A, Crépin S, et al. NsrR, GadE, and GadX interplay in repressing expression of the *Escherichia coli* O157:H7 LEE pathogenicity island in response to nitric oxide. *PLoS Pathog* 2014;10:e1003874. <https://doi.org/10.1371/journal.ppat.1003874>.

SUPPLEMENTARY DATA

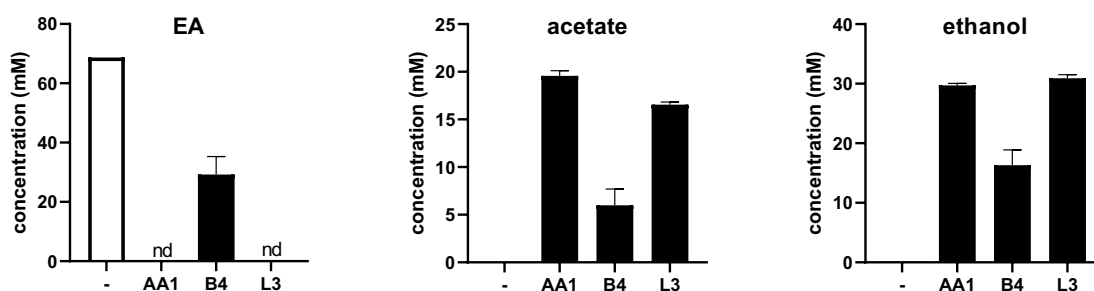


Figure S1

NMR quantification of EA and main products of the EA metabolism. *In vitro* growth of AA1, B4 and L3 commensal *E. coli* EA^{+(N/C)} strains performed in M9 minimal medium supplemented with EA as a sole carbon source. After 72 hours of culture, supernatants were collected, and EA and the main products of the EA metabolism were quantified by NMR (MetaToul, Toulouse, France). Results are presented as mean and standard deviations from three independent experiments. The concentration of the different products is represented before growth (white bars) and after growth (black bars). nd: not detected

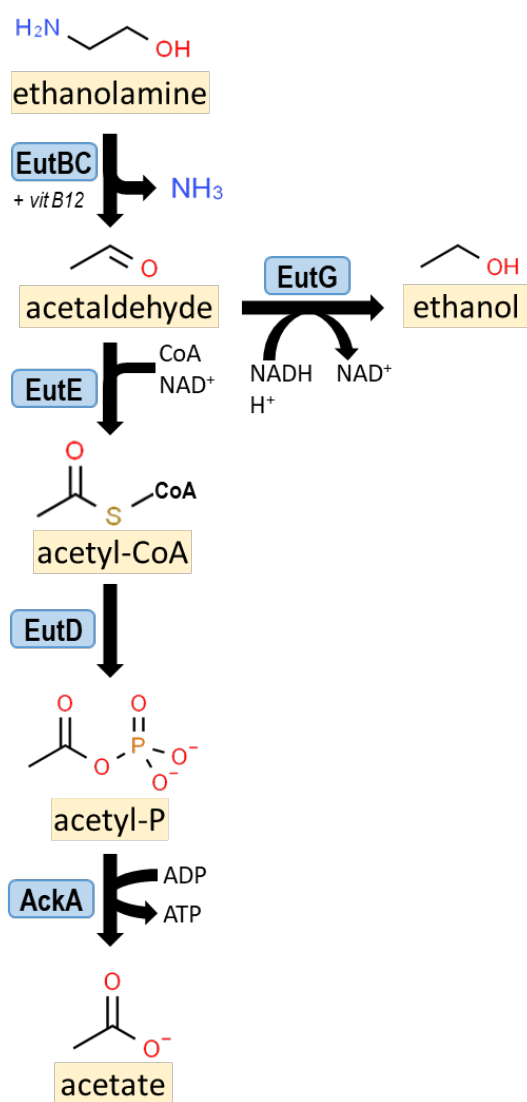


Figure S2

Model of EA catabolism pathway in *E. coli*. Schematic representation of the metabolic pathway involved in EA degradation. Enzymes are indicated in blue and metabolic intermediates in yellow. Co-factors and accessory molecules are not all indicated.

identity % (vs HS)	EA ⁺					EA ^{+(N)}								EA ^{+(N/C)}			mean	<i>Eut protein fonction</i>
	AA2	N3	P3	S1	U1	C2	D2	F3	I1	M2	R2	Y1	AA1	B4	L3			
<i>eutS</i>	100	100	100	99.1	100	100	100	100	99.1	100	99.1	100	99.1	100	100	99.8	Microcompartment	
<i>eutP</i>	100	99.37	98.74	93.71	98.74	97.48	98.11	98.11	97.48	98.11	97.48	98.11	97.48	99.37	99.37	98.1	enzyme	
<i>eutQ</i>	99.14	99.57	99.14	94.85	96.57	97	96.57	97	97.42	96.57	97.42	98.28	97.42	98.71	99.14	97.7	enzyme	
<i>eutT</i>	100	100	100	95.51	99.25	100	99.25	100	99.25	99.25	99.25	99.25	99.25	100	100	99.4	facilitator	
<i>eutI</i>	100	99.7	100	96.45	99.41	99.11	100	99.7	99.41	99.11	99.41	99.41	99.41	100	100	99.4	enzyme	
<i>eutM</i>	100	100	100	97.3	99.1	100	100	100	100	100	100	100	100	100	100	99.8	Microcompartment	
<i>eutN</i>	100	100	100	100	100	100	100	100	100	100	100	98.95	100	100	100	99.9	Microcompartment	
<i>eutE</i>	99.14	99.57	99.57	97.86	98.93	98.5	99.36	98.93	98.93	98.93	98.93	99.14	98.93	99.57	99.57	99.1	enzyme	
<i>eutJ</i>	99.28	99.28	99.28	96.04	98.92	97.48	97.84	97.48	98.2	98.2	98.2	98.92	98.2	99.28	98.56	98.3	facilitator	
<i>eutG</i>	99.24	99.24	99.24	96.46	98.99	99.24	99.49	99.24	97.97	97.72	97.97	98.73	97.97	99.24	99.49	98.7	enzyme	
<i>eutH</i>	-	100	100	98.53	99.02	96.08	99.02	99.26	99.02	99.02	99.02	99.02	99.02	100	100	99.1	EA transporter	
<i>eutA</i>	97.86	99.57	99.36	96.36	98.5	98.29	98.29	98.29	98.07	98.07	98.29	98.29	98.29	99.57	99.57	98.4	facilitator	
<i>eutB</i>	100	100	100	99.12	99.78	100	99.78	100	100	100	100	100	100	100	100	99.9	enzyme	
<i>eutC</i>	99.66	99.66	100	96.61	99.32	99.66	99.66	99.66	99.32	100	99.32	99.32	99.32	99.32	100	99.4	enzyme	
<i>eutL</i>	97.72	100	100	98.17	100	99.54	99.54	99.09	99.09	99.54	99.09	100	99.09	100	100	99.4	Microcompartment	
<i>eutK</i>	100	100	100	92.77	98.19	99.4	99.4	99.4	99.4	99.4	99.4	98.8	99.4	100	100	99.0	Microcompartment	
<i>eutR</i>	-	98.57	98.29	95.43	99.14	98.86	98.57	98.57	98.86	98.29	98.86	99.14	98.86	98.29	98.57	98.4	regulator	
mean	99.5	99.7	99.6	96.7	99.1	98.9	99.1	99.1	98.9	99	98.9	99.1	98.9	99.6	99.7			

Figure S3

Identity percentage of *eut* genes between strains. Values in the table indicate the identity percentage of all *eut* genes for each strain using HS as a reference genome. A color code is associated with the identity percentage from red (low identity) to green (high identity). Each *eut* gene is associated with an Eut protein function.

```

HS      tttttttaactctcacgcttaccctgaatattcagggtaagcagtttagctgcaatatattagtaaagcttattactgagtttgcaataataaaa-aa
AA1    tcattcttaactctcacgcttaccctgaatattcagggtaagcgggttttcctgaaatataattagtaaagcttattactgagtttacgaataataaa-aa
L3     tcTTTTTTAACTCTCACGCTTACCCTGAATATTCAGGGTAAGCGGTTTGGGCTGCAATATATTAGTAAAGCTTATTACTGAGTTTGCGAATAATAAAA-AA
B4     tcTTTTTTAACTCTCACGCTTACCCTGAATATTCAGGGTAAGCGGTTTGGGCTGCAATATATTAGTAAAGCTTATTACTGAGTTTGCGAATAATAAAA-AA
P3     tcTTTTTTAACTCTCACGCTTACCCTGAATATTCAGGGTAAGCGGTTTGGGCTGCAATATATTAGTAAAGCTTATTACTGAGTTTGCGAATAATAAAA-AA

HS      aagcagtcataataatctcgatattatTTTATTATATTCATGCGTTGCATATGAAAGTTTATGCACCACAGCGAATATACTCCATTCTTAGTGAT
AA1    aagcagtcataataatctcgatattatTTTATTATATTCATGCGTTGCATATGAAAGTTTATGCATAACAGTAAATATCTCTCCATTCTTAGTGAT
L3     aagcagtcataataatctcgatattatTTTATTATATTCATGCGTTGCATATGAAAGTTTATGCACCACAGCGAATACCTCTCCATTCTTAGTGAT
B4     aagcagtcataataatctcgatattatTTTATTATATTCATGCGTTGCATATGAAAGTTTATGCACCACAGCGAATATCTCTCCATTCTTAGTGAT
P3     aagcagtcataataatctcgatattatTTTATTATATTCATGCGTTGCATATGAAAGTTTATGCACCACAGCGAATATCTCTCCATTCTTAGTGAT

HS      ctacctcacctTTTAAACGCGCTTGCCGAATTTGTATTACTCTGACGAAAAATTGTCACGATACACGAAAGTTTTTCACAGGCGGGGACTCATG
AA1    ctacctcacctTTTAAACGCGCTTGCCGAATTTGTATTACTCTGACGAAAAATTGTCACGATACACGAAAGTTTTTCACAGGCGGGGACTCATG
L3     ctacctcacctTTTAAACGCGCTTGCCGAATTTGTATTACTCTGACGAAAAATTGTCACGATACACGAAAGTTTTTCACAGGCGGGGACTCATG
B4     ctacctcacctTTTAAACGCGCTTGCCGAATTTGTATTACTCTGACGAAAAATTGTCACGATACACGAAAGTTTTTCACAGGCGGGGACTCATG
P3     ctacctcacctTTTAAACGCGCTTGCCGAATTTGTATTACTCTGACGAAAAATTGTCACGATACACGAAAGTTTTTCACAGGCGGGGACTCATG

```

Figure S4

Comparison of the *eut* promoter region between strains. Alignment of the nucleotidic sequence upstream of *eutS* of strain HS, the EA^{+(N/C)} strains AA1, L3 and B4 and the EA⁻ strain P3.

Table S1. Bacterial strains and plasmids used in this study

Strains or plasmids	Description	References
<i>Bacterial strains</i>		
A1, A2; B3; B4; C1; C2; D2; E3; F1; F2; F3; G1; G4; H2; I1; J1; L3; L4; M1; M2; N2; N3; O1; O2; P3; Q1; Q4; R2; S1; S2; S4; T1; U1; U3; V1; V2; V3; Y1; AA1; AA2	<i>E. coli</i> strains isolated from fecal samples of healthy individuals	this study
HS	<i>E. coli</i> commensal strain isolated from healthy human	[4]
L3 Δ <i>eutB</i>	L3 Δ <i>eutB</i> ; Kan ^R	this study
L3 Δ <i>eutE</i>	L3 Δ <i>eutE</i> ; Kan ^R	this study
L3 Δ <i>eutG</i>	L3 Δ <i>eutG</i> ; Kan ^R	this study
<i>Plasmids</i>		
pKD46-Gm	λ Red recombinase expression, Gm ^R	[5]
pKD4	<i>kan</i> cassette template, Amp ^R , Kan ^R	[6]
pMS	Cloning vector, Sm ^R	Life Technologies

Table S2. Primers used in this study

Purpose	Name	Sequence 5'-3'	References
16S	8F	AGAGTTTGATCCTGGCTCAG	[7]
	1492R	CGGTTACCTTGTTACGACTT	[8]
RAPD	1283	GCGATCCCCA	[9]
RT-qPCR	eutS_qPC R_F	GCATCATTCAGGAATTTGTGCC	this study
	eutS_qPC R_R	AACGCCGATTTTCTTCGCC	this study
	eutE_qPC R_F	ATGCAAAGCAGTGACACGCC	this study
	eutE_qPC R_R	CGCATTGCCACGCTTTTAAACC	this study
	eutB_qPC R_F	CTGGCTAAAGCCAACGAAC	this study
	eutB_qPC R_R	GCAGTCATCTTCATAGGCAATC	this study
	eutG_qPC R_F	CATGCCATTGAAGCATAACAGCG	this study
	eutG_qPC R_R	AAACGCCATTCCCGCCATAC	this study
	rpoA_qua ntif_F	GGTGAGAGTTCAGGGCAAAG	[10]
	rpoA_qua ntif_R	ACCGCGCTGAACTTTGATAC	[10]
<i>eut</i> mutants	eutB_Km _F	GTGCCGGTGACGGTGAAATCACTCGCATTTCCTT CCTGAGGGAACGACTTAGCCACGTTGTGTCTCA AAATC	[11]
	eutB_Km _R	TACAATTTCTTCAATCTGTTTTTGATCCATGATGT GTTATCTCCGCGTCATTAGAAAACTCATCGAGC A	[11]
	eutE_Km _F	GAGGTGGTGTCTGGCGGTCAGGTAATTTCCAC AAATAAGGCAGAACATAGCCACGTTGTGTCTCA AAATC	[11]
	eutE_Km _R	TTGCAGACGTGGGGTGAGCCATTGTTTCGTCGTGC GCCATCTGTTACTCCTTAGAAAACTCATCGAGC A	[11]
	eutG_Km _F	ATTTACCGCAGCACAGTTTGTATGACCCCGCT GGCGATCGCCAGTAAGCCACGTTGTGTCTCAAA ATC	[11]
	eutG_Km _R	ACTCCGCAATCAGCTCACTTACCGCGTTAATAG CGTCACGATCGTCGGATTAGAAAACTCATCGA GCA	this study
	eutA_F Km_verif _R	GATTCGTCGCGCGTTTTCC AGCCAGTTTAGTCTGACCATC	this study this study
	eutN_F	CAAATTGCTGATGGTGGAAA	[11]
	eutJ_F	GATCTGCTGCAACTGGACAA	[11]

MANUSCRIPT 3

Design and validation of a dual-fluorescence reporter system to monitor bacterial gene expression *in situ*

Maria Ines Moreira de Gouveia¹, Audrey Reuter¹, Annie Garrivier¹, Julien Daniel¹, Annick Bernalier-Donadille¹, Gregory Jubelin^{1*}

¹ *Université Clermont Auvergne, INRAE, MEDIS UMR454, F-63000 Clermont-Ferrand, France*

* corresponding author: Gregory Jubelin, INRAE Auvergne Rhône-Alpes, UMR454 MEDIS, site de Theix, 63122 Saint-Genès-Champanelle, France. Email: gregory.jubelin@inrae.fr

ABSTRACT

Fluorescence-based reporter systems are valuable tools for studying gene expression dynamics in living cells. However, available strategies to follow gene expression in bacteria within their natural ecosystem that can be typically rich and complex, are scarce. In this work, we designed and developed a new plasmid-based tool ensuring both the identification of a strain of interest in complex environments and the monitoring of gene expression through the combination of two distinct fluorescent proteins as reporter genes. The tool was validated in *Escherichia coli* to monitor the expression of *eut* genes involved in the catabolism of ethanolamine. We demonstrated that the constructed reporter strain gradually responds to increasing ethanolamine concentrations during *in vitro* cultures. The reporter strain was next inoculated to mice and a flow cytometry strategy was developed to specifically detect the reporter strain among the dense microbiota of intestinal samples, allowing the analysis of *eut* expression level *in vivo*. This novel dual-fluorescent reporter system would be helpful to evaluate transcriptional processes in bacteria within complex environment.

INTRODUCTION

During the last fifty years, multiple techniques have been developed to monitor gene expression in microorganisms. Among them, transcriptional fusions using fluorescent protein (FP) encoding genes are tools extensively used to record the expression dynamics of genes in living bacterial cells [1,2]. Since the discovery of green fluorescent protein (GFP), numerous GFP-variants and other FPs with unique spectral properties have been engineered, resulting in a diverse set of colors that could be used separately or combined together. Among their multiple benefits, FPs allow non-destructive spatiotemporal analysis of gene expression at the single-cell level in various conditions both *in vitro* and *in vivo*. Despite these advances, there is still an important need for improved strategies to analyse bacterial behaviours in their natural ecosystem which usually corresponds to complex environments including various and often dense populations of microorganisms. If numerous works developed strategy to detect specific bacteria in natural environment by immuno-staining with appropriate antibody or by the use of fluorescent tracers [3–7], few combine detection of a bacterial strain with transcriptional tools to report gene expression in individual cells. In this report, we designed a new plasmid-based tool ensuring both the identification of a strain of interest in complex environments and the monitoring of gene expression through the combination of two distinct FPs as reporter genes. The gene expression monitoring system includes a genetic amplifier to warrant transcription measurement of weakly expressed genes from their native chromosomal position [8]. The signal amplification is generated by insertion of the T7 RNA polymerase (T7pol) encoding gene at a locus of interest in the bacterial chromosome. Synthesis of T7pol, which reflects the native expression level of target gene, initiates the transcription of a plasmid-based FP gene controlled by the T7 RNAP-specific promoter. The creation of multiple FP transcripts by each T7pol molecule increases the fluorescence signal and improves the detection of poorly expressed genes. As an additional improvement, the FP used to record gene expression is a derivative of

GFP obtained by the addition of a specific C-terminal peptide recognized by intracellular tail-specific proteases found in numerous Gram-negative and Gram-positive bacteria [9]. This protein extension leads to a destabilized variant of GFP, which is initially a highly stable protein, and allows an accurate monitoring of up- and down-promoter activities for a better study of temporal gene expression. The new designed tool was validated in *Escherichia coli* to monitor the expression of *eut* genes involved in the catabolism of ethanolamine (EA). EA is a small organic molecule found in the gastro-intestinal tract of mammals that can be used as nutrients by intestinal bacteria, especially pathogens [10]. EA catabolism requires the expression of 17 genes clustered in the *eut* operon and transcription of this unit is mainly controlled by the AraC/XylS-type transcriptional regulatory protein EutR. Activity of this specific regulator is induced by the presence of EA and vitamin B₁₂ and activation of *eut* operon by EutR creates a positive-feedback loop since *eutR* is the last gene of the *eut* operon [11]. In this work, we first demonstrated that the *eut* reporter *E. coli* strain gradually responds to known EA concentrations during *in vitro* cultures. The strain was next inoculated to mice and a flow cytometry strategy was developed to specifically detect the reporter strain among the dense microbiota of intestinal samples and to analyse the expression level of *eut* operon *in situ*.

MATERIAL & METHODS

Bacterial strains, plasmids and growth conditions

Strains and plasmids used in this study are listed in [Table S1](#). Bacterial strains were cultured in LB with shaking at 37 °C. In some conditions, LB was supplemented with various concentrations of ethanolamine hydrochloride and 150 nM of cyanocobalamin (vitamin B₁₂). When required, antibiotics were used at the following concentrations: chloramphenicol (25 µg.ml⁻¹); streptomycin (Sm) (50 µg.ml⁻¹); gentamicin (15 µg.ml⁻¹).

Construction of pHL51 plasmid

The reporter plasmid pHL51 was generated through modifications of pHL32 [8]. First, the *kan* sequence, obtained by *SacI* and *AatII* digestion, was replaced with a DNA fragment including *spcR* gene (from pMS vector (Thermofisher)) and mTagBFP2 (from Addgene plasmid#34632) under the control of a strong constitutive promoter (IGEM database, sequence#BBa_J23119). The DNA fragment was generated by PCR using primers 2019_02_F and 2019_02_R ([Table S2](#)) and ligated with pHL32 backbone using In-Fusion cloning kit (Takara Bio). *gfp* gene from the resulting plasmid was next modified by inverse PCR using primers pHL50_AAV_F and pHL50_AAV_R in order to obtain the destabilized GFP variant (GFP[AAV]) through addition of the peptide RPAANDENYAAAV at the C-terminal end of GFP [12]. Sequence of pHL51 was validated by DNA sequencing.

*Construction of the *eut* reporter strain*

The *eut* reporter strain was constructed by inserting the coding sequence of the T7 RNA polymerase gene (T7pol) downstream of the *eutR* gene in *E. coli* strain HS. Briefly, a T7pol-*cam* cassette was amplified from cHNL135 strain with primers 2019_03_F_L3 and 2019_03_R ([Table S2](#)). PCR product was purified and next introduced into HS (pKD46-Gm) by

electroporation following the method of Datsenko and Wanner [13]. Correct integration of the T7pol-*cam* cassette was verified by PCR and sequencing.

Fluorescence quantification

For expression quantification assays, fluorescence was measured with a Spark microplate reader (Tecan) using 200 µl of bacterial suspension in black 96-well plates with clear bottoms. Green fluorescence from GFP was analysed using excitation at 485±20 nm and emission at 530±20 nm and blue fluorescence from mTagBFP2 using excitation at 399±20 nm and emission at 454±20 nm. When cultures were performed directly within the microplate reader, 96-well plates were incubated at 37 °C with orbital shaking and OD_{600nm} and fluorescence signals were recorded every 20 minutes. Fluorescence quantification was also performed by flow cytometry (Cytotflex, Beckman Coulter). FSC (forward scatter), SSC (side scatter), FITC (488nm exciter, 525±20 nm emitter) and PB450 (405 nm exciter, 450±22 nm emitter) axes were set to log display. For sample analysis, the following gating strategy was applied: bacterial cells were gated based on a FSC and SSC plot, and then analysed for BFP signal. A new gate covering BFP⁺ cells, which correspond to bacterial cells of our reporter strain was selected and then analysed for GFP signal. At least 10,000 events of BFP⁺ cells were recorded for each sample and data were analysed with CytExpert software v2.2 (Beckman Coulter) or with Floreada tool (<https://floreada.io>).

Real-time microscopy

For microscopy assays, the *eut* reporter strain was grown into M9 minimal medium supplemented with glucose (0.2%), casamino acids (0.4%) (M9 CASA) and streptomycin 50 µg/ml in order to limit medium autofluorescence. Overnight culture of the *HS eutR*-T7pol (pHL51) strain was diluted to an OD_{600nm} of 0.05, grown until an OD_{600nm} of 0.4, and finally

diluted in M9 CASA by a factor 4 to make the sample. Agarose-mounted slides are prepared as described by Cayron and Lesterlin [14]. 1% agarose (g/v) were melted into M9 CASA medium. A blue frame (In-Situ Hybridisation Frame, ABGENE) was stucked on a microscope slide, leaving only the opened protecting plastic at the top of the blue frame. ~300 μ L of melted agarose were poured into the blue frame and a cover glass was added to remove the exceeding agarose medium and to flatten the surface of the mounted slide. After 30 min, the cover glass and the protective plastic were removed, 10 μ L of culture were dropped at the surface of the agarose-mounted slide and the slide was tilted gently to spread the sample. After letting the sample dry, a new cover glass was added on the blue frame to perform the observation. For all the samples, the agarose mounted-slides were supplemented with streptomycin 50 μ g/ml and when mentioned with Ethanolamine 30 mM and cyanocobalamin 150 nM. Microscopy imaging was carried out on a Zeiss Cell Observer Spinning Disk (YOKOGAWA CSU 1X unit and Hamamatsu Orca Flash 4.0 cameras) equipped with CApochromat 40X NA 1.2 Water objective using Zen Blue software for image acquisition. Acquisition were performed using 50 % power of laser at 405 nm (BFP) and 25 % power 488 nm (GFP) excitation wavelengths. Emissions were filtered at 450 nm (BFP) and 509 nm (GFP) wavelengths. Exposure settings were 100 ms for BFP and 100 ms for GFP. Images were taken every 10 min during 4 hours. The microscope is equipped with a thermostatic chamber preheated to 37°C before the experiment. Images were visualized and treated with ImageJ/Fiji (<https://fiji.sc/>) [4] and the Hyper Stack Reg plugin [5] to align time-lapse experiments.

Mouse experiments

C3H/HeOuj with specific-pathogen-free (SPF) status were purchased from Charles River laboratories. Five-to six-week-old female mice were housed in cages containing no more than five animals, maintained under a 12 h light/dark cycle at a temperature of 21 ± 2 °C, and fed

with standard diet and water ad libitum. The experiments performed herein were reviewed and approved by the Auvergne Committee for Animal Experimentation C2EA and received the agreement number #31333-2021042809157789v2. Mice were given drinking water supplemented with streptomycin (5 g/L) throughout the experiment, starting from one day prior inoculation with 10^7 cells of HS (pHL51) or HS *eutR*-T7pol (pHL51). The drinking water of some groups of mice was also supplemented with 30 mM of ethanolamine hydrochloride and 150 nM of vitamin B₁₂. Five days post-inoculation, mice were euthanized to collect faeces and cecal contents. Samples were resuspended in PBS, centrifuged at 1000 rpm for 1 min and supernatants were filtered using a 40- μ m PET membrane before analysis by flow cytometry.

RESULTS AND DISCUSSION

Rationale design and general description of the reporter system

A map of the designed reporter plasmid pHL51 is shown in [Figure 1](#). This plasmid originates from pHL32, a low copy vector with a p15A origin of replication, hosting *gfp* gene under the control of the T7pol promoter [8]. The following modifications have been implemented to yield pHL51. First, gene encoding the blue fluorescent protein mTagBFP2 placed under a constitutive promoter was inserted in opposite orientation to the P_{T7pol}-*gfp* fusion. Consequently, bacterial cells carrying pHL51 permanently produce mTagBFP2 and emit blue fluorescence. Secondly, the *kan* region of pHL32 was replaced by gene *aadA* conferring resistance to spectinomycin/streptomycin. This antibiotic resistance offers the possibility to work *in vivo* with streptomycin-pretreated mice, a common animal model to study colonization and/or infection of the gastro-intestinal tract by *Enterobacteriaceae* [17]. Finally, the gene *gfp* mut3.1 was modified to insert the peptide RPAANDENYAAAV at the C-terminal end of GFP, leading to the unstable variant GFP[AAV] which allows the studies of transient gene expression in bacteria [9].

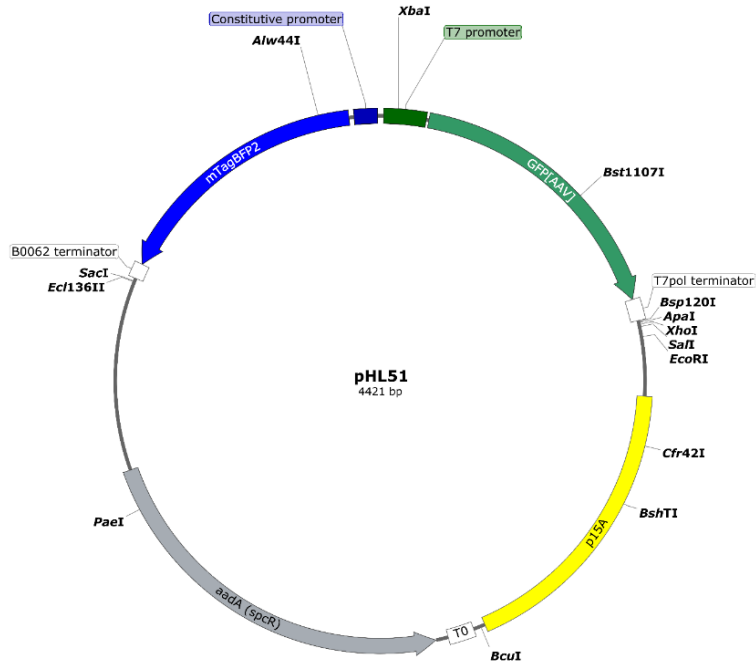


Figure 1 Map of the reporter plasmid pHL51. ORFs are represented by arrows and other elements (promoter, terminator, origin of replication) are represented by rectangles. Indicated enzymes correspond to unique cutters.

The functionality of pHL51 was evaluated using the BL21(DE3) *E. coli* strain carrying the T7pol encoding gene under the control of the IPTG inducible promoter P_{lacUV5} . The strain was grown in LB and IPTG 1mM was added at mid-log phase to induce the production of T7pol. Four hours after the induction, bacterial cells were analysed by flow cytometry to quantify the levels of BFP and GFP fluorescence (**Fig 2A**). When compared to the BL21(DE3) strain alone used as a negative control, the BFP signal from BL21(DE3) pHL51 was high and uniform between cells. In addition, the BFP signal was similar between uninduced and IPTG induced conditions. In contrast, GFP fluorescence strongly increased by a factor 60 in IPTG induced culture. This result demonstrates that *E. coli* carrying pHL51 constitutively emits blue fluorescence and expresses GFP in response to the level of T7pol. Next, the stability of GFP[AAV] and mTagBFP2 was assessed using a culture of BL21(DE3) carrying pHL51. The strain was grown in LB until mid-log phase, induced 1 hour with IPTG and then shifted to PBS in order to remove the inducer and to prevent bacterial growth and protein synthesis. BFP and

GFP signals were quantified after the shift to estimate the stability of both proteins. As shown in Figure 2B, the BFP signal stayed stable over the 3-hour experiment. In contrast, the GFP signal decreased overtime and 50 % of the initial signal was lost after 3 hours. These data indicate that the GFP[AAV] expressed from pHL51 has a limited half-life as initially reported [9].

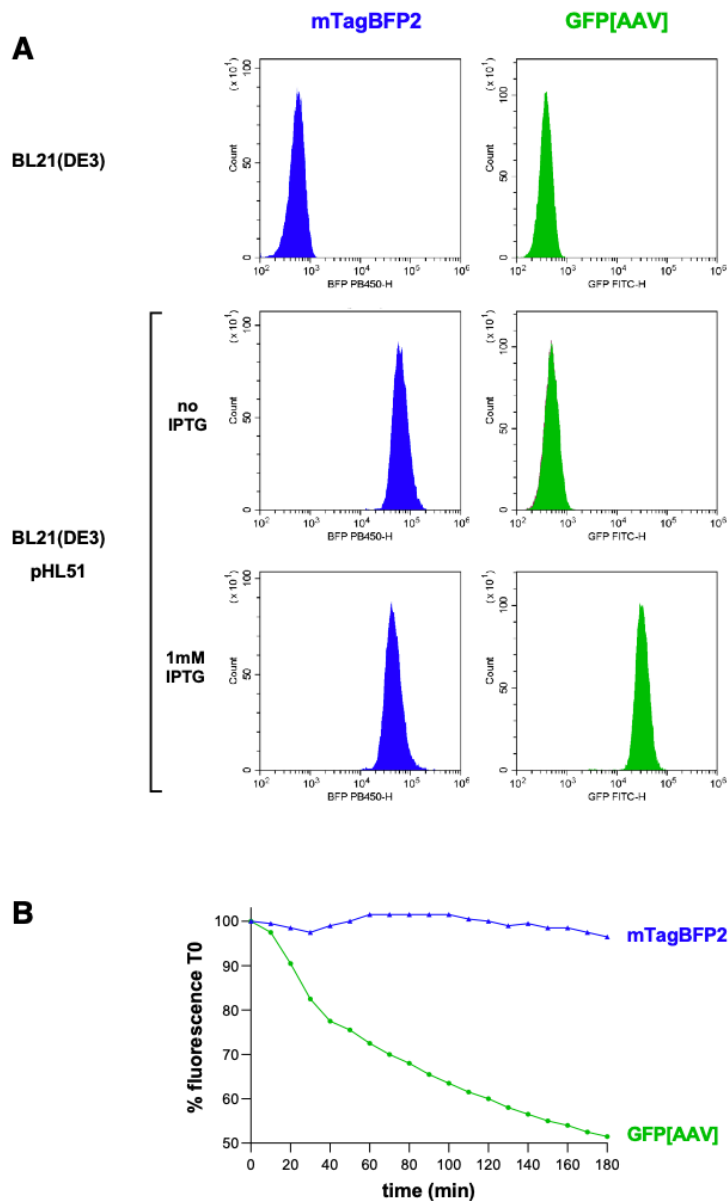


Figure 2 Functional validation of pHL51 reporter plasmid in *E. coli* BL21(DE3). **(A)** The BL21(DE3) strain carrying pHL51 (or without plasmid) were grown in LB supplemented or not with IPTG. Fluorescence emitted from mTagBFP2 and GFP[AAV] were quantified by flow cytometry. Data are shown on histograms with BFP (left panels) or GFP (right panels) signals on the x-axis. **(B)** The BL21(DE3) strain carrying pHL51 was grown in LB + IPTG and the mid-log phase culture was washed once with PBS before quantification of BFP and GFP signals by fluorimetry. Data are presented as percentages of fluorescence relative to the signal recorded after the shift (T0).

Use of the reporter system to study the expression of *eut* genes associated with EA catabolism in commensal *E. coli*

The reporter plasmid pHL51 was next used to monitor the expression of *eut* genes in the *E. coli* commensal strain HS. To this end, we inserted a promoterless T7pol encoding gene immediately downstream *eutR*, which is the last gene of the *eut* operon and this construct was next transformed with pHL51 (**Fig. 3A**). The strain was cultured in LB supplemented with the various concentrations of EA and the co-factor vitamin B₁₂. As expected, the GFP signal gradually increased with EA concentrations (**Fig. 3B**). Moreover, absence of vitamin B₁₂ impede *eut* expression despite the addition of a high concentration of EA. Vitamin B₁₂ is required for EA catabolism because it is an essential co-factor of the ammonia lyase EutBC which transforms EA into acetaldehyde and ammonium [18]. It has been suggested that vitamin B₁₂ is needed for expression of the *eut* operon in *Salmonella enterica* serovar Typhimurium through the binding, together with EA, to EutR yielding an active form of the master regulatory protein [19]. The different samples were next analysed by flow cytometry to quantify GFP and BFP signals in individual cells (**Fig 3C**). First, the BFP signal was homogeneous among cells and did not vary between conditions. In contrast, the GFP signal strongly fluctuated in response to EA concentration in the culture medium. Only 1 % of cells were GFP⁺ in absence of EA, suggesting that few cells express *eut* genes without supplementation of the inducer. The percentage of GFP⁺ cells increased along with the concentration of EA and reached 43% when EA was added at a final concentration of 30 mM (**Fig 3C**). Again, no bacterial cells express the *eut* operon in absence of vitamin B₁₂. It is noteworthy that expression of GFP follows a bimodal distribution with the existence of GFP⁻ and GFP⁺ populations. Indeed, a majority of cells were GFP⁻ even at the highest concentration of EA, indicating that *eut* genes are not transcribed in this population.

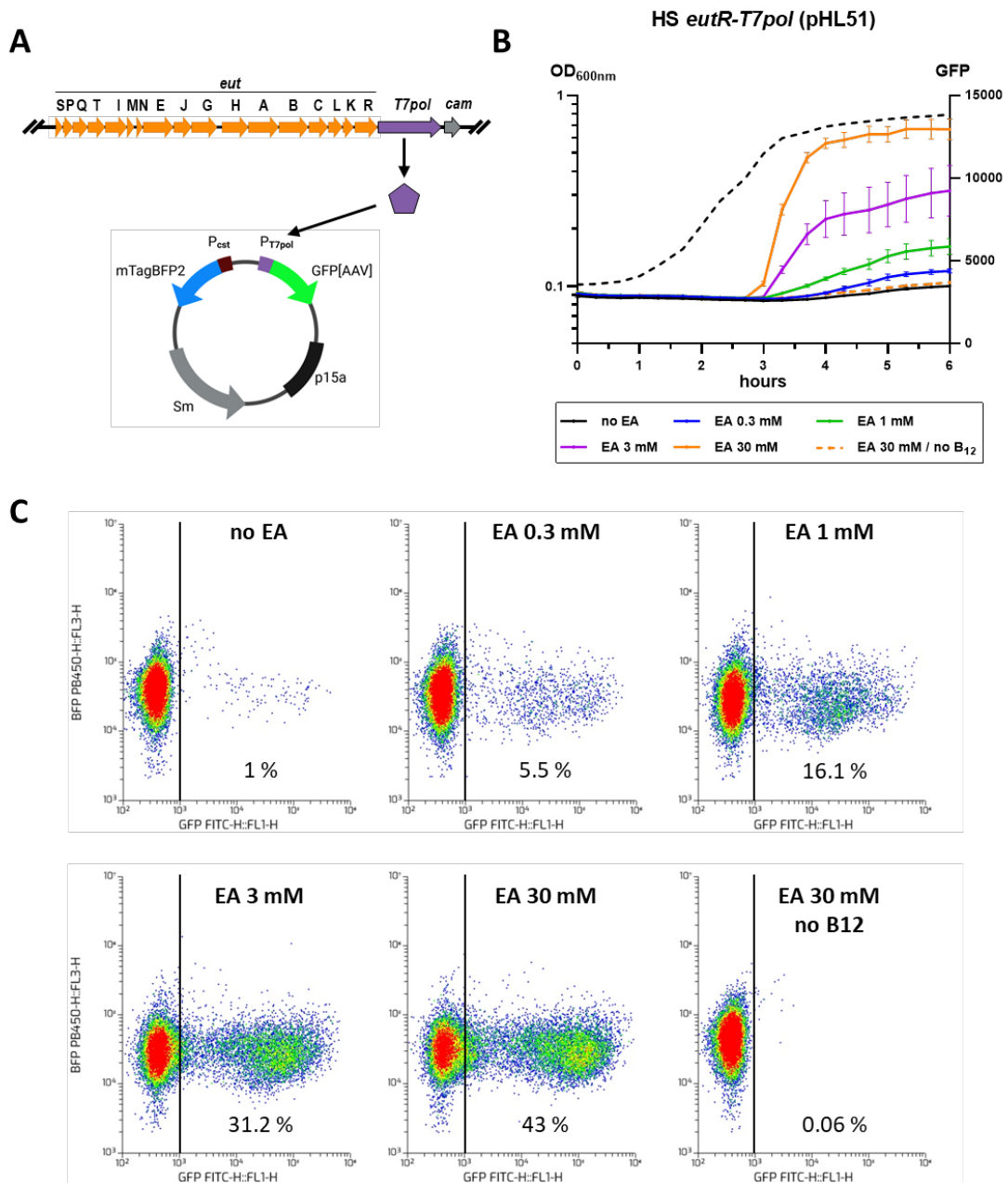


Figure 3 Quantification of *eut* gene expression during *in vitro* growth of *E. coli*.

(A) Schematic representation of the EA reporter strain. (B) The HS *eutR*-T7pol (pHL51) strain was grown in LB supplemented with cyanocobalamin and indicated concentrations of EA. GFP fluorescence and OD_{600nm} was measured every 20 minutes. GFP signals are represented by solid lines and a representative growth curve is shown (dashed line). (C) Six-hour samples of cultures shown in (B) were analysed by flow cytometry. Data are represented on two-dimensional dot plots with BFP and GFP signals on the y-axis and x-axis, respectively. Gates corresponding to GFP-negative and GFP-positive populations are shown and the percentages of the GFP+ cells are indicated at the bottom of each image.

To verify this assumption, we also performed time-lapse microscopy to follow the fluorescence signal kinetic in dividing bacteria. The *eut* reporter strain was grown to mid-log phase and few cells were dropped on an agarose pad supplemented or not with EA + vitamin B₁₂. Fluorescence signals were then recorded for 3 hours under the microscope (Fig 4 and Fig S1). Because no EA was added to the initial culture, all cells were BFP⁺ and GFP⁻ at the beginning of the time-lapse procedure. However, few cells became GFP⁺ very quickly (~20 min) before shutting down again. Interestingly, a second wave of GFP signal occurred later during the time-lapse (~180 min) only for some cells. As expected, no cells turned green when cells were grown on an agarose pad without EA and vitamin B₁₂ (Fig S1 and S2). These data confirm the heterogeneous expression of *eut* genes in strain HS and also demonstrate that *eut* expression is highly dynamic despite the availability of EA in excess. Over the last 20 years, many works demonstrate that clonal microbial cells can display significant diversity in their gene expression level, a phenomenon called phenotypic heterogeneity [20,21]. Moreover, it has been recently suggested that cellular metabolism inherently follows stochastic processes and is a generic source of phenotypic heterogeneity [22]. Accordingly, many metabolic pathways give rise to distinct single-cell behaviours in *E. coli* [23]. Observation of a heterogeneous expression of the *eut* operon in this work add a new example of phenotypic heterogeneity related to metabolic processes in *E. coli*.

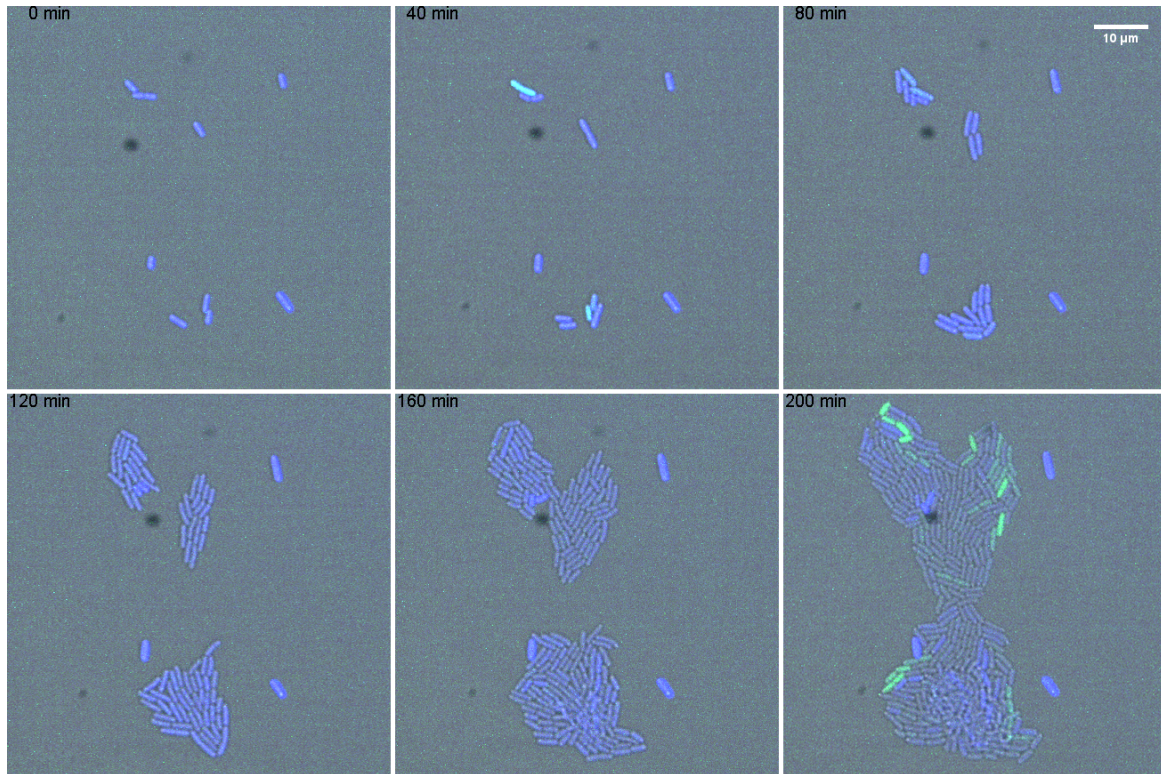


Figure 4 *eut* expression during *in vitro* growth in time-lapse microscopy. An agarose pad of M9 CASA supplemented with vitamin B₁₂ and EA 30 mM was inoculated with cells of HS *eutR*-T7pol (pHL51) grown in liquid culture. Time-lapse microscopy was performed during 200 minutes on several fields and a representative field is shown here. Images correspond to overlays of BFP, GFP and brightfield acquisitions.

Monitoring *eut* expression in vivo during colonization of the mouse gut

Our reporter system was designed to monitor the expression of a gene of interest in complex environment. Indeed, the constitutive expression of mTagBFP2 should allow the tracking of our strain even in dense ecosystems with high level of diverse microorganisms like within the gastro-intestinal tract. We evaluated this feasibility by using the streptomycin mouse model. Mice were given drinking water supplemented with the antibiotic in order to remove the facultative anaerobes while keeping the anaerobe populations [24]. The HS *eutR-T7pol* (pHL51) was next given to mice by oral gavage. The presence of *aat* gene on the pHL51 plasmid allows the selection of our reporter strain in the gut of streptomycin treated mice. EA and vitamin B₁₂ were also added to the drinking water of some animals to artificially increase the level of both factors within the gut. Five days post-inoculation, fecal and cecal contents of animals were recovered and analysed by flow cytometry following the gating strategy depicted in figure 5A. A first gate covering events with FSC/SSC parameters usually defined for bacilli was applied to further analyse BFP fluorescence. A small population exhibited a strong BFP signal and this corresponds to our reporter strain constitutively expressing mTagBFP2 (Fig 5A). These events usually represent less than 0.5 % of the total number of detected events, corroborating with an average of $\sim 1.10^9$ CFU of *E. coli* / g faeces observed in the streptomycin mouse model [25]. These BFP⁺ cells were next gated to analyse green fluorescence in order to report the expression level of *eut* genes. As expected, negligible GFP⁺ events were detected in samples recovered from mice inoculated with HS (pHL51) used as a negative control (Fig 5B). In contrast, up to 5 % of GFP⁺ cells were detected in the cecal content of mice inoculated with the reporter strain HS *eutR-T7pol* (pHL51). This percentage increased up to 35 % when EA and vitamin B₁₂ were added to the drinking water of animals, demonstrating that more HS cells in the gut express the *eut* operon in response to the supplementation. This was confirmed by a quantitative analysis of the GFP signal with a 6-fold induction of fluorescence in samples

recovered from EA supplemented mice (**Fig 5C**). We also analysed bacteria recovered from fecal samples. Whereas the effect of EA supplementation on GFP fluorescence was observed like for the cecal content, the percentage of GFP⁺ cells in untreated mice stayed negligible, suggesting that HS do not encounter anymore EA and/or vitamin B₁₂ in faeces to express *eut* genes (**Fig 5B** and **5C**). *In vivo* expression of the *eut* operon has been demonstrated for gut pathogens such as *Listeria monocytogenes* or *Salmonella enterica* serovar Typhimurium [26,27]. Both pathogens induce a strong inflammation in the gut, contributing to epithelial cell death [28] and release of high level of EA in the intestinal lumen. In addition, EA is used both as a nutrient source and as an intestinal signal to control virulence program of pathogens [26,27,29-31] . Here we show that commensal *E. coli* also express *eut* operon in the mouse gut in absence of an infectious context. This is correlated with a recent finding showing that EA confers a nutritional advantage to commensals in the gastro-intestinal tract [32]. In conclusion, we developed a new dual-fluorescence plasmid-based tool which allows the spatio-temporal monitoring of gene expression in bacteria within their natural complex ecosystem such as the gastro-intestinal tract.

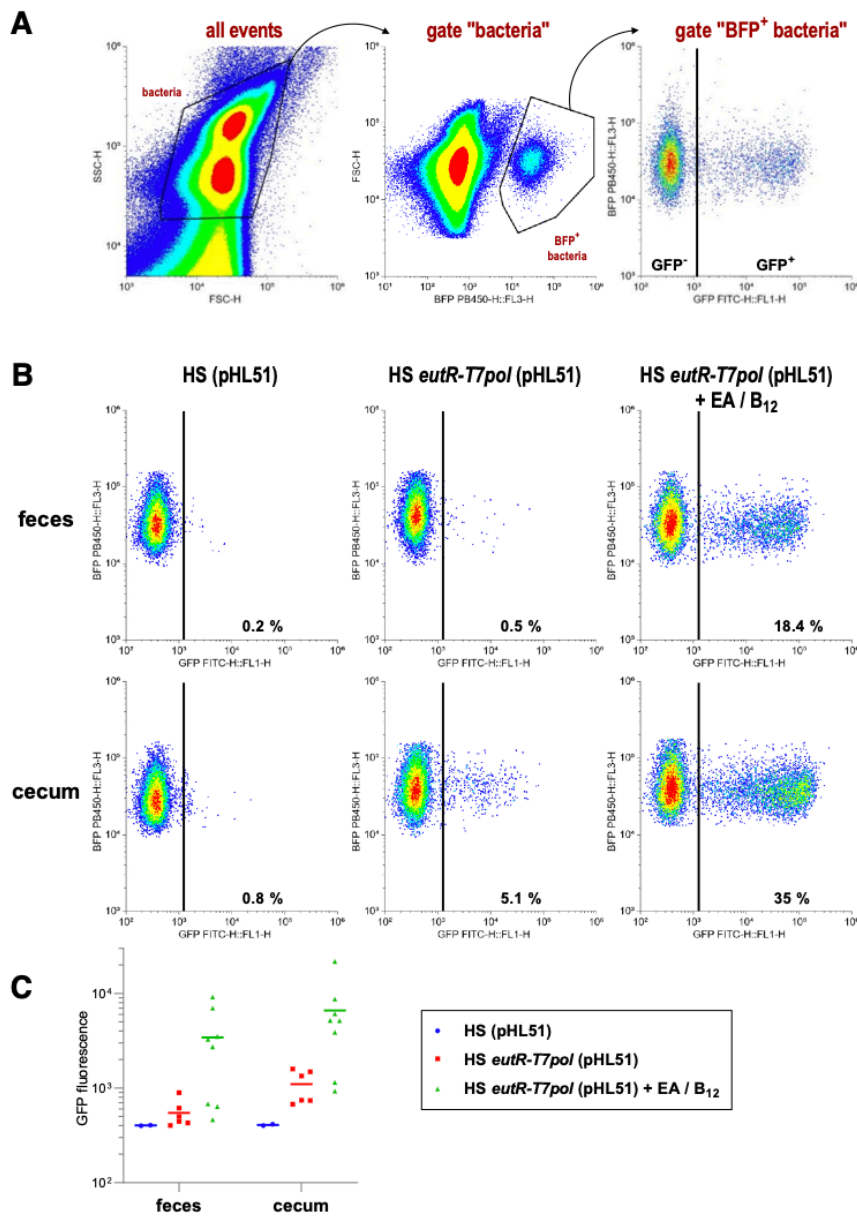


Figure 5 Quantification of *eut* gene expression in the gut of mice inoculated with the *eut* reporter strain. (A) Gating strategy to evaluate *eut* expression in *E. coli* HS sampled from gut contents. Gates are represented by black polygons on the left and middle panels, and gates corresponding to GFP⁻ or GFP⁺ cells are separated by a vertical black line on the right panel. Data are represented on two-dimensional dot plots with BFP and GFP signals on the y-axis and x-axis, respectively. Dot plots were generated using all events or from indicated gates. (B) Fecal or cecal samples were analysed by flow cytometry following the gating strategy presented in (A). Data are represented on two-dimensional dot plots with BFP and GFP signals on the y-axis and x-axis, respectively. Gates corresponding to GFP-negative and GFP-positive populations are shown and the percentages of the GFP⁺ cells are indicated at the bottom of each image. (C) Quantification of GFP signal measured from gut samples of indicated groups of mice. Eat dot represents one mouse and means are indicated as a line.

ACKNOWLEDGMENTS

We thank the staff of CLIC (Clermont Confocal Imaging) and especially Caroline VACHIAS for the excellent assistance in time-lapse microscopy experiments. This work was supported by fundings from INRAE. Maria Ines Moreira de Gouveia was a PhD Research Fellow funded by the French ministry of Education and Research. Audrey Reuter was a Postdoctoral Research Fellow granted by the “Microbiology and Food Chain” Division of INRAE.

REFERENCES

- [1] Rosochacki SJ, Matejczyk M. Green fluorescent protein as a molecular marker in microbiology. *Acta Microbiol Pol* 2002;51.
- [2] Campbell-Valois FX, Sansonetti PJ. Tracking bacterial pathogens with genetically-encoded reporters. *FEBS Lett* 2014;588. <https://doi.org/10.1016/j.febslet.2014.05.022>.
- [3] Packer D. The history of the antibody as a tool. *Acta Histochem* 2021;123. <https://doi.org/10.1016/j.acthis.2021.151710>.
- [4] Yoon SA, Park SY, Cha Y, Gopala L, Lee MH. Strategies of detecting bacteria using fluorescence-based dyes. *Front Chem* 2021;9. <https://doi.org/10.3389/fchem.2021.743923>.
- [5] Yang C, Ren C, Zhou J, Liu J, Zhang Y, Huang F, et al. Dual fluorescent- and isotopic-labelled self-assembling vancomycin for in vivo imaging of bacterial infections. *Angew Chemie - Int Ed* 2017;56:2356–60. <https://doi.org/10.1002/anie.201610926>.
- [6] Nielsen AT, Dolganov NA, Rasmussen T, Otto G, Miller MC, Felt SA, et al. A bistable switch and anatomical site control *Vibrio cholerae* virulence gene expression in the intestine. *PLoS Pathog* 2010;6. <https://doi.org/10.1371/journal.ppat.1001102>.
- [7] Peñate-Medina O, Tower RJ, Peñate-Medina T, Will O, Saris PEJ, Suojanen J, et al. Universal membrane-labeling combined with expression of *Katushka* far-red fluorescent protein enables non-invasive dynamic and longitudinal quantitative 3D dual-color fluorescent imaging of multiple bacterial strains in mouse intestine. *BMC Microbiol* 2019;19. <https://doi.org/10.1186/s12866-019-1538-z>.
- [8] Lim HN, Van Oudenaarden A. A multistep epigenetic switch enables the stable inheritance of DNA methylation states. *Nat Genet* 2007;39. <https://doi.org/10.1038/ng1956>.
- [9] Andersen JB, Sternberg C, Poulsen LK, Bjørn SP, Givskov M, Molin S. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* 1998;64. <https://doi.org/10.1128/aem.64.6.2240-2246.1998>.

- [10] Garsin DA. Ethanolamine utilization in bacterial pathogens: Roles and regulation. *Nat Rev Microbiol* 2010;8. <https://doi.org/10.1038/nrmicro2334>.
- [11] Kaval KG, Garsin DA. Ethanolamine utilization in bacteria. *MBio* 2018;9. <https://doi.org/10.1128/mBio.00066-18>.
- [12] Miller WG, Leveau JHJ, Lindow SE. Improved gfp and inaZ broad-host-range promoter-probe vectors. *Mol Plant-Microbe Interact* 2000;13. <https://doi.org/10.1094/MPMI.2000.13.11.1243>.
- [13] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 2000;97:6640–5. <https://doi.org/10.1073/pnas.120163297>.
- [14] Cayron J, Lesterlin C. Multi-scale analysis of bacterial growth under stress treatments. *J Vis Exp* 2019;2019. <https://doi.org/10.3791/60576>.
- [15] Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: An open-source platform for biological-image analysis. *Nat Methods* 2012;9. <https://doi.org/10.1038/nmeth.2019>.
- [16] Sharma V. ImageJ plugin HyperStackReg V5.6 2018. <https://doi.org/10.5281/zenodo.2252521>.
- [17] Conway T, Krogfelt KA, Cohen PS. The life of commensal *Escherichia coli* in the mammalian intestine. *EcoSal Plus* 2004;1. <https://doi.org/10.1128/ecosalplus.8.3.1.2>.
- [18] Chang GW, Chang JT. Evidence for the B12-dependent enzyme ethanolamine deaminase in *Salmonella*. *Nature* 1975;254. <https://doi.org/10.1038/254150a0>.
- [19] Roof DM, Roth JR. Autogenous regulation of ethanolamine utilization by a transcriptional activator of the eut operon in *Salmonella typhimurium*. *J Bacteriol* 1992;174. <https://doi.org/10.1128/jb.174.20.6634-6643.1992>.

- [20] Martin C Saint, Jubelin G, Darsonval M, Leroy S, Leneveu-Jenvrin C, Hmidene G, et al. Genetic, physiological, and cellular heterogeneities of bacterial pathogens in food matrices: Consequences for food safety. *Compr Rev Food Sci Food Saf* 2022. <https://doi.org/10.1111/1541-4337.13020>.
- [21] Avery S V. Microbial cell individuality and the underlying sources of heterogeneity. *Nat Rev Microbiol* 2006;4. <https://doi.org/10.1038/nrmicro1460>.
- [22] Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. Stochasticity of metabolism and growth at the single-cell level. *Nature* 2014;514. <https://doi.org/10.1038/nature13582>.
- [23] Afroz T, Biliouris K, Kaznessis Y, Beisel CL. Bacterial sugar utilization gives rise to distinct single-cell behaviours. *Mol Microbiol* 2014;93. <https://doi.org/10.1111/mmi.12695>.
- [24] Maltby R, Leatham-Jensen MP, Gibson T, Cohen PS, Conway T. Nutritional basis for colonization resistance by human commensal *Escherichia coli* strains HS and Nissle 1917 against *E. coli* O157:H7 in the mouse intestine. *PLoS One* 2013;8. <https://doi.org/10.1371/journal.pone.0053957>.
- [25] Gardette M, Daniel J, Loukiadis E, Jubelin G. Role of the nitric oxide reductase NorVW in the survival and virulence of enterohaemorrhagic *Escherichia coli* during infection. *Pathogens* 2020;9:1–14. <https://doi.org/10.3390/pathogens9090683>.
- [26] Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, et al. The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 2009;459. <https://doi.org/10.1038/nature08080>.
- [27] Anderson CJ, Clark DE, Adli M, Kendall MM. Ethanolamine signaling promotes *Salmonella* niche recognition and adaptation during infection. *PLoS Pathog* 2015;11. <https://doi.org/10.1371/journal.ppat.1005278>.

- [28] Blander JM. Death in the intestinal epithelium—basic biology and implications for inflammatory bowel disease. *FEBS J* 2016. <https://doi.org/10.1111/febs.13771>.
- [29] Garsin DA. Ethanolamine: A signal to commence a host-associated lifestyle? *MBio* 2012;3. <https://doi.org/10.1128/mBio.00172-12>.
- [30] Rowley CA, Sauder AB, Kendall MM. The ethanolamine-sensing transcription factor EutR promotes virulence and transmission during *Citrobacter rodentium* intestinal infection. *Infect Immun* 2020;88. <https://doi.org/10.1128/IAI.00137-20>.
- [31] Kendall MM, Gruber CC, Parker CT, Sperandio V. Ethanolamine controls expression of genes encoding components involved in interkingdom signaling and virulence in enterohemorrhagic *Escherichia coli* O157:H7. *MBio* 2012;3. <https://doi.org/10.1128/mBio.00050-12>.
- [32] Moreira de Gouveia MI, Daniel J, Garrivier A, Bernalier-Donadille A, Jubelin G. Diversity of ethanolamine utilization by human commensal *Escherichia coli*. *Res Microbiol* 2022;103989. <https://doi.org/10.1016/j.resmic.2022.103989>.
- [33] Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, et al. The pangenome structure of *Escherichia coli*: Comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J Bacteriol* 2008;190. <https://doi.org/10.1128/JB.00619-08>.
- [34] Doublet B, Douard G, Targant H, Meunier D, Madec JY, Cloeckaert A. Antibiotic marker modifications of λ Red and FLP helper plasmids, pKD46 and pCP20, for inactivation of chromosomal genes using PCR products in multidrug-resistant strains. *J Microbiol Methods* 2008;75. <https://doi.org/10.1016/j.mimet.2008.06.010>.
- [35] Subach OM, Cranfill PJ, Davidson MW, Verkhusha V V. An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. *PLoS One* 2011;6. <https://doi.org/10.1371/journal.pone.0028674>.

SUPPLEMENTARY DATA

<https://drive.google.com/file/d/11nvq7Us8v0AYocbwjb81rv6YpXC0W2Ef/view?usp=sharing>

https://drive.google.com/file/d/1_Acq-xwnx4lpeqi0Oerm4jyvW-YYUfFa/view?usp=sharing

Figure S1 Time-lapse movies of HS *eutR*-T7pol (pHL51) strain on agarose pads supplemented or not with EA 30mM and vitamin B₁₂

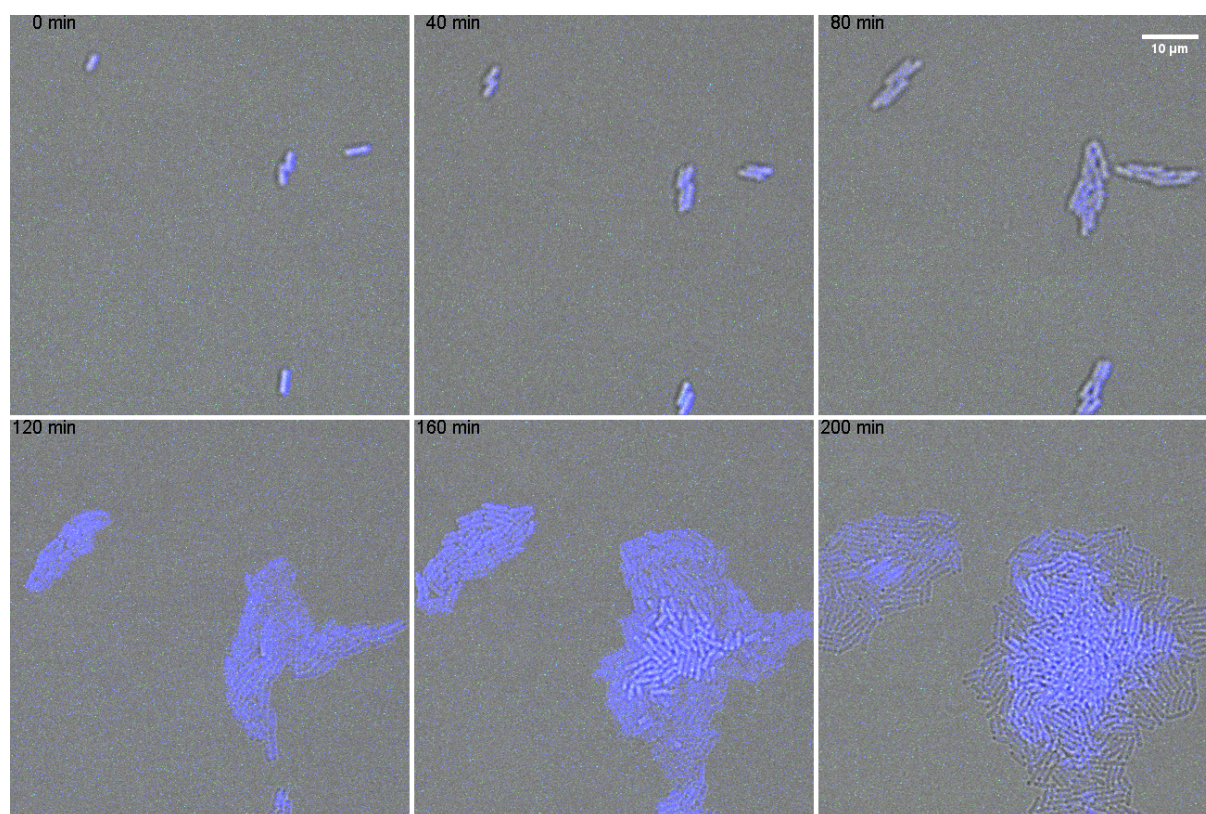


Figure S2 Time-lapse microscopy of HS *eutR*-T7pol (pHL51) strain grown on an agarose pad without EA and vitamin B₁₂ supplementation. Time-lapse microscopy was performed during 200 minutes on several fields and a representative field is shown here. Images correspond to overlays of BFP, GFP and brightfield acquisitions.

Table S1. Bacterial strains and plasmids used in this study

Strains or plasmids	Description	References
Bacterial strains		
HS	<i>E. coli</i> commensal strain isolated from healthy human	[4]
HS <i>eutR</i> -T7pol- <i>cam</i>	HS carrying T7pol- <i>cam</i> cassette inserted downstream of <i>eutR</i> ; Cm ^R	this study
cHNL135	MC4100, agn43:T7RNApol:cam	[14]
Plasmids		
pKD46-Gm	λRed recombinase expression, Gm ^R	[5]
pMS	Cloning vector, Sm ^R	Life Technologies
pBAD-mTagBFP2	pBAD expressing mTagBFP2 protein; Addgene plasmid # 34632	[15]
pHL32	P _{T7RNApol} -GFP3.1mut, Kan ^R	[14]
pHL51	Derivative of pHL32 carrying P _{T7RNApol} -GFP[AAV] and P _{cst} -mTagBFP2; Sm ^R	this study

Table S2. Primers used in this study

Purpose	Name	Sequence 5'-3'	References
Construction of HS <i>eutR</i> -T7pol- <i>cam</i> strain	2019_03_F_L3	CCGAGAAGCCGTCGTTGACGTTGC ATCAGCGGATGCGGGAATGGGGGT GATAACAGGAGGAATTAACCATGC	this study
	2019_03_R	ACACGATTAACATCGCTAAGAAC GAAGCGGCGTTAAGTGAGTTTATT AAGGTCAGGGATTGGGTGTAAGT CCATGGAGTTCTGAGGTCATTACTG	this study
	<i>eutR</i> -F	GGCTGAAACGCATTTCGCCTG	this study
	T7seqR	CGTTATCCGCAACCTCACCAG	this study
Construction of pHL51	2019_02_F	ACAGGAGTCCAAGCGCTTGGTCTG ACAGTTATTATTTACC	this study
	2019_02_R	AGTGAGTCGTATTAGATGCATTG ACAGCTAGC	this study
	pHL50_AAV_F	GAAAACACTACGCTGCAGCAGTTTAA GCTTAATTAGCTGACCATAACC	this study
	pHL50_AAV_R	GTCGTTTGCTGCAGGCCTTTTGTAT AGTTCATCCATGCCATG	this study

DISCUSSION

Through the years, the accumulation of knowledge on gut microbiota composition and function shed light on its importance and impact on the health of an individual. Indeed, several studies on the literature have established a connection between gut dysbiosis and intestinal and extra-intestinal pathologies such as CD and RA. Data suggest that incidence and prevalence of IBD are still increasing worldwide, and approximately 3.7 million European individuals suffer from IBD at the present time [309]. Whilst approximately 2.3 million individuals are diagnosed with RA in Europe each year [310]. For CD patients, several studies have shown an increase in Proteobacteria, mainly *E. coli* associated to the intestinal mucosa [149],[143],[139],[311]. For RA patients, the expansion of *Enterobacteriaceae* appears to be more variable but has been detected in some studies as well as in our small collection of *E. coli* isolates [230]. A bloom of *Enterobacteriaceae* appears to be a signature of dysbiosis for both these inflammatory diseases.

Several theories (described on chapter 3 of the bibliographical review of my manuscript) have been proposed to explain the bloom of *Enterobacteriaceae*. One of these theories is the metabolic adaptation of *Enterobacteriaceae* to the dysbiotic gut environment. The changing conditions of the GIT in case of disease might influence the compounds available in the gut lumen that can be used by commensal bacteria. The main goal of my PhD work was to identify substrates that could give *Enterobacteriaceae* a nutritional advantage to promote their expansion in case of dysbiosis associated with CD and RA.

Part 1 : Identification of compounds that may improve fitness of *Enterobacteriaceae* in the gut of patients suffering from CD or RA.

The first aim of my PhD project was to identify substrates that are differentially consumed by *Enterobacteriaceae* isolated from the gut of healthy individuals or from the gut of CD or RA patients. To that end, screening of our collection of *Enterobacteriaceae* HI, CD and RA strains was performed with the OmniLog system to compare the utilization of 190 different substrates.

The OmniLog analysis showed that RA strains had a better capacity to metabolize α -Ketobutyric acid and D-Malic acid than HI strains. Even though no link exists between D-malic acid and RA in the literature, there are few studies that connect RA to α -Ketobutyric acid. This ketoacid called α -Ketobutyric acid or 2-Oxobutanoate, is a degradation product of several amino acids such as methionine and threonine. This compound can be transformed into propanoyl-CoA that can further participate in the TCA cycle. Very recently Jutley *et al.* (2021) analyzed the serum and urinary metabolome of patients suffering from RA [312]. This study showed a positive correlation between C-reactive protein and the presence of Ketobutyric acid.

C-reactive protein is a protein for which concentrations rise in response to inflammation, suggesting a link between systemic inflammation and increased concentrations of Ketobutyric acid in plasma of early RA patients. In addition, Wang *et al.* analyzed the metabolic profile of plasma on rats that were treated with sodium urate to induce acute gout arthritis, a form of arthritis that is very painful. The plasma metabolic profiling revealed 22 endogenous metabolites associated with acute gout arthritis including α -Ketobutyric acid [313].

Even though RA is an extra intestinal inflammatory disease, alterations of the gut microbiota have been well described on the literature. However, the mechanisms involved in the gut-articulations axis remain unknown. There are complex interactions between the innate and acquired immune responses, involving antigen-presenting cells, autoreactive T cell formation, and production of autoantibodies directed against host cells such as RF and ACPAs. These autoantibodies are often detected in the blood, before any sign of inflammation in the joints, suggesting that the triggering of autoimmunity might occur at different sites to the joints, for example in the GIT [314]. Moreover, Kalinkovich *et al.* demonstrated in a study from 2019 that intestinal dysbiosis precedes auto-immunity and inflammation of the articulations [315]. These data suggest that intestinal dysbiosis might play an important role on the onset of RA meaning that dysbiosis might be more a cause than a consequence of the disease (**Figure 21**). If dysbiosis precedes inflammation at the articulations, we could speculate that the expansion of Proteobacteria might be detectable at the onset of RA. Our study cannot answer to this question because strains from our RA collection were isolated from individuals suffering from RA for several years and treated with different medications. We should also keep in mind that medication could also bias selection of microbial populations in the gut of patients, in addition to the effects induced by the pathology itself. RA is characterized by a constant immune response located at the joints leading to inflammation of synovial tissue and subsequent cartilage and bone erosion [316]. But inflammation is not restricted to articulations, as it was shown by Nissinen *et al.* that demonstrated an augmentation of mRNA expression of inflammatory markers showing that inflammatory cells are also activated in the small intestine suggesting that the gut of RA patients is inflamed [317]. However, this study needs to be interpreted with caution though because it might have been influenced by treatments for RA such as anti-inflammatory treatments that could have biased the results.

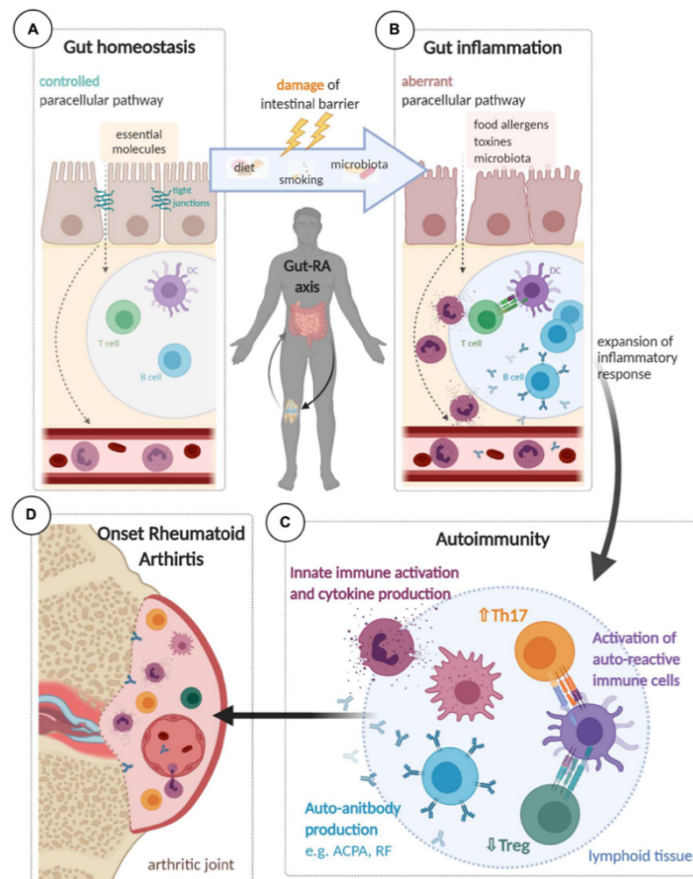


Figure 21 The gut-articulation axis. Under healthy conditions, tight junction molecules control the entry of toxins, food allergens, or microbiota from the gut lumen into the blood stream (A). Environmental factors, like diet, smoking, or the microbiota itself, can lead to damage of the epithelium and to the opening of the paracellular pathway. Toxins, food allergens, or microbiota can cross the epithelium and get in contact with the immune cells leading to inflammation (B). Autoreactive cells are activated in the lymphoid tissues leading to inflammatory cytokine responses and autoantibody (ACPAs and RF) production whereas Tregs are reduced (C). Furthermore, autoreactive cells migrate to the joints and cause cartilage and bone damage (D). From Brandl et al., 2021.

A lot of questions remain unanswered regarding α -Ketobutyric acid. Studies in the literature analyzed the plasma of early RA patients and found a possible link between systemic inflammation and this compound. Whilst our study focused on *E. coli* isolated from the gut of late RA patients and showed that our isolates had a better capacity to catabolize α -Ketobutyric acid. Further studies need to be done to better explore the role of this compound on the different stages of RA and as a selective compound for *E. coli* population in the dysbiotic gut. Firstly, it would be interesting to determine the prevalence of the genes responsible for the metabolism of this compound in our RA collection of strains. In *E. coli* strains, genes *pflB* and *tdcE* code for the enzyme pyruvate-formate lyase responsible for the degradation of this compound into propanoyl-CoA [318]. *In vitro* competitions could also be performed on minimal media

supplemented with α -Ketobutyric acid in order to see if this compound could give an advantage to RA strains over HI strains. As a rat model suffering from arthritis is well-developed in our lab, it would be also of strong interest to perform competition assays with RA and HI strains and to quantify α -Ketobutyric acid in feces and plasma. Identification of substrates used by *Enterobacteriaceae* to expand during dysbiosis appears important as it could contribute to the development of novel approaches to prevent RA since dysbiosis seems to play an important role at the onset of the disease.

By contrast with the comparison HI vs RA strains, the OmniLog results showed more differences for the comparison between HI and CD strains. The gut inflammation status is quite different between these two diseases, and inflammation is known to have a deep impact on the gut environment. That might be the reason why we found more differential compounds for the comparison HI vs CD. OmniLog analysis with the CD strains revealed a better capacity to metabolize D-serine by this group while HI strains had a better capacity to metabolize sucrose. Utilization of D-serine is possible thanks to the presence of the *dsdCXA* locus encoding the activator DsdC, the inner-membrane transporter DsdX and DsdA which is the D-serine deaminase/dehydratase transforming D-serine to pyruvate and ammonium [319]. The *dsdCXA* locus is localized in the *argW* chromosomal region that corresponds to the *E. coli* K-12 MG1655 53' genetic map position [320]. The locus *argW* is a hotspot site for integration of mobile genetic elements. Interestingly, a study from 2002 highlighted the opposite link between the catabolism of D-serine and sucrose. Indeed, when comparing the *argW* regions of several *E. coli* strains, authors found a surprising level of diversity. While *E. coli* strain K-12 possesses the *dsdCXA* locus in the *argW* site, other strains including pathogenic strains O157:H7, have the *csc* locus responsible for the catabolism of sucrose [321] (**Figure 22**). The strains from our work also present an opposite relationship between the presence of the *dsd* and the *csc* clusters [321],[319] with the finding that most CD strains harbour the *dsd* locus whereas most HI strains harbor the *csc* locus.

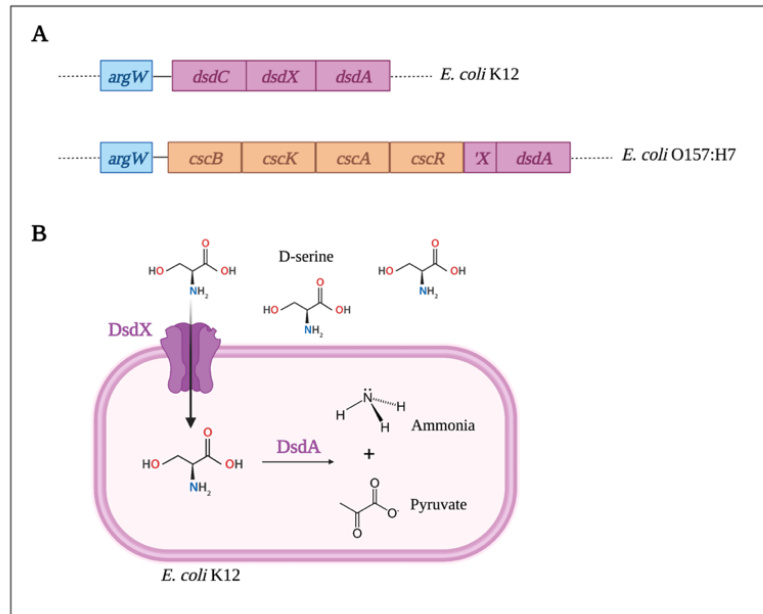


Figure 22 Schematic illustration of the differences between the *argW* region and D-serine catabolism pathway. (A) Differences detected on the *argW* region of *E. coli* K12 and O157:H7. The EHEC strain is unable to tolerate D-Ser due to an insertion that truncates the D-Ser catabolism locus whilst the locus is intact on the K12 strain. (B) Schematic illustrating of the D-serine catabolism pathway in *E. coli* K12. Adapted from O'Boyle et al., 2020 ; Jahreis et al., 2002. Figure made with biorender.com.

In addition, the ability to use D-serine by CD (and some HI) strains was systematically associated with the presence of the *dsd* locus without exception. Besides, some strains (205, V3, D2, F3, M2) presented a *dsd* locus but were able to consume both D-serine and sucrose. This could be explained by the fact that sucrose catabolism is not strictly dependent of the presence of the *csc* locus. Indeed, *scrB* gene encoding a sucros-6-phosphate hydrolase, can be found in *Enterobacteriaceae* including *E. coli* either on plasmids or on mobile elements within the chromosome. Moreover, our OmniLog results showed that raffinose was also better catabolized by HI strains than CD strains. The fact that sucrose and raffinose phenotypes are correlated in our HI collection make sense. Indeed, raffinose is a trisaccharide composed of galactose, glucose, and fructose while sucrose is composed of glucose and fructose. Raffinose can be processed by an α -galactosidase to produce sucrose and D-galactose.

The concentration of D-serine in the gut is thought to be around 1 μ M and the daily intake of D-amino acids from food around 100 mg per day [319]. D-amino acids can be found in a variety of nutrients such as vegetables, fruits, or rice, and at high concentrations in some aquatic invertebrates, and in fermented foods like cheese, sour milk and black beans [322]. In my PhD, we have demonstrated that most *E. coli* strains isolated from the gut of CD patients are able to catabolize D-serine and that this phenotype confers a competitive advantage of the CD strain

204 over HI strains T1, R2 and A1 *in vitro*, although we failed to demonstrate this advantage *in vivo* in our model of gnotobiotic rats. From these data, we can speculate that, if D-serine concentration is higher in the gut of CD patients, this would select strains having the capacity to metabolize D-serine and potentially lead to an expansion of such strains in the gut of CD patients. Interestingly, a recent work linked D-serine, IBD and inflammation in a mouse model since the authors demonstrated that the oral administration of D-serine prevented the onset and progression of colitis in mice [323]. Another study also revealed that dietary L-serine conferred a competitive fitness advantage to AIEC but only in the inflamed gut [294]. Similarly to L-serine, utilization of the D-enantiomer by *E. coli* might be associated with inflammatory conditions of the gut, explaining why we failed to observe a fitness advantage of the CD strain in our experimental rat model. In the future, it would be interesting to perform competition assays in rodents treated with dextran sodium sulfate to induce colitis as well as to quantify D-serine in the digestive tract to evaluate a potential correlation between inflammation and D-serine concentration in the gut. D-serine has been also described as a signal molecule sensed by gut pathogens to regulate gene expression [324]. It would be interesting to evaluate if such mechanisms also exist in commensal *E. coli*, and to determine the role of D-serine perception on the expression of genes potentially involved in the fitness of CD strains.

Our results can be compared with a study where the authors isolated commensal *E. coli* strains from fecal samples of young children suffering from cystic fibrosis and showed an increased growth rate in the presence of glycerol as a sole carbon source [291]. Glycerol is present in high concentrations in the gut of patients with cystic fibrosis and selection of *E. coli* strains efficiently consuming glycerol probably explain the massive expansion of *E. coli* in the gut of cystic fibrosis patients. In this case, it seems to be a genetic adaptation of *E. coli* strains because authors identified adaptive mutations allowing a very efficient utilization of glycerol. In our case, high prevalence of D-serine⁺ strains in the gut of CD patients appears to be more related to a selection of strains rather than a genetic adaptation of resident strains to the environment. Even if the *argW* cluster is prone to genetic rearrangements, presence of the *dsd* cluster conferring D-serine catabolism phenotype corresponds to the ancestral configuration of the *argW* region.

Taken together, our work highlights an increase of prevalence of *dsd* cluster in CD strains and suggests that this phenotypic trait potentially confers a competitive advantage over other members of the gut microbiota, contributing to the bloom of *Enterobacteriaceae*. Further works are however warranted to confirm such hypothesis.

Part 2 : Focus on the EA catabolism pathway in commensal *E. coli* from HI collection.

The second aim of my PhD project was to explore the phenotypic diversity of EA catabolism in commensal *E. coli* strains. Our work highlighted the existence of 3 different EA utilization profiles among our HI collection of strains showing a great phenotypic diversity of EA utilization. Indeed, our findings on EA phenotypes in commensal *E. coli* reflects the high variability of *E. coli* in terms of phenotypic traits. This bacterium has a great genomic plasticity and *Enterobacteriaceae* family has been shown to be the highest contributor to gene diversity in the human gut [249].

In this part of my PhD work, I focused only on commensal *E. coli* from the HI collection. Among our collection, strains that could not catabolize EA had a different structure of the *eut* locus or a default on its expression. Yet our study was not able to demonstrate what differs between strains able to use EA as a sole nitrogen source and the 3 strains capable of using EA as a nitrogen and carbon source. No differences were detected on the organization of the *eut* locus or the expression of the tested genes. It would be interesting to explore if other mechanisms not yet identified are involved in EA metabolism for instances on the sequestration of the compound itself and/or the co-factor vitamin B₁₂. As there is a great phenotypic diversity of EA catabolism among commensal *E. coli*, maybe the bacterial microcompartments structure varies between the EA utilization profiles. But these remain only a hypothesis and further work in this field is needed to answer these questions. Our studies also demonstrated that *in vivo* utilization of EA brings a fitness advantage to commensal *E. coli*. Using the new approach developed during my PhD of a dual-fluorescence reporter system, further studies could be carried *in vivo* to follow EA utilization in the mouse gut both in time and space, in homeostatic or dysbiotic conditions.

Several studies in the literature and previous work on our laboratory had focused on EA utilization by pathogenic strains of *E. coli*. Indeed, EA has been mainly described by its central role on adaptation of pathogenic strains in the gut [325],[326]. EA can serve as a carbon and/or nitrogen source to promote growth as well as a signal to influence virulence during host infection [327]. Although the presence of the *eut* locus is widespread on bacteria, less attention has been given to the role of this compound on commensal *E. coli* strains. Although a recent study from Rowley *et al.* demonstrates that two commensal strains grow more rapidly and even outcompete the pathogen *E. coli* O157:H7 specifically when EA is provided as the sole nitrogen source [298]. These findings and my PhD work suggest a more extensive role for EA utilization in commensal *E. coli* strains. A better understanding of EA catabolism in commensal *E. coli* could open doors on new approaches to limit intestinal infection by pathogens for which EA

utilization is important such as EHEC, *Salmonella* or *Listeria*. Indeed, EA⁺ commensal strains identified thanks to my work, could be evaluated in the future as competitive strains to limit gut colonization by intestinal pathogens and associated disease.

Even though EA was not present in plates PM01 and PM02 tested with the OmniLog technique, we tested our collection of CD and RA strains to the utilization of this compound. We observed that most strains consume EA as a nitrogen source whatever the strain groups but the phenotype able to use EA as a carbon source was absent among our CD and RA collection of strains. However, we should keep in mind that such absence potentially results from the small number of strains in CD and RA collections. Interestingly EA is a compound that was already highlighted in the literature to have its importance on both pathologies. Moreover, EA is abundant in the inflamed gut since desquamation of epithelial cells in the lumen is more important. In a study from 2019, Ormsby *et al.* determined that the intestinal SCFA propionic acid acts as a signal for AIEC and alters their metabolism and increase their use of EA [326]. To date, the rapid replication of AIEC in macrophages has been unexplained, but apparently growth can be facilitated by the presence of extracellular levels of EA comparable to those in the human gut. On the other hand, Shan *et al.* (2018) identified EA as a differential marker present on the gut of rats developing RA in comparison to the control group [328]. This study even highlighted a correlation between the increase of *Enterobacteriaceae* observed in the *in vivo* model and the variation of the EA concentration in the gut. The role of EA seems thus to have its importance in the physiopathology of both diseases.

In summary, the aim of my PhD work was to evaluate if metabolic abilities differ between *Enterobacteriaceae* isolated from HI or from CD or RA patients as a hypothesis to explain their expansion in the gut of patients. Among substrates differentially used between HI and disease-associated strains, the preferential utilization of D-serine by CD strains is a promising hypothesis to follow. More studies are needed though, in order to assess if D-serine confers a nutritional advantage to CD strains in the gut and in which conditions it occurs. Likewise, it would be of main interest to analyze the efficiency of CD and RA strains to use EA over HI strains, a characteristic that could participate to a better fitness of CD or RA strains over HI strains. Understand the mechanisms by which *Enterobacteriaceae* proliferates in the gut of patients is a prerequisite to envisage strategies aiming at limiting their expansion and by extension, the associated harmful consequences for the host.

SCIENTIFIC CONTRIBUTIONS

PUBLICATIONS

Moreira de Gouveia MI, Daniel J, Garrivier A, Bernalier-Donadille A, Jubelin G. Diversity of ethanolamine utilization by human commensal *Escherichia coli*. Res Microbiol. 2022 Aug 18:103989. doi: 10.1016/j.resmic.2022.103989. Epub ahead of print. PMID: 35988812.

Jubelin G, **Moreira de Gouveia MI**, Bernalier-Donadille A. Microbiote et maladies infectieuses intestinales. Chapitre 20 de l'ouvrage « Microbiote intestinal et santé humaine » (2021). Elsevier Masson.

ORAL PRESENTATIONS

Moreira de Gouveia MI, Daniel J, Garrivier A, Bernalier-Donadille A, Jubelin G. Ethanolamine catabolism by commensal *Escherichia coli* strains *in vitro* and *in vivo*. 18ème Rencontre des Microbiologistes du Pôle Clermontois, April 7th 2022, Clermont-Ferrand, France.

Moreira de Gouveia MI, Daniel J, Garrivier A, Bernalier-Donadille A, Jubelin G. Diversity of ethanolamine utilization by commensal *Escherichia coli*. **Doctoral School Days at Clermont Auvergne University**, May 18th 2022, Clermont-Ferrand, France.

POSTERS

Moreira de Gouveia MI, Daniel J, Garrivier A, Bernalier-Donadille A, Jubelin G. Ethanolamine catabolism by *Escherichia coli* commensal strains. **SYMPOSIUM**, October 13th-15th 2021, Online.

Moreira de Gouveia MI, Daniel J, Garrivier A, Bernalier-Donadille A, Jubelin G. Ethanolamine catabolism by *Escherichia coli* commensal strains. **SFM**, October 10th-12th 2022, Montpellier, France.

Moreira de Gouveia MI, Garrivier A, Daniel J, Bernalier-Donadille A, Jubelin G. Design and validation of a dual-fluorescence reporter system to monitor bacterial gene expression *in situ*. **SFM**, October 10th-12th 2022, Montpellier, France.

ANNEX

Microbiote et maladies infectieuses intestinales

CHAPITRE

20

Grégory Jubelin, Maria-Ines Moreira De Gouveia,
Annick Bernalier-Donadille

Points forts

Le microbiote intestinal joue un rôle majeur dans la protection de l'hôte vis-à-vis des agents pathogènes exogènes ou présents en faible quantité dans le côlon (pathobiontes). Les mécanismes impliqués dans cette résistance à la colonisation du microbiote intestinal sont multiples :

- compétition pour les nutriments ou les sites d'adhésion entre espèces commensales et pathogènes ;
- productions de composés antimicrobiens (bactériocines, acides gras à chaîne courte, acides biliaires, etc.) ;
- activation des cellules de l'hôte : stimulation de la production de mucus, renforcement des jonctions serrées, etc. et stimulation du système immunitaire muqueux.

Les agents pathogènes stricts ont, quant à eux, développé des stratégies pour contrer l'effet barrière du microbiote intestinal :

- induction d'un état inflammatoire ;
- utilisation de nutriments alternatifs non utilisés par les espèces commensales ;
- production de molécules antimicrobiennes ciblant les espèces commensales.

Les déséquilibres du microbiote intestinal (dysbioses), qui peuvent apparaître de façon transitoire (exemple : après une antibiothérapie) ou être présents de manière chronique (exemple : maladies inflammatoires), rendent l'hôte plus susceptible aux infections par :

- les agents pathogènes externes, dont l'effet sera maximisé par la présence d'un microbiote dysbiotique ;
- les pathobiontes, qui profitent de la dysbiose pour occuper la niche écologique d'autres espèces commensales affectées et dont le niveau de population va donc augmenter.

Le microbiote intestinal est une communauté microbienne complexe hébergée par le tractus gastro-intestinal depuis la naissance. Cet écosystème se compose de diverses espèces bactériennes ayant de multiples activités métaboliques et immunologiques, potentiellement bénéfiques pour la santé humaine. Le microbiote intestinal développe ainsi des relations symbiotiques avec l'organisme en étant étroitement lié à l'épithélium intestinal et à la réponse immunitaire innée, avec lesquels il contribue à former une barrière complexe contre les bactéries pathogènes intestinales.

L'une des fonctions majeures du microbiote intestinal est en effet d'assurer la protection contre les agents pathogènes, concept connu sous le nom de résistance à la colonisation [1]. À l'état d'équilibre (eubiose), chez les individus en bonne santé, le microbiote intestinal déploie divers mécanismes, comme la sécrétion de produits antimicrobiens, la compétition pour certains nutriments, le maintien de l'intégrité de la barrière épithéliale et l'activation immunitaire, pour protéger l'organisme contre l'infection. Cependant, certains agents pathogènes dits « stricts » sont toutefois capables, lorsque les conditions sont réunies, de surmonter l'effet protecteur du microbiote et de provoquer des infections gastro-intestinales. Par ailleurs, le microbiote intestinal d'un individu peut subir au cours de sa vie des altérations (dysbioses) en réponse à différents facteurs tels que l'alimentation, le stress, la maladie, la prise de médicaments ou bien encore le changement brutal d'environnement. Ces modifications sont susceptibles de favoriser la multiplication de bactéries pathogènes ou de pathobiontes,

micro-organismes sous-dominants du microbiote qui, en proliférant, sont capables d'induire une pathologie. Les dysbioses du microbiote intestinal observées dans de nombreuses pathologies digestives et extradiigestives [2] peuvent ainsi être des facteurs de susceptibilité aux infections.

Ce chapitre présente les mécanismes de la résistance à la colonisation développés par le microbiote intestinal, les stratégies des bactéries pathogènes pour surmonter cette résistance et finalement provoquer une infection entérique, ainsi que l'impact des dysbioses du microbiote intestinal dans la susceptibilité à l'infection.

Mécanisme de la résistance à la colonisation du microbiote intestinal

Les mécanismes de la résistance à la colonisation du microbiote intestinal (figure 20.1) incluent la compétition pour les nutriments ou les sites d'adhésion avec les agents pathogènes et la production de divers composés ayant des effets antimicrobiens. Le microbiote agit également sur les cellules de l'hôte et stimule indirectement la barrière intestinale en permettant le maintien de la couche de mucus protectrice de l'épithélium ou en contribuant à la réparation de ce dernier en cas

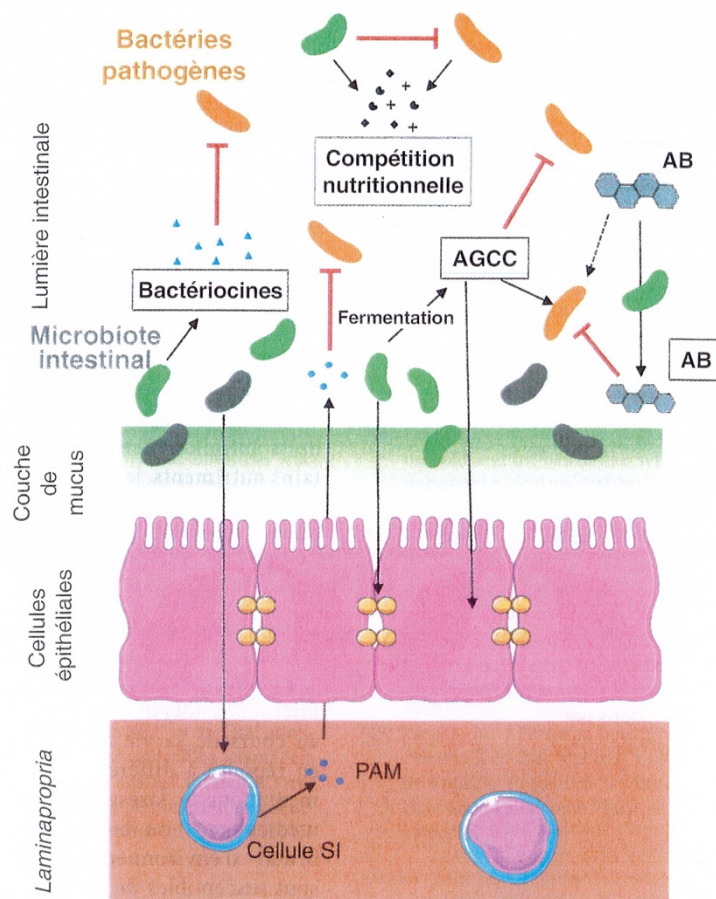


Figure 20.1. Principaux mécanismes de la résistance à la colonisation du microbiote intestinal. PAM : peptide antimicrobien ; SI : système immunitaire ; AB : acides biliaires ; AGCC : acides gras à chaîne courte.

d'agression. Enfin, les interrelations du microbiote avec le système immunitaire inné ne seront pas décrites ici, celles-ci étant développées dans le chapitre 3.

Compétition pour les substrats avec les agents pathogènes

Les bactéries composant le microbiote intestinal sont adaptées à leur environnement : elles utilisent les nutriments disponibles, occupent les sites d'adhésion et modulent l'écosystème à leur avantage grâce aux produits de leur métabolisme. Pour qu'une bactérie exogène puisse coloniser l'écosystème, il est nécessaire qu'elle possède un avantage écologique lui permettant d'occuper une niche de façon plus efficace que les espèces commensales. Toute perturbation du microbiote intestinal, comme celle observée lors d'une antibiothérapie, sera ainsi susceptible de favoriser l'agent pathogène au détriment des espèces commensales. Dans ce contexte, la compétition pour l'utilisation des nutriments présents dans l'intestin joue un rôle important dans le mécanisme de résistance à la colonisation. Ceci a été en particulier décrit pour les souches bactériennes appartenant à la même espèce qui nécessitent souvent les mêmes substrats de croissance. Par exemple, les souches commensales d'*Escherichia coli* entrent en compétition avec les *E. coli* entérohémorragiques (*enterohemorrhagic Escherichia coli* [EHEC]) pour l'acide aminé proline [3]. La compétition nutritionnelle ne se limite pas aux macronutriments mais peut également intéresser les micronutriments, comme le fer dans le cas d'infection à *Salmonella enterica* serovar *Typhimurium* (*S. typhimurium*) [1].

Production de composés antimicrobiens par le microbiote intestinal

Acides gras à chaîne courte

Les acides gras à chaîne courte (AGCC), produits par les bactéries lors de la fermentation des substrats non digestibles, sont principalement l'acétate, le propionate et le butyrate. Ces AGCC jouent un rôle trophique sur la muqueuse intestinale : ils permettent la prolifération des cellules intes-

tinales et maintiennent l'intégrité des jonctions épithéliales serrées. En diminuant la perméabilité épithéliale, les AGCC s'opposent à la translocation de bactéries et du lipopolysaccharide (LPS). Ainsi, l'acétate produit par les bifidobactéries inhibe l'apoptose des cellules épithéliales et empêche la translocation des shigatoxines produites par les EHEC dans le sang [4]. Les AGCC peuvent également altérer la croissance bactérienne en affectant le pH intracellulaire et le fonctionnement métabolique [5]. Les AGCC peuvent en effet diffuser dans le cytoplasme où ils se dissocient en ions, entraînant une baisse du pH intracellulaire. Ainsi, le métabolisme d'*E. coli* pourrait être altéré en présence d'acétate en empêchant la biosynthèse de la méthionine, entraînant une accumulation d'homocystéine toxique et une inhibition de la croissance [6]. Le butyrate pourrait toutefois favoriser l'infection par certains pathogènes en modulant l'expression des récepteurs cellulaires (récepteurs de la shigatoxine), des facteurs de virulence ou des facteurs d'adhésion aux cellules épithéliales [5]. Certaines espèces commensales pourraient donc soutenir des bactéries pathogènes via la libération d'AGCC, alors que d'autres renforceraient la fonction de barrière de l'épithélium intestinal par cette même production d'AGCC.

Acides biliaires

Les acides biliaires (AB), qui possèdent des propriétés antimicrobiennes, sont produits par le foie et excrétés dans le tractus intestinal pour faciliter la digestion des lipides alimentaires. Le foie synthétise l'acide cholique et l'acide chénodésoxycholique, AB primaires qui sont déconjugués par le microbiote intestinal grâce à des hydrolases spécifiques (BSH) et convertis en AB secondaires, que sont l'acide désoxycholique et l'acide lithocholique [5]. L'acide désoxycholique est bactéricide pour de nombreuses espèces pathogènes, comme *Staphylococcus aureus* et *Clostridium difficile*, mais aussi des espèces commensales comme les bifidobactéries et les lactobacilles, par rupture de la membrane et perte subéquente du contenu cellulaire. Très peu d'espèces (moins de 0,025 % du microbiote) sont capables de réaliser la 7-déhydroxylation des AB. Cette déconjugaison est essentiellement retrouvée chez les bacilles à Gram positif, en particulier des espèces commensales de *Clostridium* et de *Lactobacillus* et assure la résistance à la colonisation envers d'autres

espèces [7]. Il a ainsi été montré que *Clostridium scindens* induisait une résistance à la colonisation de *C. difficile* chez la souris [7]. Ces espèces productrices de BSH sont également protégées de la toxicité des AB et sont donc avantagées dans la compétition pour la survie [1].

Bactériocines

Les bactériocines sont des polypeptides toxiques qui sont produits par certaines bactéries pour inhiber la colonisation et la croissance d'autres espèces [8] par de multiples mécanismes d'action. Elles sont structurellement et fonctionnellement très diverses. Les bactériocines des bactéries à Gram négatif sont principalement produites par des bactéries lactiques (par exemple : *Lactococcus* et *Lactobacillus*) et certaines espèces de *Streptococcus* [9]. Les bactériocines produites par les bactéries à Gram négatif le sont principalement par les entérobactéries et peuvent être divisées en protéines de poids moléculaire élevé (colicines) et peptides de poids moléculaire inférieur (microcines) [10]. Les bactériocines jouent un rôle dans la régulation des populations commensales du microbiote mais elles peuvent également intervenir dans les mécanismes de résistance à la colonisation, en inhibant les bactéries pathogènes.

Parmi les espèces productrices de bactériocines, les plus étudiées sont celles issues de la famille des *Lactobacillaceae* et du genre *Bifidobacterium*. Les espèces bactériennes de ces groupes produisent plusieurs types de bactériocines [11]. Ces bactériocines limitent de manière significative la colonisation de nombreux pathogènes comme les entérocoques multirésistants, *Listeria monocytogenes*, *C. difficile*, *S. aureus* et autres [1]. Certains probiotiques libèrent ainsi des bactériocines responsables de l'élimination de certaines espèces pathogènes et de la restauration des communautés intestinales commensales.

Effet protecteur du microbiote intestinal via l'hôte

Modulation de la production de mucus

Le mucus se compose d'une couche interne et d'une couche externe. La couche de mucus interne, proche des cellules de l'épithélium, est impéné-

trable et forme une barrière physique contre les bactéries, en empêchant toute interaction directe avec la couche épithéliale [5]. À l'opposé, certaines espèces commensales métabolisent les nutriments dans la couche de mucus externe où elles résident (*Akkermansia muciniphila*, *Bacteroides caccae*, etc.). Une réduction de la couche de mucus, comme observé lors d'une alimentation de type occidental carencée en glucides indigestibles, conduit à une susceptibilité accrue à la colonisation par des agents pathogènes [12].

Le maintien de la couche de mucus est partiellement sous le contrôle du microbiote intestinal. Les bactéries commensales, et en particulier leurs métabolites fermentaires, régulent en effet l'expression des gènes codant les mucines (muc2). Ainsi, le butyrate stimule la production de ces mucines via muc2 [13]. L'administration de *Bifidobacterium longum* permettrait de réduire les dommages causés à la couche de mucus, probablement en stimulant la synthèse de ce composé [14]. Enfin, le microbiote intestinal peut influencer également l'activité des glycosyltransférases cellulaires, modifiant ainsi la glycosylation des mucines et donc leur efficacité contre les pathogènes [13].

Maintien de l'intégrité de l'épithélium intestinal

Afin de garantir l'étanchéité de la paroi intestinale, les cellules épithéliales sont scellées par des jonctions serrées (complexes protéiques). Ces jonctions empêchent la diffusion paracellulaire de micro-organismes commensaux ou pathogènes. Le microbiote intestinal contrôle partiellement l'intégrité de cette barrière, en particulier en influençant l'expression des gènes codant pour la claudine, protéine majeure des jonctions serrées, ainsi que la distribution de ces jonctions. L'expression de la claudine est en effet diminuée chez les souris axéniques [13]. Certains pathogènes, comme les *E. coli* entéro-pathogènes (*enteropathogenic Escherichia coli* [EPEC]), peuvent altérer l'intégrité des jonctions serrées alors que certains probiotiques (*Saccharomyces boulardii*, *E. coli* Nissle) peuvent, au contraire, induire l'expression des protéines de jonctions et/ou prévenir les altérations induites au niveau de ces jonctions par un pathogène.

Le microbiote intestinal contribue également largement au processus de réparation épithéliale

suite à une agression [13]. Bien que les mécanismes ne soient pas complètement connus à ce jour, ils impliquent la production de cytokines et la libération de facteurs cicatrisants via des voies de signalisation impliquant des récepteurs spécifiques (TLR).

Microbiote intestinal et infection

Les agents bactériens responsables d'infection intestinale chez l'homme peuvent être classifiés en deux groupes :

- les pathogènes dits « stricts » tels que *Salmonella enterica*, EPEC ou *Campylobacter jejuni*, qui sont capables de provoquer une pathologie chez des individus sains possédant un microbiote intestinal à l'état d'eubiose;
- les pathogènes dits « opportunistes » ou « pathobiontes » tels que *Clostridium difficile* ou *E. coli* adhérent-invasif (AIEC), à l'origine d'infections consécutives à un déséquilibre du microbiote intestinal.

Infections à l'état d'eubiose

Au cours des dernières années, l'utilisation de rongeurs comme modèle animal d'infection a permis de comprendre comment les pathogènes stricts réussissent à surmonter l'effet barrière du microbiote intestinal et induire une pathologie chez un individu en bonne santé. Trois principales stratégies ont été décrites : induction d'un état inflammatoire, utilisation de nutriments alternatifs et production de molécules antimicrobiennes.

Induction d'un état inflammatoire

Il a été clairement démontré que la virulence de plusieurs pathogènes stricts est étroitement corrélée à leur capacité à induire une inflammation du tractus gastro-intestinal. C'est typiquement le cas de *S. typhimurium*, l'agent responsable de la salmonellose (infection alimentaire provoquant fièvre, diarrhée, vomissements et douleurs abdominales). L'expression de ses facteurs de virulence, en particulier deux systèmes de sécrétion de type 3, permet à *S. typhimurium* d'envahir l'épithélium intestinal et provoque une forte réaction inflammatoire caractérisée par une recrudescence

de cellules immunitaires sur le site d'infection. Cette inflammation contribue à l'expansion du pathogène qui peut représenter jusqu'à 90 % des bactéries intestinales au pic de l'infection [15], affectant ainsi grandement l'homéostasie du tractus digestif. L'inflammation se manifeste notamment par le relargage d'espèces réactives de l'oxygène (*reactive oxygen species* [ROS]) ou de l'azote (*reactive nitrogen species* [RNS]) dans la lumière intestinale qui vont, d'une part, affaiblir de nombreuses bactéries commensales du microbiote en tant que molécules antibactériennes et, d'autre part, former des composés tels que le tétrathionate ou le nitrate. Ces molécules sont utilisées par *Salmonella* en tant qu'accepteur d'électrons pour réaliser de la respiration anaérobie, un procédé plus efficace que la fermentation réalisée par la plupart des bactéries du microbiote. Ces ROS/RNS vont également oxyder des glucides présents dans l'intestin et ces composés oxydés, en particulier le glucarate et le galactarate, sont efficacement consommés par l'agent pathogène, favorisant ainsi son développement [16]. L'inflammation va également déplacer le métabolisme des cellules épithéliales intestinales de la phosphorylation oxydative mitochondriale vers la glycolyse aérobie limitant ainsi la consommation de dioxygène (O_2) par l'épithélium. La diffusion d' O_2 dans le tractus intestinal va en conséquence fortement affecter la multiplication de nombreuses bactéries commensales anaérobies strictes et favoriser la prolifération des bactéries anaérobies facultatives telles que *Salmonella*.

Utilisation de nutriments alternatifs

Le tractus gastro-intestinal, en particulier le côlon, est un écosystème complexe caractérisé par la présence d'une variété importante de molécules résultant des interactions entre les aliments, le microbiote et la muqueuse intestinale (dégradation, transformation, absorption, etc.). Au cours de la dernière décennie, plusieurs études ont montré que les pathogènes intestinaux stricts sont capables d'utiliser des nutriments peu ou pas catabolisés par les bactéries du microbiote, leur conférant ainsi un avantage sélectif pour coloniser la niche écologique. En particulier, certains substrats provenant de la muqueuse intestinale tels que l'acide sialique et l'éthanolamine, ou générés par le microbiote

intestinal tel que le 1,2-propanediol, sont efficacement consommés par *S. typhimurium* [17,18]. Dans le cas d'une infection à EHEC, bactérie responsable de troubles intestinaux pouvant s'accompagner d'un dysfonctionnement rénal grave, l'agent pathogène a la capacité d'utiliser certains glucides ou acides aminés de façon plus efficace que la plupart des *E. coli* commensaux du microbiote intestinal [19]. Ces particularités métaboliques confèrent un avantage compétitif des agents pathogènes vis-à-vis des bactéries du microbiote intestinal, contribuant ainsi à leur expansion dans le tractus gastro-intestinal au cours de l'infection.

Production de molécules antimicrobiennes

À l'instar des bactériocines produites par les bactéries commensales, certaines bactéries pathogènes synthétisent également des molécules inhibitrices affectant directement la prolifération ou la survie des bactéries commensales. Dans le cas des infections à *L. monocytogenes*, l'agent responsable de la listériose, le pathogène synthétise une bactériocine, la listériolysine S, qui affecte la croissance de différents genres bactériens. Dans un modèle infectieux de souris, la production de listériolysine S est associée à une altération du microbiote intestinal et une plus forte expansion de *L. monocytogenes* dans le tractus gastro-intestinal, aboutissant à une pathologie plus prononcée chez l'animal infecté [20]. De nombreux pathogènes produisent également une machinerie particulière, le système de sécrétion de type 6, qui permet d'injecter des molécules toxiques directement à l'intérieur de bactéries avoisinantes. Ce système est exprimé par *S. typhimurium* et *Vibrio cholerae*, l'agent responsable du choléra, au cours de leur processus infectieux et, à travers son activité antibactérienne, il limite l'effet barrière du microbiote intestinal et favorise le développement du pathogène [21,22].

Infections consécutives à un déséquilibre du microbiote intestinal

Au cours de la vie, le microbiote intestinal d'un individu peut subir des altérations dans certaines

situations. En fonction des facteurs de déséquilibre rencontrés, le degré ainsi que la durée de la dysbiose intestinale seront plus ou moins prononcés. Parmi les troubles associés aux déséquilibres du microbiote intestinal, il est désormais clairement établi que cette dysbiose peut conduire à l'émergence de populations bactériennes à l'origine de pathologies digestives.

Dysbiose transitoire et infection

Dans un contexte de dysbiose transitoire, par exemple lors d'un traitement antibiotique par voie orale, le déséquilibre du microbiote intestinal peut se manifester par une multiplication démesurée de *Clostridium difficile* (renommé en 2016 *Clostridioides difficile*), une bactérie fréquemment retrouvée en faible proportion dans le tractus digestif de l'homme. *C. difficile* est une bactérie sporulante dont certaines souches produisent des toxines appelées Tcd à l'origine de symptômes cliniques tels que diarrhée et inflammation aiguë du côlon. Ces dernières années, plusieurs travaux ont identifié les mécanismes expliquant pourquoi une dysbiose intestinale peut conduire à une prolifération excessive de *C. difficile*. En particulier, la germination des spores de *C. difficile* est induite par la présence d'acides biliaires (AB) primaires tels que le cholate. Or, ces AB sont normalement transformés par le microbiote intestinal en AB secondaires qui, au contraire, ont un effet inhibiteur sur la multiplication de *C. difficile* (figure 20.1). Au cours d'une dysbiose intestinale, l'altération de ce processus de transformation conduit à une augmentation du ratio AB primaires/AB secondaires et favorise donc la germination des spores et la multiplication des cellules végétatives de *C. difficile* [23]. D'un point de vue métabolique, la dysbiose intestinale va également altérer la disponibilité de nombreux substrats et ceci participe à l'émergence de *C. difficile*. Il a par exemple été montré dans un modèle de souris que l'expansion de *C. difficile* observée suite à un traitement antibiotique implique l'utilisation d'acide sialique et de succinate par le pathobionte, deux métabolites dont la concentration augmente au cours de la dysbiose [24,25]. Un autre élément participant aussi à la multiplication de *C. difficile* lors d'une dysbiose concerne la réduction

de bactéries du genre *Bacillus*, *Lactococcus* et *Enterococcus* connues pour produire des molécules antimicrobiennes efficaces contre *C. difficile* [26].

Dysbiose chronique et infection

Pour de nombreuses pathologies chroniques, en particulier inflammatoires, le microbiote intestinal des individus est durablement altéré, avec en général une réduction globale de la diversité bactérienne, une diminution d'espèces bactériennes considérées comme bénéfiques (*Faecalibacterium*, *Roseburia*) et une augmentation significative des *Proteobacteria* telles qu'*E. coli* [27]. Dans le cas de la maladie de Crohn, une pathologie inflammatoire de l'intestin, des souches particulières d'*E. coli* sont retrouvées au niveau des muqueuses iléale et colique chez 20 à 60 % des patients alors que ces souches sont peu associées aux muqueuses de l'individu sain (0–20 % des individus) [28]. Ces isolats, appelés AIEC, ont la capacité d'envahir l'épithélium intestinal, et ceci participe au maintien de l'état inflammatoire de l'intestin, à la déstabilisation de la barrière épithéliale et à la fibrose des tissus [29]. Même si la forte prévalence des AIEC chez les patients atteints de maladie de Crohn n'est pas encore totalement expliquée à ce jour, différents facteurs semblent participer à l'expansion de ce pathobionte chez les patients. Les AIEC tirent notamment avantage de l'inflammation intestinale à travers :

- l'augmentation de la teneur en oxygène dans le tractus digestif, une condition propice au développement des bactéries aérotolérantes tels qu'*E. coli* [30];
- l'altération importante du microbiote intestinal, ce qui réduit drastiquement la résistance à la colonisation;
- l'utilisation efficace de substrats en conditions inflammatoires tels que l'éthanolamine ou la L-sérine [28,31].

Certains déterminants génétiques sont également critiques dans la capacité des AIEC à coloniser la muqueuse intestinale avec notamment la protéase Vat qui dégrade le mucus ou encore les pili de type 1 qui assurent l'adhésion puis l'invasion des entérocytes [29]. De façon intéressante, la forte capacité des AIEC à adhérer aux entérocytes

provient de mutations récentes du gène *fimH* codant l'adhésine des pili de type 1, suggérant l'implication de mutations patho-adaptatives dans le pouvoir virulent de ce pathobionte [32].

Conclusion

Le microbiote intestinal contribue grandement à protéger l'hôte des agents pathogènes, en agissant directement contre ces derniers ainsi qu'en stimulant la réponse de l'hôte. Les agents pathogènes stricts ont toutefois développé des stratégies pour contrer cet effet et conduire à la pathologie infectieuse. Les états de dysbiose du microbiote, transitoires ou chroniques, rendent également l'hôte plus susceptible à l'infection par des pathogènes externes ou des pathobiontes. Afin de lutter contre ces infections, l'utilisation de probiotiques capables de limiter l'inflammation, d'utiliser des substrats préférentiels des pathogènes et/ou de produire des molécules antibactériennes sont des pistes à investiguer. Celles-ci seraient particulièrement adaptées en cas de dysbiose transitoire du microbiote (par exemple : traitement antibiotique). Dans les cas d'infections sévères et résistantes aux traitements (par exemple : infection à *C. difficile*), la transplantation fécale s'avère un outil thérapeutique essentiel.

Références

- [1] Jacob S, Jacob DG, Luminos LM. Intestinal Microbiota as a host defense mechanism to infection threats. *Front Microbiol* 2019;10:1676.
- [2] Weiss GA, Hennes T. Mechanisms and consequences of intestinal dysbiosis. *Cell Mol Life Sci* 2017;74:2959–77.
- [3] Momose Y, Hirayama K, Itoh K. Competition for proline between indigenous *Escherichia coli* and *E. coli* O157:H7 in gnotobiotic mice associated with infant intestinal microbiota and its contribution to the colonization resistance against *E. coli* O157:H7. *Antonie Van Leeuwenhoek* 2008;94:165–71.
- [4] Fukuda S, Toh H, Hase K, et al. Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* 2011;469:543–7.
- [5] Ducarmon QR, Zwittink RD, Hornung BVH, et al. Gut microbiota and colonization resistance against bacterial enteric infection. *Microbiol mol Biol Rev* 2019;83:e00007–19.
- [6] O'Byrne C, Booth IR, Roe AJ, et al. Inhibition of *Escherichia coli* growth by acetic acid a problem with

- methione biosynthesis and homocysteine toxicity. *Microbiology* 2002;148:2215–22.
- [7] Buffie CG, Bucci V, Stein RR, et al. Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* 2015;517:205–8.
- [8] Rea MC, Sit CS, Clayton E, et al. Thuricin CD, a post translationally modified bacteriocin with a narrow spectrum of activity against *Clostridium difficile*. *Proc Natl Acad Sci* 2010;107:9352–7.
- [9] Rea MC, Ross RP, Cotter PD, et al. Classification of bacteriocins from Gram-positive bacteria. In: Drider D, Rebuffat S (éd.). *Prokaryotic antimicrobial peptides: from genes to applications*. New York: Springer; 2011. p. 29–53.
- [10] Rebuffat S. Bacteriocins from Gram-negative bacteria: a classification? In: Drider D, Rebuffat S (éd.). *Prokaryotic antimicrobial peptides: from genes to applications*. New York: Springer; 2011. p. 55–72.
- [11] Kleerebezem M. Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis. *Peptides* 2004;25:1405–14.
- [12] Desai MS, Seekatz AM, Koropatkin NM, et al. A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell* 2016;167:1339–53.
- [13] Sokol H. Microbiote et effet de barrière. In: Marteau P, Doré J (éd.). *Le Microbiote intestinal, un organe à part entière*. Montrouge: John Libbey Eurotext; 2017. p. 65–9.
- [14] Schroeder BO, Birchenough GMH, Stahlman M, et al. Bifidobacteria or fiber protects against diet-induced microbiota-mediated colonic mucus deterioration. *Cell Host Microbe* 2018;23:27–40.
- [15] Stecher B, Robbiani R, Walker AW, et al. *Salmonella enterica* serovar typhimurium exploits inflammation to compete with the intestinal microbiota. *PLoS Biology* 2007;5:2177–89.
- [16] Faber F, Tran L, Byndloss MX, et al. Host-mediated sugar oxidation promotes post-antibiotic pathogen expansion. *Nature* 2016;534:697–9.
- [17] Vonaesch P, Anderson M, Sansonetti PJ. Pathogens, microbiome and the host: Emergence of the ecological Koch's postulates. *FEMS Microbiol Rev* 2018;42:273–92.
- [18] Sorbara MT, Pamer EG. Interbacterial mechanisms of colonization resistance and the strategies pathogens use to overcome them. *Mucosal Immunol* 2019;12:1.
- [19] Ducarmon QR, Zwittink RD, Hornung BVH, et al. Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. *Microbiol Mol Biol Rev* 2019;83:e00007–19.
- [20] Quereda JJ, Dussurget O, Nahori MA, et al. Bacteriocin from epidemic *Listeria* strains alters the host intestinal microbiota to favor infection. *Proc Natl Acad Sci U S A* 2016;113:5706–11.
- [21] Fu Y, Waldor MK, Mekalanos JJ. Tn-seq analysis of *Vibrio cholerae* intestinal colonization reveals a role for T6SS-mediated antibacterial activity in the host. *Cell Host Microbe* 2013;14:652–63.
- [22] MacIntyre DL, Miyata ST, Kitaoka M, et al. The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc Natl Acad Sci U S A* 2010;107:19520–4.
- [23] Abbas A, Zackular JP. Microbe-microbe interactions during *Clostridioides difficile* infection. *Curr Opin Microbiol* 2020;53:19–25.
- [24] Ng KM, Ferreyra JA, Higginbottom SK, et al. Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* 2013;502:96–9.
- [25] Ferreyra JA, Wu KJ, Hryckowian AJ, et al. Gut microbiota-produced succinate promotes *C. Difficile* infection after antibiotic treatment or motility disturbance. *Cell Host Microbe* 2014;16:770–7.
- [26] Schäffler H, Breitrück A. *Clostridium difficile* – From colonization to infection. *Front Microbiol* 2018;9:646.
- [27] Nishida A, Inoue R, Inatomi O, et al. Gut microbiota in the pathogenesis of inflammatory bowel disease. *Clin J Gastroenterol* 2018;11:1–10.
- [28] Ormsby MJ, Logan M, Johnson SA, et al. Inflammation associated ethanalamine facilitates infection by Crohn's disease-linked adherent-invasive *Escherichia coli*. *EBioMedicine* 2019;43:325–32.
- [29] Chervy M, Barnich N, Denizot J. Adherent-invasive *E. coli*: Update on the lifestyle of a troublemaker in Crohn's disease. *Int J Mol Sci* 2020;21:3734.
- [30] Shaler CR, Elhenawy W, Coombes BK. The Unique Lifestyle of Crohn's Disease-Associated Adherent-Invasive *Escherichia coli*. *J Mol Biol* 2019;431:2970–81.
- [31] Kitamoto S, Alteri CJ, Rodrigues M, et al. Dietary l-serine confers a competitive fitness advantage to Enterobacteriaceae in the inflamed gut. *Nat Microbiol* 2020;5:116–25.
- [32] Dreux N, Denizot J, Martinez-Medina M, et al. Point Mutations in FimH Adhesin of Crohn's Disease-Associated Adherent-Invasive *Escherichia coli* Enhance Intestinal Inflammatory Response. *PLoS Pathog* 2013;9, e1003141.

REFERENCES

- [1] C. Landman and E. Quévrain, “[Gut microbiota: Description, role and pathophysiologic implications],” *La Rev. Med. interne*, vol. 37, no. 6, pp. 418–423, Jun. 2016, doi: 10.1016/J.REVMED.2015.12.012.
- [2] P. J. Turnbaugh, R. E. Ley, M. Hamady, C. M. Fraser-Liggett, R. Knight, and J. I. Gordon, “The Human Microbiome Project,” *Nature*, vol. 449, no. 7164, pp. 804–810, 2007, doi: 10.1038/nature06244.
- [3] “Metagenomics of the Human Intestinal Tract | METAHIT Project | Fact Sheet | FP7 | CORDIS | European Commission.” <https://cordis.europa.eu/project/id/201052/fr>. (accessed Oct. 21, 2022).
- [4] I. Rowland *et al.*, “Gut microbiota functions: metabolism of nutrients and other food components,” *Eur. J. Nutr.*, vol. 57, no. 1, Feb. 2018, doi: 10.1007/S00394-017-1445-8.
- [5] S. M. Jandhyala, R. Talukdar, C. Subramanyam, H. Vuyyuru, M. Sasikala, and D. N. Reddy, “Role of the normal gut microbiota,” *World J. Gastroenterol.*, vol. 21, no. 29, p. 8787, Aug. 2015, doi: 10.3748/WJG.V21.I29.8787.
- [6] C. M. Guinane and P. D. Cotter, “Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ,” *Therap. Adv. Gastroenterol.*, vol. 6, no. 4, pp. 295–308, 2013, doi: 10.1177/1756283X13482996.
- [7] E. M. Bik *et al.*, “Molecular analysis of the bacterial microbiota in the human stomach,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 3, p. 732, Jan. 2006, doi: 10.1073/PNAS.0506655103.
- [8] T. C. Martinsen, K. Bergh, and H. L. Waldum, “Gastric juice: a barrier against infectious diseases,” *Basic Clin. Pharmacol. Toxicol.*, vol. 96, no. 2, pp. 94–102, Feb. 2005, doi: 10.1111/J.1742-7843.2005.PTO960202.X.
- [9] J. M. Manson, M. Rauch, and M. S. Gilmore, “The commensal microbiology of the gastrointestinal tract,” *Adv. Exp. Med. Biol.*, vol. 635, pp. 15–28, 2008, doi: 10.1007/978-0-387-09550-9_2.
- [10] H. Sharif, D. Devadason, N. Abreheart, R. Stevenson, and L. Marciani, “Imaging Measurement of Whole Gut Transit Time in Paediatric and Adult Functional Gastrointestinal Disorders: A Systematic Review and Narrative Synthesis,” *Diagnostics (Basel, Switzerland)*, vol. 9, no. 4, Dec. 2019, doi: 10.3390/DIAGNOSTICS9040221.
- [11] S. Sjöstedt, A. Heimdahl, L. Kager, and C. E. Nord, “Microbial colonization of the oropharynx, esophagus and stomach in patients with gastric diseases,” *Eur. J. Clin.*

- Microbiol.*, vol. 4, no. 1, pp. 49–51, Feb. 1985, doi: 10.1007/BF02148660.
- [12] B. Zilberstein *et al.*, “Digestive tract microbiota in healthy volunteers,” *Clinics (Sao Paulo)*, vol. 62, no. 1, pp. 47–54, Feb. 2007, doi: 10.1590/S1807-59322007000100008.
- [13] S. E. Aidy, B. van den Bogert, and M. Kleerebezem, “The small intestine microbiota, nutritional modulation and relevance for health,” *Curr. Opin. Biotechnol.*, vol. 32, pp. 14–20, Apr. 2015, doi: 10.1016/J.COPBIO.2014.09.005.
- [14] A. J. Kastl, N. A. Terry, G. D. Wu, and L. G. Albenberg, “The Structure and Function of the Human Small Intestinal Microbiota: Current Understanding and Future Directions,” *Cell. Mol. Gastroenterol. Hepatol.*, vol. 9, no. 1, p. 33, Jan. 2020, doi: 10.1016/J.JCMGH.2019.07.006.
- [15] M. Wang, S. Ahrné, B. Jeppsson, and G. Molin, “Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes,” *FEMS Microbiol. Ecol.*, vol. 54, no. 2, pp. 219–231, Oct. 2005, doi: 10.1016/J.FEMSEC.2005.03.012.
- [16] X. Wang, S. P. Heazlewood, D. O. Krause, and T. H. J. Florin, “Molecular characterization of the microbial species that colonize human ileal and colonic mucosa by using 16S rDNA sequence analysis,” *J. Appl. Microbiol.*, vol. 95, no. 3, pp. 508–520, 2003, doi: 10.1046/J.1365-2672.2003.02005.X.
- [17] T. Kanno *et al.*, “Gastric acid reduction leads to an alteration in lower intestinal microflora,” *Biochem. Biophys. Res. Commun.*, vol. 381, no. 4, pp. 666–670, Apr. 2009, doi: 10.1016/J.BBRC.2009.02.109.
- [18] J. Walter and R. Ley, “The human gut microbiome: Ecology and recent evolutionary changes,” *Annual Review of Microbiology*, vol. 65. *Annu Rev Microbiol*, pp. 411–429, 2011, doi: 10.1146/annurev-micro-090110-102830.
- [19] T. P. M. Scheithauer, G. M. Dallinga-Thie, W. M. de Vos, M. Nieuwdorp, and D. H. van Raalte, “Causality of small and large intestinal microbiota in weight regulation and insulin resistance,” *Mol. Metab.*, vol. 5, no. 9, pp. 759–770, Sep. 2016, doi: 10.1016/J.MOLMET.2016.06.002.
- [20] F. Bäckhed, R. E. Ley, J. L. Sonnenburg, D. A. Peterson, and J. I. Gordon, “Host-bacterial mutualism in the human intestine,” *Science*, vol. 307, no. 5717, pp. 1915–1920, Mar. 2005, doi: 10.1126/SCIENCE.1104816.
- [21] N. Ottman, S. Y. Geerlings, S. Aalvink, W. M. de Vos, and C. Belzer, “Action and function of *Akkermansia muciniphila* in microbiome ecology, health and disease,” *Best Pract. Res. Clin. Gastroenterol.*, vol. 31, no. 6, pp. 637–642, Dec. 2017, doi:

- 10.1016/J.BPG.2017.10.001.
- [22] S. M. Lee, G. P. Donaldson, Z. Mikulski, S. Boyajian, K. Ley, and S. K. Mazmanian, “Bacterial colonization factors control specificity and stability of the gut microbiota,” *Nature*, vol. 501, no. 7467, pp. 426–429, 2013, doi: 10.1038/NATURE12447.
- [23] L. V. Hooper and A. J. MacPherson, “Immune adaptations that maintain homeostasis with the intestinal microbiota,” *Nat. Rev. Immunol.*, vol. 10, no. 3, pp. 159–169, Mar. 2010, doi: 10.1038/NRI2710.
- [24] A. Swidsinski, V. Loening-Baucke, H. Lochs, and L. P. Hale, “Spatial organization of bacterial flora in normal and inflamed intestine: a fluorescence in situ hybridization study in mice,” *World J. Gastroenterol.*, vol. 11, no. 8, pp. 1131–1140, Feb. 2005, doi: 10.3748/WJG.V11.I8.1131.
- [25] S. Tuddenham and C. L. Sears, “The intestinal microbiome and health,” *Curr. Opin. Infect. Dis.*, vol. 28, no. 5, pp. 464–470, Jan. 2015, doi: 10.1097/QCO.000000000000196.
- [26] H. Hayashi, M. Sakamoto, and Y. Benno, “Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods,” *Microbiol. Immunol.*, vol. 46, no. 8, pp. 535–548, 2002, doi: 10.1111/J.1348-0421.2002.TB02731.X.
- [27] E. B. Hollister, C. Gao, and J. Versalovic, “Compositional and Functional Features of the Gastrointestinal Microbiome and Their Effects on Human Health,” *Gastroenterology*, vol. 146, no. 6, p. 1449, 2014, doi: 10.1053/J.GASTRO.2014.01.052.
- [28] D. Stanley, M. S. Geier, H. Chen, R. J. Hughes, and R. J. Moore, “Comparison of fecal and cecal microbiotas reveals qualitative similarities but quantitative differences,” *BMC Microbiol.*, vol. 15, no. 1, Dec. 2015, doi: 10.1186/S12866-015-0388-6.
- [29] A. M. O’Hara and F. Shanahan, “The gut flora as a forgotten organ,” *EMBO Rep.*, vol. 7, no. 7, p. 688, Jul. 2006, doi: 10.1038/SJ.EMBOR.7400731.
- [30] M. Rajilić-Stojanović and W. M. de Vos, “The first 1000 cultured species of the human gastrointestinal microbiota,” *FEMS Microbiol. Rev.*, vol. 38, no. 5, pp. 996–1047, Sep. 2014, doi: 10.1111/1574-6976.12075.
- [31] J. Doré and G. Corthier, “[The human intestinal microbiota],” *Gastroenterol. Clin. Biol.*, vol. 34 Suppl 1, no. SUPPL. 1, Sep. 2010, doi: 10.1016/S0399-8320(10)70015-4.
- [32] M. Shapira, “Gut Microbiotas and Host Evolution: Scaling Up Symbiosis,” *Trends Ecol. Evol.*, vol. 31, no. 7, pp. 539–549, Jul. 2016, doi: 10.1016/J.TREE.2016.03.006.
- [33] J. Qin *et al.*, “A human gut microbial gene catalogue established by metagenomic

- sequencing,” *Nature*, vol. 464, no. 7285, pp. 59–65, 2010, doi: 10.1038/NATURE08821.
- [34] E. Le Chatelier *et al.*, “Richness of human gut microbiome correlates with metabolic markers,” *Nature*, vol. 500, no. 7464, pp. 541–546, 2013, doi: 10.1038/NATURE12506.
- [35] M. Fassarella, E. E. Blaak, J. Penders, A. Nauta, H. Smidt, and E. G. Zoetendal, “Gut microbiome stability and resilience: elucidating the response to perturbations in order to modulate gut health,” *Gut*, vol. 70, no. 3, pp. 595–605, Mar. 2021, doi: 10.1136/GUTJNL-2020-321747.
- [36] T. Yatsunenko *et al.*, “Human gut microbiome viewed across age and geography,” *Nature*, vol. 486, no. 7402, pp. 222–227, Jun. 2012, doi: 10.1038/NATURE11053.
- [37] A. D. L. A. W. M. A. V. J. A. G. M. de V. W. M. de V. Erwin G. Zoetendal, “The Host Genotype Affects the Bacterial Community in the Human Gastrointestinal Tract,” <http://dx.doi.org/10.1080/089106001750462669>, vol. 13, no. 3, pp. 129–134, Jan. 2009, doi: 10.1080/089106001750462669.
- [38] N. Arpaia *et al.*, “Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation,” *Nature*, vol. 504, no. 7480, pp. 451–455, 2013, doi: 10.1038/NATURE12726.
- [39] M. J. Gosalbes, S. Llop, Y. Vallès, A. Moya, F. Ballester, and M. P. Francino, “Meconium microbiota types dominated by lactic acid or enteric bacteria are differentially associated with maternal eczema and respiratory problems in infants,” *Clin. Exp. Allergy*, vol. 43, no. 2, pp. 198–211, Feb. 2013, doi: 10.1111/CEA.12063.
- [40] R. I. Mackie, A. Sghir, and H. R. Gaskins, “Developmental microbial ecology of the neonatal gastrointestinal tract,” *Am. J. Clin. Nutr.*, vol. 69, no. 5, 1999, doi: 10.1093/AJCN/69.5.1035S.
- [41] M. G. Dominguez-Bello *et al.*, “Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 26, pp. 11971–11975, Jun. 2010, doi: 10.1073/PNAS.1002601107.
- [42] M. M. Grönlund, O. P. Lehtonen, E. Eerola, and P. Kero, “Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery,” *J. Pediatr. Gastroenterol. Nutr.*, vol. 28, no. 1, pp. 19–25, Jan. 1999, doi: 10.1097/00005176-199901000-00007.
- [43] M. Kuitunen *et al.*, “Probiotics prevent IgE-associated allergy until age 5 years in cesarean-delivered children but not in the total cohort,” *J. Allergy Clin. Immunol.*, vol. 123, no. 2, pp. 335–341, Feb. 2009, doi: 10.1016/J.JACI.2008.11.019.
- [44] A. Marcobal and J. L. Sonnenburg, “Human milk oligosaccharide consumption by

- intestinal microbiota,” *Clin. Microbiol. Infect.*, vol. 18 Suppl 4, no. 0 4, pp. 12–15, 2012, doi: 10.1111/J.1469-0691.2012.03863.X.
- [45] S. Matamoros, C. Gras-Leguen, F. Le Vacon, G. Potel, and M. F. De La Cochetiere, “Development of intestinal microbiota in infants and its impact on health,” *Trends Microbiol.*, vol. 21, no. 4, pp. 167–173, Apr. 2013, doi: 10.1016/J.TIM.2012.12.001.
- [46] A. R. Pacheco, D. Barile, M. A. Underwood, and D. A. Mills, “The impact of the milk glycobioime on the neonate gut microbiota,” *Annu. Rev. Anim. Biosci.*, vol. 3, pp. 419–445, Feb. 2015, doi: 10.1146/ANNUREV-ANIMAL-022114-111112.
- [47] W. A. Walker and R. S. Iyengar, “Breast milk, microbiota, and intestinal immune homeostasis,” *Pediatr. Res.*, vol. 77, no. 1–2, pp. 220–228, Jan. 2015, doi: 10.1038/PR.2014.160.
- [48] C. J. Stewart *et al.*, “Temporal development of the gut microbiome in early childhood from the TEDDY study,” *Nature*, vol. 562, no. 7728, pp. 583–588, Oct. 2018, doi: 10.1038/S41586-018-0617-X.
- [49] D. Artis, “Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut,” *Nat. Rev. Immunol.*, vol. 8, no. 6, pp. 411–420, Jun. 2008, doi: 10.1038/NRI2316.
- [50] L. Dethlefsen, M. McFall-Ngai, and D. A. Relman, “An ecological and evolutionary perspective on human-microbe mutualism and disease,” *Nature*, vol. 449, no. 7164, pp. 811–818, Oct. 2007, doi: 10.1038/NATURE06245.
- [51] M. Zhang, K. Sun, Y. Wu, Y. Yang, P. Tso, and Z. Wu, “Interactions between Intestinal Microbiota and Host Immune Response in Inflammatory Bowel Disease,” *Front. Immunol.*, vol. 8, no. AUG, Aug. 2017, doi: 10.3389/FIMMU.2017.00942.
- [52] Y. Maeda and K. Takeda, “Host–microbiota interactions in rheumatoid arthritis,” *Exp. Mol. Med.*, vol. 51, no. 12, Dec. 2019, doi: 10.1038/S12276-019-0283-6.
- [53] M. Zarepour *et al.*, “The mucin Muc2 limits pathogen burdens and epithelial barrier dysfunction during *Salmonella enterica* serovar Typhimurium colitis,” *Infect. Immun.*, vol. 81, no. 10, pp. 3672–3683, 2013, doi: 10.1128/IAI.00854-13.
- [54] D. Artis *et al.*, “RELMbeta/FIZZ2 is a goblet cell-specific immune-effector molecule in the gastrointestinal tract,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 37, pp. 13596–13600, Sep. 2004, doi: 10.1073/PNAS.0404034101.
- [55] L. A. David *et al.*, “Diet rapidly and reproducibly alters the human gut microbiome,” *Nature*, vol. 505, no. 7484, pp. 559–563, 2014, doi: 10.1038/NATURE12820.
- [56] X. Gui, Z. Yang, and M. D. Li, “Effect of Cigarette Smoke on Gut Microbiota: State of

- Knowledge,” *Front. Physiol.*, vol. 12, Jun. 2021, doi: 10.3389/FPHYS.2021.673341.
- [57] L. J. Mailing, J. M. Allen, T. W. Buford, C. J. Fields, and J. A. Woods, “Exercise and the Gut Microbiome: A Review of the Evidence, Potential Mechanisms, and Implications for Human Health,” *Exerc. Sport Sci. Rev.*, vol. 47, no. 2, pp. 75–85, Apr. 2019, doi: 10.1249/JES.000000000000183.
- [58] A. Madison and J. K. Kiecolt-Glaser, “Stress, depression, diet, and the gut microbiota: human-bacteria interactions at the core of psychoneuroimmunology and nutrition,” *Curr. Opin. Behav. Sci.*, vol. 28, pp. 105–110, Aug. 2019, doi: 10.1016/J.COBEHA.2019.01.011.
- [59] F. Sommer, J. M. Anderson, R. Bharti, J. Raes, and P. Rosenstiel, “The resilience of the intestinal microbiota influences health and disease,” *Nat. Rev. Microbiol.*, vol. 15, no. 10, pp. 630–638, Oct. 2017, doi: 10.1038/NRMICRO.2017.58.
- [60] F. Raymond *et al.*, “The initial state of the human gut microbiome determines its reshaping by antibiotics,” *ISME J.*, vol. 10, no. 3, pp. 707–720, Mar. 2016, doi: 10.1038/ISMEJ.2015.148.
- [61] A. Goethel *et al.*, “Nod2 influences microbial resilience and susceptibility to colitis following antibiotic exposure,” *Mucosal Immunol.*, vol. 12, no. 3, pp. 720–732, May 2019, doi: 10.1038/S41385-018-0128-Y.
- [62] J. Moltzau Anderson *et al.*, “NOD2 Influences Trajectories of Intestinal Microbiota Recovery After Antibiotic Perturbation,” *Cell. Mol. Gastroenterol. Hepatol.*, vol. 10, no. 2, pp. 365–389, Jan. 2020, doi: 10.1016/J.JCMGH.2020.03.008.
- [63] C. L. Gentile and T. L. Weir, “The gut microbiota at the intersection of diet and human health,” *Science*, vol. 362, no. 6416, pp. 776–780, Nov. 2018, doi: 10.1126/SCIENCE.AAU5812.
- [64] R. D. Hills, B. A. Pontefract, H. R. Mishcon, C. A. Black, S. C. Sutton, and C. R. Theberge, “Gut Microbiome: Profound Implications for Diet and Disease,” *Nutrients*, vol. 11, no. 7, Jul. 2019, doi: 10.3390/NU11071613.
- [65] K. Ray, “Gut microbiota: Filling up on fibre for a healthy gut,” *Nat. Rev. Gastroenterol. Hepatol.*, vol. 15, no. 2, pp. 67–67, 2018, doi: 10.1038/NRGASTRO.2018.2.
- [66] L. Wrzosek *et al.*, “Bacteroides thetaiotaomicron and Faecalibacterium prausnitzii influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent,” *BMC Biol.*, vol. 11, May 2013, doi: 10.1186/1741-7007-11-61.
- [67] J. W. McRorie and N. M. McKeown, “Understanding the Physics of Functional Fibers

- in the Gastrointestinal Tract: An Evidence-Based Approach to Resolving Enduring Misconceptions about Insoluble and Soluble Fiber,” *J. Acad. Nutr. Diet.*, vol. 117, no. 2, pp. 251–264, Feb. 2017, doi: 10.1016/J.JAND.2016.09.021.
- [68] A. Tomova *et al.*, “The Effects of Vegetarian and Vegan Diets on Gut Microbiota,” *Front. Nutr.*, vol. 6, Apr. 2019, doi: 10.3389/FNUT.2019.00047.
- [69] N. G. Forouhi, R. M. Krauss, G. Taubes, and W. Willett, “Dietary fat and cardiometabolic health: evidence, controversies, and consensus for guidance,” *BMJ*, vol. 361, 2018, doi: 10.1136/BMJ.K2139.
- [70] P. Jethwani and K. Grover, “Gut Microbiota in Health and Diseases-A Review,” *Int.J.Curr.Microbiol.App.Sci*, vol. 8, no. 8, p. 8, 2019, doi: 10.20546/ijcmas.2019.808.187.
- [71] H. Tilg, A. R. Moschen, and A. Kaser, “Obesity and the microbiota,” *Gastroenterology*, vol. 136, no. 5, pp. 1476–1483, 2009, doi: 10.1053/J.GASTRO.2009.03.030.
- [72] G. De Palma *et al.*, “Transplantation of fecal microbiota from patients with irritable bowel syndrome alters gut function and behavior in recipient mice,” *Sci. Transl. Med.*, vol. 9, no. 379, Mar. 2017, doi: 10.1126/SCITRANSLMED.AAF6397.
- [73] J. R. Brestoff and D. Artis, “Commensal bacteria at the interface of host metabolism and the immune system,” *Nat. Immunol.*, vol. 14, no. 7, p. 676, Jul. 2013, doi: 10.1038/NI.2640.
- [74] E. N. Bergman, “Energy contributions of volatile fatty acids from the gastrointestinal tract in various species,” *Physiol. Rev.*, vol. 70, no. 2, pp. 567–590, 1990, doi: 10.1152/PHYSREV.1990.70.2.567.
- [75] T. L. Miller and M. J. Wolin, “Fermentations by saccharolytic intestinal bacteria,” *Am. J. Clin. Nutr.*, vol. 32, no. 1, pp. 164–172, 1979, doi: 10.1093/AJCN/32.1.164.
- [76] J. K. Nicholson *et al.*, “Host-gut microbiota metabolic interactions,” *Science*, vol. 336, no. 6086, pp. 1262–1267, Jun. 2012, doi: 10.1126/SCIENCE.1223813.
- [77] J. Marchix, G. Goddard, and M. A. Helmuth, “Host-Gut Microbiota Crosstalk in Intestinal Adaptation,” *Cell. Mol. Gastroenterol. Hepatol.*, vol. 6, no. 2, pp. 149–162, Jan. 2018, doi: 10.1016/J.JCMGH.2018.01.024.
- [78] G. C. van Zanten *et al.*, “Synbiotic *Lactobacillus acidophilus* NCFM and cellobiose does not affect human gut bacterial diversity but increases abundance of lactobacilli, bifidobacteria and branched-chain fatty acids: a randomized, double-blinded cross-over trial,” *FEMS Microbiol. Ecol.*, vol. 90, no. 1, pp. 225–236, Oct. 2014, doi: 10.1111/1574-6941.12397.

- [79] P. Lepage *et al.*, “A metagenomic insight into our gut’s microbiome,” *Gut*, vol. 62, no. 1, pp. 146–158, Jan. 2013, doi: 10.1136/GUTJNL-2011-301805.
- [80] G. T. Macfarlane and S. Macfarlane, “Fermentation in the human large intestine: its physiologic consequences and the potential contribution of prebiotics,” *J. Clin. Gastroenterol.*, vol. 45 Suppl, no. SUPPL. 3, Nov. 2011, doi: 10.1097/MCG.0B013E31822FECFE.
- [81] P. Das, P. Babaei, and J. Nielsen, “Metagenomic analysis of microbe-mediated vitamin metabolism in the human gut microbiome,” *BMC Genomics*, vol. 20, no. 1, Mar. 2019, doi: 10.1186/S12864-019-5591-7.
- [82] V. Sharma *et al.*, “B-Vitamin Sharing Promotes Stability of Gut Microbial Communities,” *Front. Microbiol.*, vol. 10, no. JUL, 2019, doi: 10.3389/FMICB.2019.01485.
- [83] M. Rossi, A. Amaretti, and S. Raimondi, “Folate production by probiotic bacteria,” *Nutrients*, vol. 3, no. 1, pp. 118–134, 2011, doi: 10.3390/NU3010118.
- [84] H. M. Said and E. Nexo, “Gastrointestinal Handling of Water-Soluble Vitamins,” *Compr. Physiol.*, vol. 8, no. 4, pp. 1291–1311, Sep. 2018, doi: 10.1002/CPHY.C170054.
- [85] Y. Saitoh *et al.*, “Tight junctions. Structural insight into tight junction disassembly by *Clostridium perfringens* enterotoxin,” *Science*, vol. 347, no. 6223, pp. 775–778, Feb. 2015, doi: 10.1126/SCIENCE.1261833.
- [86] E. Cario, G. Gerken, and D. K. Podolsky, “Toll-like receptor 2 controls mucosal inflammation by regulating epithelial barrier function,” *Gastroenterology*, vol. 132, no. 4, pp. 1359–1374, 2007, doi: 10.1053/J.GASTRO.2007.02.056.
- [87] J. Xu *et al.*, “A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis,” *Science*, vol. 299, no. 5615, pp. 2074–2076, Mar. 2003, doi: 10.1126/SCIENCE.1080029.
- [88] F. Lutgendorff, L. M. A. Akkermans, and J. D. Söderholm, “The role of microbiota and probiotics in stress-induced gastro-intestinal damage,” *Curr. Mol. Med.*, vol. 8, no. 4, pp. 282–98, 2008, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/18537636>.
- [89] M. Banasaz, E. Norin, R. Holma, and T. Midtvedt, “Increased Enterocyte Production in Gnotobiotic Rats Mono-Associated with *Lactobacillus rhamnosus* GG,” *Appl. Environ. Microbiol.*, vol. 68, no. 6, p. 3031, 2002, doi: 10.1128/AEM.68.6.3031-3034.2002.
- [90] T. S. Stappenbeck, L. V. Hooper, and J. I. Gordon, “Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 24, pp. 15451–15455, Nov. 2002, doi: 10.1073/PNAS.202604299.

- [91] H. M. Hamer, D. Jonkers, K. Venema, S. Vanhoutvin, F. J. Troost, and R. J. Brummer, "Review article: the role of butyrate on colonic function," *Aliment. Pharmacol. Ther.*, vol. 27, no. 2, pp. 104–119, Jan. 2008, doi: 10.1111/J.1365-2036.2007.03562.X.
- [92] Y. Momose, K. Hirayama, and K. Itoh, "Competition for proline between indigenous *Escherichia coli* and *E. coli* O157:H7 in gnotobiotic mice associated with infant intestinal microbiota and its contribution to the colonization resistance against *E. coli* O157:H7," *Antonie Van Leeuwenhoek*, vol. 94, no. 2, pp. 165–171, Aug. 2008, doi: 10.1007/S10482-008-9222-6.
- [93] A. J. Fabich *et al.*, "Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine," *Infect. Immun.*, vol. 76, no. 3, pp. 1143–1152, Mar. 2008, doi: 10.1128/IAI.01386-07.
- [94] R. Maltby, M. P. Leatham-Jensen, T. Gibson, P. S. Cohen, and T. Conway, "Nutritional basis for colonization resistance by human commensal *Escherichia coli* strains HS and Nissle 1917 against *E. coli* O157:H7 in the mouse intestine," *PLoS One*, vol. 8, no. 1, Jan. 2013, doi: 10.1371/JOURNAL.PONE.0053957.
- [95] A. L. Lewis and W. G. Lewis, "Host sialoglycans and bacterial sialidases: a mucosal perspective," *Cell. Microbiol.*, vol. 14, no. 8, pp. 1174–1182, Aug. 2012, doi: 10.1111/J.1462-5822.2012.01807.X.
- [96] K. M. Ng *et al.*, "Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens," *Nature*, vol. 502, no. 7469, pp. 96–99, 2013, doi: 10.1038/NATURE12503.
- [97] E. Deriu *et al.*, "Probiotic bacteria reduce salmonella typhimurium intestinal colonization by competing for iron," *Cell Host Microbe*, vol. 14, no. 1, pp. 26–37, Jul. 2013, doi: 10.1016/J.CHOM.2013.06.007.
- [98] E. F. O'Shea, P. D. Cotter, C. Stanton, R. P. Ross, and C. Hill, "Production of bioactive substances by intestinal bacteria as a basis for explaining probiotic mechanisms: bacteriocins and conjugated linoleic acid," *Int. J. Food Microbiol.*, vol. 152, no. 3, pp. 189–205, Jan. 2012, doi: 10.1016/J.IJFOODMICRO.2011.05.025.
- [99] S. M. Asaduzzaman and K. Sonomoto, "Lantibiotics: diverse activities and unique modes of action," *J. Biosci. Bioeng.*, vol. 107, no. 5, pp. 475–487, May 2009, doi: 10.1016/J.JBIOOSC.2009.01.003.
- [100] S. Duquesne, D. Destoumieux-Garzón, J. Peduzzi, and S. Rebuffat, "Microcins, gene-encoded antibacterial peptides from enterobacteria," *Nat. Prod. Rep.*, vol. 24, no. 4, pp. 708–734, 2007, doi: 10.1039/B516237H.

- [101] S. I. Patzer, M. R. Baquero, D. Bravo, F. Moreno, and K. Hantke, “The colicin G, H and X determinants encode microcins M and H47, which might utilize the catecholate siderophore receptors FepA, Cir, Fiu and Iron,” *Microbiology*, vol. 149, no. Pt 9, pp. 2557–2570, Sep. 2003, doi: 10.1099/MIC.0.26396-0.
- [102] M. Sassone-Corsi *et al.*, “Microcins mediate competition among Enterobacteriaceae in the inflamed gut,” *Nature*, vol. 540, no. 7632, pp. 280–283, Dec. 2016, doi: 10.1038/NATURE20557.
- [103] H. L. Cash, C. V. Whitham, C. L. Behrendt, and L. V. Hooper, “Symbiotic bacteria direct expression of an intestinal bactericidal lectin,” *Science*, vol. 313, no. 5790, pp. 1126–1130, Aug. 2006, doi: 10.1126/SCIENCE.1127119.
- [104] L. V. Hooper, T. S. Stappenbeck, C. V. Hong, and J. I. Gordon, “Angiogenins: a new class of microbicidal proteins involved in innate immunity,” *Nat. Immunol.*, vol. 4, no. 3, pp. 269–273, Mar. 2003, doi: 10.1038/NI888.
- [105] L. Caetano *et al.*, “Antivirulence activity of the human gut metabolome,” *MBio*, vol. 5, no. 4, pp. 1–13, 2014, doi: 10.1128/MBIO.01183-14.
- [106] T. De Sablet, C. Chassard, A. Bernalier-Donadille, M. Vareille, A. P. Gobert, and C. Martin, “Human microbiota-secreted factors inhibit shiga toxin synthesis by enterohemorrhagic *Escherichia coli* O157:H7,” *Infect. Immun.*, vol. 77, no. 2, pp. 783–790, Feb. 2009, doi: 10.1128/IAI.01048-08.
- [107] C. Cordonnier, G. Le Bihan, J. G. Emond-Rheault, A. Garrivier, J. Harel, and G. Jubelin, “Vitamin B12 Uptake by the Gut Commensal Bacteria *Bacteroides thetaiotaomicron* Limits the Production of Shiga Toxin by Enterohemorrhagic *Escherichia coli*,” *Toxins (Basel)*, vol. 8, no. 1, Jan. 2016, doi: 10.3390/TOXINS8010014.
- [108] K. Honda and D. R. Littman, “The microbiota in adaptive immune homeostasis and disease,” *Nature*, vol. 535, no. 7610, pp. 75–84, Jul. 2016, doi: 10.1038/NATURE18848.
- [109] C. A. Thaiss, N. Zmora, M. Levy, and E. Elinav, “The microbiome and innate immunity,” *Nature*, vol. 535, no. 7610, pp. 65–74, Jul. 2016, doi: 10.1038/NATURE18847.
- [110] E. F. Verdu, H. J. Galipeau, and B. Jabri, “Novel players in coeliac disease pathogenesis: role of the gut microbiota,” *Nat. Rev. Gastroenterol. Hepatol.*, vol. 12, no. 9, p. 497, Sep. 2015, doi: 10.1038/NRGASTRO.2015.90.
- [111] J. Grigat, A. Soruri, U. Forssmann, J. Riggert, and J. Zwirner, “Chemoattraction of macrophages, T lymphocytes, and mast cells is evolutionarily conserved within the human alpha-defensin family,” *J. Immunol.*, vol. 179, no. 6, pp. 3958–3965, Sep. 2007,

- doi: 10.4049/JIMMUNOL.179.6.3958.
- [112] B. P. Willing, A. Vacharaksa, M. Croxen, T. Thanachayanont, and B. B. Finlay, “Altering host resistance to infections through microbial transplantation,” *PLoS One*, vol. 6, no. 10, 2011, doi: 10.1371/JOURNAL.PONE.0026988.
- [113] D. C. Baumgart and W. J. Sandborn, “Inflammatory bowel disease: clinical aspects and established and evolving therapies,” *Lancet (London, England)*, vol. 369, no. 9573, pp. 1641–1657, May 2007, doi: 10.1016/S0140-6736(07)60751-X.
- [114] J. M. Dahlhamer, E. P. Zammitti, B. W. Ward, A. G. Wheaton, and J. B. Croft, “Prevalence of Inflammatory Bowel Disease Among Adults Aged ≥ 18 Years - United States, 2015,” *MMWR. Morb. Mortal. Wkly. Rep.*, vol. 65, no. 42, pp. 1166–1169, Oct. 2016, doi: 10.15585/MMWR.MM6542A3.
- [115] G. G. Kaplan and S. C. Ng, “Understanding and Preventing the Global Increase of Inflammatory Bowel Disease,” *Gastroenterology*, vol. 152, no. 2, pp. 313-321.e2, Feb. 2017, doi: 10.1053/J.GASTRO.2016.10.020.
- [116] N. A. Molodecky *et al.*, “Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review,” *Gastroenterology*, vol. 142, no. 1, 2012, doi: 10.1053/J.GASTRO.2011.10.001.
- [117] J. Torres *et al.*, “ECCO Guidelines on Therapeutics in Crohn’s Disease: Medical Treatment,” *J. Crohns. Colitis*, vol. 14, no. 1, pp. 4–22, Jan. 2020, doi: 10.1093/ECCO-JCC/JJZ180.
- [118] A. Rehman *et al.*, “Nod2 is essential for temporal development of intestinal microbial communities,” *Gut*, vol. 60, no. 10, pp. 1354–1362, Oct. 2011, doi: 10.1136/GUT.2010.216259.
- [119] D. Knights *et al.*, “Complex host genetics influence the microbiome in inflammatory bowel disease,” *Genome Med.*, vol. 6, no. 12, Dec. 2014, doi: 10.1186/S13073-014-0107-1.
- [120] Y. Ogura *et al.*, “A frameshift mutation in NOD2 associated with susceptibility to Crohn’s disease,” *Nature*, vol. 411, no. 6837, pp. 603–606, May 2001, doi: 10.1038/35079114.
- [121] J. Lamoril, J. C. Deybach, and P. Bouizegarène, “Maladie de Crohn et génétique : connaissances actuelles,” *Immuno-analyse Biol. Spécialisée*, vol. 22, no. 3, pp. 137–150, Jun. 2007, doi: 10.1016/J.IMMBIO.2007.01.002.
- [122] A. Franke *et al.*, “Genome-wide meta-analysis increases to 71 the number of confirmed Crohn’s disease susceptibility loci,” *Nat. Genet.*, vol. 42, no. 12, pp. 1118–1125, 2010,

- doi: 10.1038/NG.717.
- [123] K. M. De Lange *et al.*, “Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease,” *Nat. Genet.*, vol. 49, no. 2, pp. 256–261, Jan. 2017, doi: 10.1038/NG.3760.
- [124] J. R. F. Cummings *et al.*, “Confirmation of the role of ATG16L1 as a Crohn’s disease susceptibility gene,” *Inflamm. Bowel Dis.*, vol. 13, no. 8, pp. 941–946, Aug. 2007, doi: 10.1002/IBD.20162.
- [125] M. P. Buisine *et al.*, “Abnormalities in mucin gene expression in Crohn’s disease,” *Inflamm. Bowel Dis.*, vol. 5, no. 1, pp. 24–32, 1999, doi: 10.1097/00054725-199902000-00004.
- [126] J. D. Söderholm *et al.*, “Augmented increase in tight junction permeability by luminal stimuli in the non-inflamed ileum of Crohn’s disease,” *Gut*, vol. 50, no. 3, pp. 307–313, 2002, doi: 10.1136/GUT.50.3.307.
- [127] E. V. Loftus, “Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences,” *Gastroenterology*, vol. 126, no. 6, pp. 1504–1517, 2004, doi: 10.1053/J.GASTRO.2004.01.063.
- [128] J. K. Hou, H. El-Serag, and S. Thirumurthi, “Distribution and manifestations of inflammatory bowel disease in Asians, Hispanics, and African Americans: a systematic review,” *Am. J. Gastroenterol.*, vol. 104, no. 8, pp. 2100–2109, Aug. 2009, doi: 10.1038/AJG.2009.190.
- [129] C. N. Bernstein, A. Kraut, J. F. Blanchard, P. Rawsthorne, N. Yu, and R. Walld, “The relationship between inflammatory bowel disease and socioeconomic variables,” *Am. J. Gastroenterol.*, vol. 96, no. 7, pp. 2117–2125, Jul. 2001, doi: 10.1111/J.1572-0241.2001.03946.X.
- [130] G. G. Kaplan *et al.*, “The inflammatory bowel diseases and ambient air pollution: a novel association,” *Am. J. Gastroenterol.*, vol. 105, no. 11, pp. 2412–2419, Nov. 2010, doi: 10.1038/AJG.2010.252.
- [131] J. K. Hou, B. Abraham, and H. El-Serag, “Dietary intake and risk of developing inflammatory bowel disease: a systematic review of the literature,” *Am. J. Gastroenterol.*, vol. 106, no. 4, pp. 563–573, Apr. 2011, doi: 10.1038/AJG.2011.44.
- [132] E. Lindberg, G. Jarnerot, and B. Huitfeldt, “Smoking in Crohn’s disease: effect on localisation and clinical course,” *Gut*, vol. 33, no. 6, pp. 779–782, 1992, doi: 10.1136/GUT.33.6.779.
- [133] M. G. V. M. Russel *et al.*, “Inflammatory bowel disease: is there any relation between

- smoking status and disease presentation? European Collaborative IBD Study Group,” *Inflamm. Bowel Dis.*, vol. 4, no. 3, pp. 182–186, 1998, doi: 10.1097/00054725-199808000-00002.
- [134] J. Cosnes, F. Carbonnel, F. Carrat, L. Beaugerie, S. Cattan, and J. P. Gendre, “Effects of current and former cigarette smoking on the clinical course of Crohn’s disease,” *Aliment. Pharmacol. Ther.*, vol. 13, no. 11, pp. 1403–1411, 1999, doi: 10.1046/J.1365-2036.1999.00630.X.
- [135] M. S. Desai *et al.*, “A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility,” *Cell*, vol. 167, no. 5, pp. 1339–1353.e21, Nov. 2016, doi: 10.1016/J.CELL.2016.10.043.
- [136] S. J. Ott *et al.*, “Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease,” *Gut*, vol. 53, no. 5, pp. 685–693, May 2004, doi: 10.1136/GUT.2003.025403.
- [137] N. M. J. Schwerbrock *et al.*, “Interleukin 10-deficient mice exhibit defective colonic Muc2 synthesis before and after induction of colitis by commensal bacteria,” *Inflamm. Bowel Dis.*, vol. 10, no. 6, pp. 811–823, Nov. 2004, doi: 10.1097/00054725-200411000-00016.
- [138] A. Lavelle and H. Sokol, “Gut microbiota-derived metabolites as key actors in inflammatory bowel disease,” *Nat. Rev. Gastroenterol. Hepatol.*, vol. 17, no. 4, pp. 223–237, Apr. 2020, doi: 10.1038/S41575-019-0258-Z.
- [139] D. N. Frank, A. L. St. Amand, R. A. Feldman, E. C. Boedeker, N. Harpaz, and N. R. Pace, “Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases,” *Proc. Natl. Acad. Sci.*, vol. 104, no. 34, pp. 13780–13785, Aug. 2007, doi: 10.1073/PNAS.0706625104.
- [140] R. B. Sartor, “Microbial influences in inflammatory bowel diseases,” *Gastroenterology*, vol. 134, no. 2, pp. 577–594, 2008, doi: 10.1053/J.GASTRO.2007.11.059.
- [141] B. Yilmaz *et al.*, “Microbial network disturbances in relapsing refractory Crohn’s disease,” *Nat. Med.*, vol. 25, no. 2, pp. 323–336, Feb. 2019, doi: 10.1038/S41591-018-0308-Z.
- [142] E. G. Zoetendal, A. Von Wright, T. Vilpponen-Salmela, K. Ben-Amor, A. D. L. Akkermans, and W. M. De Vos, “Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces,” *Appl. Environ. Microbiol.*, vol. 68, no. 7, pp. 3401–3407, 2002, doi: 10.1128/AEM.68.7.3401-3407.2002.

- [143] D. Gevers *et al.*, “The treatment-naïve microbiome in new-onset Crohn’s disease,” *Cell Host Microbe*, vol. 15, no. 3, p. 382, 2014, doi: 10.1016/J.CHOM.2014.02.005.
- [144] M. Lopez-Siles *et al.*, “Mucosa-associated Faecalibacterium prausnitzii and Escherichia coli co-abundance can distinguish Irritable Bowel Syndrome and Inflammatory Bowel Disease phenotypes,” *Int. J. Med. Microbiol.*, vol. 304, no. 3–4, pp. 464–475, 2014, doi: 10.1016/J.IJMM.2014.02.009.
- [145] A. Swidsinski, V. Loening-Baucke, and A. Herber, “Mucosal flora in Crohn’s disease and ulcerative colitis - an overview,” *J. Physiol. Pharmacol.*, vol. 60 Suppl 6, no. SUPPL.6, pp. 61–71, 2009, Accessed: Oct. 21, 2022. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/20224153/>.
- [146] M. S. Sadabad *et al.*, “The ATG16L1-T300A allele impairs clearance of pathosymbionts in the inflamed ileal mucosa of Crohn’s disease patients,” *Gut*, vol. 64, no. 10, pp. 1546–1552, Oct. 2015, doi: 10.1136/GUTJNL-2014-307289.
- [147] M. Tong *et al.*, “Reprogramming of gut microbiome energy metabolism by the FUT2 Crohn’s disease risk polymorphism,” *ISME J.*, vol. 8, no. 11, pp. 2193–2206, Nov. 2014, doi: 10.1038/ISMEJ.2014.64.
- [148] D. Li *et al.*, “A Pleiotropic Missense Variant in SLC39A8 Is Associated With Crohn’s Disease and Human Gut Microbiome Composition,” *Gastroenterology*, vol. 151, no. 4, pp. 724–732, Oct. 2016, doi: 10.1053/J.GASTRO.2016.06.051.
- [149] M. Baumgart *et al.*, “Culture independent analysis of ileal mucosa reveals a selective increase in invasive Escherichia coli of novel phylogeny relative to depletion of Clostridiales in Crohn’s disease involving the ileum,” *ISME J.*, vol. 1, no. 5, pp. 403–418, Sep. 2007, doi: 10.1038/ISMEJ.2007.52.
- [150] D. A. Peterson, D. N. Frank, N. R. Pace, and J. I. Gordon, “Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases,” *Cell Host Microbe*, vol. 3, no. 6, pp. 417–427, Jun. 2008, doi: 10.1016/J.CHOM.2008.05.001.
- [151] A. Darfeuille-Michaud *et al.*, “High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn’s disease,” *Gastroenterology*, vol. 127, no. 2, pp. 412–421, 2004, doi: 10.1053/j.gastro.2004.04.061.
- [152] A. Darfeuille-Michaud *et al.*, “Presence of adherent Escherichia coli strains in ileal mucosa of patients with Crohn’s disease,” *Gastroenterology*, vol. 115, no. 6, pp. 1405–1413, 1998, doi: 10.1016/S0016-5085(98)70019-8.
- [153] D. Ellinghaus *et al.*, “Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci,” *Am. J. Hum. Genet.*,

- vol. 90, no. 4, pp. 636–647, Apr. 2012, doi: 10.1016/J.AJHG.2012.02.020.
- [154] T. R. Elliott *et al.*, “Quantification and characterization of mucosa-associated and intracellular *Escherichia coli* in inflammatory bowel disease,” *Inflamm. Bowel Dis.*, vol. 19, no. 11, pp. 2326–2338, Oct. 2013, doi: 10.1097/MIB.0B013E3182A38A92.
- [155] L. Gibold *et al.*, “The Vat-AIEC protease promotes crossing of the intestinal mucus layer by Crohn’s disease-associated *Escherichia coli*,” *Cell. Microbiol.*, vol. 18, no. 5, pp. 617–631, May 2016, doi: 10.1111/CMI.12539.
- [156] J. Boudeau, N. Barnich, and A. Darfeuille-Michaud, “Type 1 pili-mediated adherence of *Escherichia coli* strain LF82 isolated from Crohn’s disease is involved in bacterial invasion of intestinal epithelial cells,” *Mol. Microbiol.*, vol. 39, no. 5, pp. 1272–1284, Feb. 2001, doi: 10.1111/J.1365-2958.2001.02315.X.
- [157] N. Dreux *et al.*, “Point mutations in FimH adhesin of Crohn’s disease-associated adherent-invasive *Escherichia coli* enhance intestinal inflammatory response,” *PLoS Pathog.*, vol. 9, no. 1, Jan. 2013, doi: 10.1371/JOURNAL.PPAT.1003141.
- [158] S. Sepehri, R. Kotlowski, C. N. Bernstein, and D. O. Krause, “Phylogenetic analysis of inflammatory bowel disease associated *Escherichia coli* and the *fimH* virulence determinant,” *Inflamm. Bowel Dis.*, vol. 15, no. 11, pp. 1737–1745, 2009, doi: 10.1002/IBD.20966.
- [159] M. Martinez-Medina *et al.*, “Biofilm formation as a novel phenotypic feature of adherent-invasive *Escherichia coli* (AIEC),” *BMC Microbiol.*, vol. 9, 2009, doi: 10.1186/1471-2180-9-202.
- [160] S. Subramanian *et al.*, “Characterization of epithelial IL-8 response to inflammatory bowel disease mucosal *E. coli* and its inhibition by mesalamine,” *Inflamm. Bowel Dis.*, vol. 14, no. 2, pp. 162–175, Feb. 2008, doi: 10.1002/IBD.20296.
- [161] N. S. Ambrose, M. Johnson, D. W. Burdon, and M. R. B. Keighley, “Incidence of pathogenic bacteria from mesenteric lymph nodes and ileal serosa during Crohn’s disease surgery,” *Br. J. Surg.*, vol. 71, no. 8, pp. 623–625, 1984, doi: 10.1002/BJS.1800710821.
- [162] E. C. Steinbach and S. E. Plevy, “The role of macrophages and dendritic cells in the initiation of inflammation in IBD,” *Inflamm. Bowel Dis.*, vol. 20, no. 1, pp. 166–175, Jan. 2014, doi: 10.1097/MIB.0B013E3182A69DCA.
- [163] E. Vazeille *et al.*, “GipA Factor Supports Colonization of Peyer’s Patches by Crohn’s Disease-associated *Escherichia Coli*,” *Inflamm. Bowel Dis.*, vol. 22, no. 1, pp. 68–81, Jan. 2016, doi: 10.1097/MIB.0000000000000609.

- [164] I. Ahmed, B. C. Roy, S. A. Khan, S. Septer, and S. Umar, “Microbiome, Metabolome and Inflammatory Bowel Disease,” *Microorganisms*, vol. 4, no. 2, Jun. 2016, doi: 10.3390/MICROORGANISMS4020020.
- [165] G. Mazzeo *et al.*, “Pathogenic Role of Associated Adherent-Invasive Escherichia coli in Crohn’s Disease,” *J. Cell. Physiol.*, vol. 232, no. 10, pp. 2860–2868, Oct. 2017, doi: 10.1002/JCP.25717.
- [166] A. Darfeuille-Michaud, “Adherent-invasive Escherichia coli: a putative new E. coli pathotype associated with Crohn’s disease,” *Int. J. Med. Microbiol.*, vol. 292, no. 3–4, pp. 185–193, Jan. 2002, doi: 10.1078/1438-4221-00201.
- [167] F. A. Carvalho *et al.*, “Crohn’s disease adherent-invasive Escherichia coli colonize and induce strong gut inflammation in transgenic mice expressing human CEACAM,” *J. Exp. Med.*, vol. 206, no. 10, pp. 2179–2189, Sep. 2009, doi: 10.1084/JEM.20090741.
- [168] F. H. Epstein and E. D. Harris, “Rheumatoid arthritis. Pathophysiology and implications for therapy,” *N. Engl. J. Med.*, vol. 322, no. 18, pp. 1277–1289, May 1990, doi: 10.1056/NEJM199005033221805.
- [169] J. Bullock *et al.*, “Rheumatoid Arthritis: A Brief Overview of the Treatment,” *Med. Princ. Pract.*, vol. 27, no. 6, pp. 501–507, Mar. 2018, doi: 10.1159/000493390.
- [170] G. J. Tobón, P. Youinou, and A. Saraux, “The environment, geo-epidemiology, and autoimmune disease: Rheumatoid arthritis,” *J. Autoimmun.*, vol. 35, no. 1, pp. 10–14, Aug. 2010, doi: 10.1016/J.JAUT.2009.12.009.
- [171] A. I. Catrina, K. D. Deane, and J. U. Scher, “Gene, environment, microbiome and mucosal immune tolerance in rheumatoid arthritis,” *Rheumatology (Oxford)*, vol. 55, no. 3, pp. 391–402, Mar. 2016, doi: 10.1093/RHEUMATOLOGY/KEU469.
- [172] M. Feldmann, F. M. Brennan, and R. N. Maini, “Rheumatoid arthritis,” *Cell*, vol. 85, no. 3, pp. 307–310, May 1996, doi: 10.1016/S0092-8674(00)81109-5.
- [173] E. H. S. Choy *et al.*, “Therapeutic benefit of blocking interleukin-6 activity with an anti-interleukin-6 receptor monoclonal antibody in rheumatoid arthritis: a randomized, double-blind, placebo-controlled, dose-escalation trial,” *Arthritis Rheum.*, vol. 46, no. 12, pp. 3143–3150, Dec. 2002, doi: 10.1002/ART.10623.
- [174] D. Aletaha *et al.*, “2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative,” *Arthritis Rheum.*, vol. 62, no. 9, pp. 2569–2581, 2010, doi: 10.1002/ART.27584.
- [175] D. L. Scott, F. Wolfe, and T. W. J. Huizinga, “Rheumatoid arthritis,” *Lancet (London, England)*, vol. 376, no. 9746, pp. 1094–1108, Sep. 2010, doi: 10.1016/S0140-

6736(10)60826-4.

- [176] K. E. Donahue *et al.*, “Systematic review: comparative effectiveness and harms of disease-modifying medications for rheumatoid arthritis,” *Ann. Intern. Med.*, vol. 148, no. 2, pp. 124–134, Jan. 2008, doi: 10.7326/0003-4819-148-2-200801150-00192.
- [177] C. Salliot and D. Van Der Heijde, “Long-term safety of methotrexate monotherapy in patients with rheumatoid arthritis: a systematic literature research,” *Ann. Rheum. Dis.*, vol. 68, no. 7, pp. 1100–1104, Jul. 2009, doi: 10.1136/ARD.2008.093690.
- [178] M. M. An, Z. Zou, H. Shen, J. D. Zhang, Y. B. Cao, and Y. Y. Jiang, “The addition of tocilizumab to DMARD therapy for rheumatoid arthritis: a meta-analysis of randomized controlled trials,” *Eur. J. Clin. Pharmacol.*, vol. 66, no. 1, pp. 49–59, Jan. 2010, doi: 10.1007/S00228-009-0754-0.
- [179] E. Hurkmans, F. J. Van Der Giesen, T. P. M. V. Vlieland, J. Schoones, and E. C. H. M. Van Den Ende, “Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis,” *Cochrane database Syst. Rev.*, vol. 2009, no. 4, 2009, doi: 10.1002/14651858.CD006853.PUB2.
- [180] A. K. Rodríguez-Elías, K. Maldonado-Murillo, L. F. López-Mendoza, and J. Ramírez-Bello, “[Genetics and genomics in rheumatoid arthritis (RA): An update],” *Gac. Med. Mex.*, vol. 152, no. 2, pp. 218–227, 2016, Accessed: Oct. 21, 2022. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/27160622/>.
- [181] A. H. M. Van Der Helm-Van Mil, T. W. J. Huizinga, G. M. T. Schreuder, F. C. Breedveld, R. R. P. De Vries, and R. E. M. Toes, “An independent role of protective HLA class II alleles in rheumatoid arthritis severity and susceptibility,” *Arthritis Rheum.*, vol. 52, no. 9, pp. 2637–2644, Sep. 2005, doi: 10.1002/ART.21272.
- [182] W. X. Liu, Y. Jiang, Q. X. Hu, and X. B. You, “HLA-DRB1 shared epitope allele polymorphisms and rheumatoid arthritis: a systemic review and meta-analysis,” *Clin. Invest. Med.*, vol. 39, no. 6, pp. E182–E203, 2016, doi: 10.25011/CIM.V39I6.27487.
- [183] H. Furukawa, S. Oka, K. Shimada, A. Hashimoto, and S. Tohma, “Human leukocyte antigen polymorphisms and personalized medicine for rheumatoid arthritis,” *J. Hum. Genet.*, vol. 60, no. 11, pp. 691–696, Nov. 2015, doi: 10.1038/JHG.2015.36.
- [184] S. Ling, E. N. Cline, T. S. Haug, D. A. Fox, and J. Holoshitz, “Citrullinated calreticulin potentiates rheumatoid arthritis shared epitope signaling,” *Arthritis Rheum.*, vol. 65, no. 3, pp. 618–626, Mar. 2013, doi: 10.1002/ART.37814.
- [185] F. Pratesi *et al.*, “HLA shared epitope and ACPA: just a marker or an active player?,” *Autoimmun. Rev.*, vol. 12, no. 12, pp. 1182–1187, Oct. 2013, doi:

- 10.1016/J.AUTREV.2013.08.002.
- [186] L. Lourido, F. J. Blanco, and C. Ruiz-Romero, “Defining the proteomic landscape of rheumatoid arthritis: progress and prospective clinical applications,” *Expert Rev. Proteomics*, vol. 14, no. 5, pp. 431–444, May 2017, doi: 10.1080/14789450.2017.1321481.
- [187] K. Kim, S. Y. Bang, H. S. Lee, and S. C. Bae, “Update on the genetic architecture of rheumatoid arthritis,” *Nat. Rev. Rheumatol.*, vol. 13, no. 1, pp. 13–24, Jan. 2017, doi: 10.1038/NRRHEUM.2016.176.
- [188] L. Klareskog, K. Amara, and V. Malmström, “Adaptive immunity in rheumatoid arthritis: anticitrulline and other antibodies in the pathogenesis of rheumatoid arthritis,” *Curr. Opin. Rheumatol.*, vol. 26, no. 1, pp. 72–79, Jan. 2014, doi: 10.1097/BOR.000000000000016.
- [189] M. J. H. De Hair, K. A. Lehmann, M. G. H. Van De Sande, K. I. Maijer, D. M. Gerlag, and P. P. Tak, “The clinical picture of rheumatoid arthritis according to the 2010 American College of Rheumatology/European League Against Rheumatism criteria: is this still the same disease?,” *Arthritis Rheum.*, vol. 64, no. 2, pp. 389–393, Feb. 2012, doi: 10.1002/ART.33348.
- [190] T. K. Kvien, T. Uhlig, S. Ødegård, and M. S. Heiberg, “Epidemiological aspects of rheumatoid arthritis: the sex ratio,” *Ann. N. Y. Acad. Sci.*, vol. 1069, pp. 212–222, 2006, doi: 10.1196/ANNALS.1351.019.
- [191] T. Frisell, S. Saevarsdottir, and J. Askling, “Family history of rheumatoid arthritis: an old concept with new developments,” *Nat. Rev. Rheumatol.*, vol. 12, no. 6, pp. 335–343, May 2016, doi: 10.1038/NRRHEUM.2016.52.
- [192] R. H. Straub, “The complex role of estrogens in inflammation,” *Endocr. Rev.*, vol. 28, no. 5, pp. 521–574, Aug. 2007, doi: 10.1210/ER.2007-0001.
- [193] H. A. Beydoun, R. El-Amin, M. McNeal, C. Perry, and D. F. Archer, “Reproductive history and postmenopausal rheumatoid arthritis among women 60 years or older: Third National Health and Nutrition Examination Survey,” *Menopause*, vol. 20, no. 9, pp. 930–935, Sep. 2013, doi: 10.1097/GME.0B013E3182A14372.
- [194] M. Pikwer, U. Bergström, J. Å. Nilsson, L. Jacobsson, and C. Turesson, “Early menopause is an independent predictor of rheumatoid arthritis,” *Ann. Rheum. Dis.*, vol. 71, no. 3, pp. 378–381, Mar. 2012, doi: 10.1136/ARD.2011.200059.
- [195] J. Y. Chen and S. P. Ballou, “The effect of antiestrogen agents on risk of autoimmune disorders in patients with breast cancer,” *J. Rheumatol.*, vol. 42, no. 1, pp. 55–59, Jan.

- 2015, doi: 10.3899/JRHEUM.140367.
- [196] K. H. Costenbader, S. C. Chang, I. De Vivo, R. Plenge, and E. W. Karlson, “Genetic polymorphisms in PTPN22, PADI-4, and CTLA-4 and risk for rheumatoid arthritis in two longitudinal cohort studies: evidence of gene-environment interactions with heavy cigarette smoking,” *Arthritis Res. Ther.*, vol. 10, no. 3, May 2008, doi: 10.1186/AR2421.
- [197] S. Kobayashi *et al.*, “A role for the aryl hydrocarbon receptor and the dioxin TCDD in rheumatoid arthritis,” *Rheumatology (Oxford)*, vol. 47, no. 9, pp. 1317–1322, 2008, doi: 10.1093/RHEUMATOLOGY/KEN259.
- [198] A. Manzel, D. N. Muller, D. A. Hafler, S. E. Erdman, R. A. Linker, and M. Kleinewietfeld, “Role of ‘Western diet’ in inflammatory autoimmune diseases,” *Curr. Allergy Asthma Rep.*, vol. 14, no. 1, Jan. 2014, doi: 10.1007/S11882-013-0404-6.
- [199] F. Oliviero, P. Spinella, U. Fiocco, R. Ramonda, P. Sfriso, and L. Punzi, “How the Mediterranean diet and some of its components modulate inflammatory pathways in arthritis,” *Swiss Med. Wkly.*, vol. 145, 2015, doi: 10.4414/SMW.2015.14190.
- [200] A. M. Minihane *et al.*, “Low-grade inflammation, diet composition and health: current research evidence and its translation,” *Br. J. Nutr.*, vol. 114, no. 7, pp. 999–1012, Oct. 2015, doi: 10.1017/S0007114515002093.
- [201] B. Qin *et al.*, “Body mass index and the risk of rheumatoid arthritis: a systematic review and dose-response meta-analysis,” *Arthritis Res. Ther.*, vol. 17, no. 1, Mar. 2015, doi: 10.1186/S13075-015-0601-X.
- [202] W. S. Chen, “Restoration of elbow flexion by latissimus dorsi myocutaneous or muscle flap,” *Arch. Orthop. Trauma Surg.*, vol. 109, no. 3, pp. 117–120, Apr. 1990, doi: 10.1007/BF00440569.
- [203] J. U. Scher *et al.*, “Periodontal disease and the oral microbiota in new-onset rheumatoid arthritis,” *Arthritis Rheum.*, vol. 64, no. 10, pp. 3083–3094, Oct. 2012, doi: 10.1002/ART.34539.
- [204] J. U. Scher *et al.*, “The lung microbiota in early rheumatoid arthritis and autoimmunity,” *Microbiome*, vol. 4, no. 1, p. 60, Nov. 2016, doi: 10.1186/S40168-016-0206-X.
- [205] X. Zhang *et al.*, “The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment,” *Nat. Med.*, vol. 21, no. 8, pp. 895–905, Aug. 2015, doi: 10.1038/NM.3914.
- [206] R. D. Haigh *et al.*, “Draft Whole-Genome Sequences of Periodontal Pathobionts *Porphyromonas gingivalis*, *Prevotella intermedia*, and *Tannerella forsythia* Contain Phase-Variable Restriction-Modification Systems,” *Genome Announc.*, vol. 5, no. 46,

- Nov. 2017, doi: 10.1128/GENOMEA.01229-17.
- [207] G. A. Schellekens, B. A. W. De Jong, F. H. J. Van Den Hoogen, L. B. A. Van De Putte, and W. J. Van Venrooij, "Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies," *J. Clin. Invest.*, vol. 101, no. 1, pp. 273–281, Jan. 1998, doi: 10.1172/JCI1316.
- [208] S. C. Bae and Y. H. Lee, "Association between anti-*Porphyromonas gingivalis* antibody, anti-citrullinated protein antibodies, and rheumatoid arthritis: A meta-analysis," *Z. Rheumatol.*, vol. 77, no. 6, pp. 522–532, Aug. 2018, doi: 10.1007/S00393-017-0328-Y.
- [209] M. Shaw, B. F. Collins, L. A. Ho, and G. Raghu, "Rheumatoid arthritis-associated lung disease," *Eur. Respir. Rev.*, vol. 24, no. 135, pp. 1–16, 2015, doi: 10.1183/09059180.00008014.
- [210] E. Goleva *et al.*, "The effects of airway microbiome on corticosteroid responsiveness in asthma," *Am. J. Respir. Crit. Care Med.*, vol. 188, no. 10, pp. 1193–1201, Nov. 2013, doi: 10.1164/RCCM.201304-0775OC.
- [211] J. R. Erb-Downward *et al.*, "Analysis of the lung microbiome in the 'healthy' smoker and in COPD," *PLoS One*, vol. 6, no. 2, 2011, doi: 10.1371/JOURNAL.PONE.0016384.
- [212] W. Kurowska, E. H. Kuca-Warnawin, A. Radzikowska, and W. Maslinski, "The role of anti-citrullinated protein antibodies (ACPA) in the pathogenesis of rheumatoid arthritis," *Cent. J. Immunol.*, vol. 42, no. 4, pp. 390–398, 2017, doi: 10.5114/CEJI.2017.72807.
- [213] J. Chen *et al.*, "An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis," *Genome Med.*, vol. 8, no. 1, Apr. 2016, doi: 10.1186/S13073-016-0299-7.
- [214] D. Kim, S. A. Yoo, and W. U. Kim, "Gut microbiota in autoimmunity: potential for clinical applications," *Arch. Pharm. Res.*, vol. 39, no. 11, pp. 1565–1576, Nov. 2016, doi: 10.1007/S12272-016-0796-7.
- [215] J. Vahtovuo, E. Munukka, M. Korkeamäki, R. Luukkainen, and P. Toivanen, "Fecal microbiota in early rheumatoid arthritis," *J. Rheumatol.*, vol. 35, no. 8, pp. 1500–1505, Aug. 2008, Accessed: Oct. 21, 2022. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/18528968/>.
- [216] X. Liu, Q. Zou, B. Zeng, Y. Fang, and H. Wei, "Analysis of fecal *Lactobacillus* community structure in patients with early rheumatoid arthritis," *Curr. Microbiol.*, vol. 67, no. 2, pp. 170–176, Aug. 2013, doi: 10.1007/S00284-013-0338-1.
- [217] E. V. Marietta *et al.*, "Suppression of Inflammatory Arthritis by Human Gut-Derived *Prevotella histicola* in Humanized Mice," *Arthritis Rheumatol. (Hoboken, N.J.)*, vol. 68, no. 12, pp. 2878–2888, Dec. 2016, doi: 10.1002/ART.39785.

- [218] A. I. Catrina, V. Joshua, L. Klareskog, and V. Malmström, “Mechanisms involved in triggering rheumatoid arthritis,” *Immunol. Rev.*, vol. 269, no. 1, pp. 162–174, Jan. 2016, doi: 10.1111/IMR.12379.
- [219] M. K. Demoruelle *et al.*, “Brief report: airways abnormalities and rheumatoid arthritis-related autoantibodies in subjects without arthritis: early injury or initiating site of autoimmunity?,” *Arthritis Rheum.*, vol. 64, no. 6, pp. 1756–1761, Jun. 2012, doi: 10.1002/ART.34344.
- [220] V. M. Holers, “Autoimmunity to Citrullinated Proteins and the Initiation of Rheumatoid Arthritis,” *Curr. Opin. Immunol.*, vol. 25, no. 6, p. 728, Dec. 2013, doi: 10.1016/J.COI.2013.09.018.
- [221] H. Hayashi, K. Shibata, M. Sakamoto, S. Tomita, and Y. Benno, “*Prevotella copri* sp. nov. and *Prevotella stercorea* sp. nov., isolated from human faeces,” *Int. J. Syst. Evol. Microbiol.*, vol. 57, no. Pt 5, pp. 941–946, May 2007, doi: 10.1099/IJS.0.64778-0.
- [222] D. Alpizar-Rodriguez *et al.*, “*Prevotella copri* in individuals at risk for rheumatoid arthritis,” *Ann. Rheum. Dis.*, vol. 78, no. 5, pp. 590–593, May 2019, doi: 10.1136/ANNRHEUMDIS-2018-214514.
- [223] Y. Maeda *et al.*, “Dysbiosis Contributes to Arthritis Development via Activation of Autoreactive T Cells in the Intestine,” *Arthritis Rheumatol. (Hoboken, N.J.)*, vol. 68, no. 11, pp. 2646–2661, Nov. 2016, doi: 10.1002/ART.39783.
- [224] J. U. Scher *et al.*, “Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis,” *Elife*, vol. 2013, no. 2, Nov. 2013, doi: 10.7554/ELIFE.01202.001.
- [225] N. J. Bernard, “Rheumatoid arthritis: *Prevotella copri* associated with new-onset untreated RA,” *Nat. Rev. Rheumatol.*, vol. 10, no. 1, p. 2, Jan. 2014, doi: 10.1038/NRRHEUM.2013.187.
- [226] K. Atarashi *et al.*, “Induction of colonic regulatory T cells by indigenous *Clostridium* species,” *Science*, vol. 331, no. 6015, pp. 337–341, Jan. 2011, doi: 10.1126/SCIENCE.1198469.
- [227] J. L. Round *et al.*, “The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota,” *Science*, vol. 332, no. 6032, pp. 974–977, May 2011, doi: 10.1126/SCIENCE.1206095.
- [228] C. Huttenhower *et al.*, “Structure, function and diversity of the healthy human microbiome,” *Nature*, vol. 486, no. 7402, pp. 207–214, Jun. 2012, doi: 10.1038/NATURE11234.

- [229] M. I. Gul'neva and S. M. Noskov, "[Colonic microbial biocenosis in rheumatoid arthritis]," *Klin. Med. (Mosk)*., vol. 89, no. 4, pp. 45–48, 2011, Accessed: Oct. 20, 2022. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/21932563/>.
- [230] X. Wu *et al.*, "Molecular Insight into Gut Microbiota and Rheumatoid Arthritis," *Int. J. Mol. Sci.*, vol. 17, no. 3, Mar. 2016, doi: 10.3390/IJMS17030431.
- [231] S. E. Winter and A. J. Bäumlner, "Why related bacterial species bloom simultaneously in the gut: Principles underlying the 'Like will to like' concept," *Cell. Microbiol.*, vol. 16, no. 2, p. 179, Feb. 2014, doi: 10.1111/CMI.12245.
- [232] N.-R. Shin, T. W. Whon, and J.-W. Bae, "Proteobacteria: microbial signature of dysbiosis in gut microbiota," *Trends Biotechnol.*, vol. 33, no. 9, pp. 496–503, Sep. 2015, doi: 10.1016/j.tibtech.2015.06.011.
- [233] L. R. Lopetuso *et al.*, "Gut Microbiota in Health, Diverticular Disease, Irritable Bowel Syndrome, and Inflammatory Bowel Diseases: Time for Microbial Marker of Gastrointestinal Disorders," *Dig. Dis.*, vol. 36, no. 1, pp. 56–65, Oct. 2018, doi: 10.1159/000477205.
- [234] "Present standing of the family name Enterobacteriaceae rahn 1937," *Int. J. Syst. Bacteriol.*, vol. 31, no. 1, p. 104, Jan. 1981, doi: 10.1099/00207713-31-1-104/CITE/REFWORKS.
- [235] M. Adeolu, S. Alnajar, S. Naushad, and R. S. Gupta, "Genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov.," *Int. J. Syst. Evol. Microbiol.*, vol. 66, no. 12, pp. 5575–5599, Dec. 2016, doi: 10.1099/IJSEM.0.001485.
- [236] C. N. Resendiz-Nava, H. V. Silva-Rojas, A. Rebollar-Alviter, D. M. Rivera-Pastrana, E. M. Mercado-Silva, and G. M. Nava, "A Comprehensive Evaluation of Enterobacteriaceae Primer Sets for Analysis of Host-Associated Microbiota," *Pathog. (Basel, Switzerland)*, vol. 11, no. 1, Jan. 2021, doi: 10.3390/PATHOGENS11010017.
- [237] M. Arumugam *et al.*, "Enterotypes of the human gut microbiome," *Nature*, vol. 473, no. 7346, pp. 174–180, 2011, doi: 10.1038/nature09944.Enterotypes.
- [238] J. G. Caporaso *et al.*, "Moving pictures of the human microbiome," *Genome Biol.*, vol. 12, no. 5, May 2011, doi: 10.1186/GB-2011-12-5-R50.
- [239] T. Escherich, "The intestinal bacteria of the neonate and breast-fed infant. 1884," *Rev. Infect. Dis.*, vol. 10, no. 6, pp. 1220–1225, 1988, doi: 10.1093/CLINIDS/10.6.1220.

- [240] J. B. Kaper, J. P. Nataro, and H. L. T. Mobley, “Pathogenic *Escherichia coli*,” *Nat. Rev. Microbiol.*, vol. 2, no. 2, pp. 123–140, Feb. 2004, doi: 10.1038/NRMICRO818.
- [241] O. Tenaillon, D. Skurnik, B. Picard, and E. Denamur, “The population genetics of commensal *Escherichia coli*,” *Nature Reviews Microbiology*, vol. 8, no. 3, pp. 207–217, Mar. 2010, doi: 10.1038/nrmicro2298.
- [242] L. Jochum and B. Stecher, “Label or Concept - What Is a Pathobiont?,” *Trends Microbiol.*, vol. 28, no. 10, pp. 789–792, Oct. 2020, doi: 10.1016/J.TIM.2020.04.011.
- [243] D. E. Chang *et al.*, “Carbon nutrition of *Escherichia coli* in the mouse intestine,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 19, pp. 7427–7432, May 2004, doi: 10.1073/PNAS.0307888101.
- [244] C. Beloin, A. Roux, and J. M. Ghigo, “*Escherichia coli* biofilms,” *Curr. Top. Microbiol. Immunol.*, vol. 322, pp. 249–289, 2008, doi: 10.1007/978-3-540-75418-3_12.
- [245] S. A. Jones *et al.*, “Respiration of *Escherichia coli* in the mouse intestine,” *Infect. Immun.*, vol. 75, no. 10, pp. 4891–4899, Oct. 2007, doi: 10.1128/IAI.00484-07.
- [246] H. Ochman and R. K. Selander, “Standard reference strains of *Escherichia coli* from natural populations,” *J. Bacteriol.*, vol. 157, no. 2, pp. 690–693, 1984, doi: 10.1128/JB.157.2.690-693.1984.
- [247] E. F. Boyd and D. L. Hartl, “Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution,” *J. Bacteriol.*, vol. 180, no. 5, pp. 1159–1165, 1998, doi: 10.1128/JB.180.5.1159-1165.1998.
- [248] S. Lu *et al.*, “Insights into the evolution of pathogenicity of *Escherichia coli* from genomic analysis of intestinal *E. coli* of *Marmota himalayana* in Qinghai-Tibet plateau of China,” *Emerg. Microbes Infect.*, vol. 5, no. 12, Dec. 2016, doi: 10.1038/EMI.2016.122.
- [249] D. Medini, C. Donati, H. Tettelin, V. Masignani, and R. Rappuoli, “The microbial pan-genome,” *Curr. Opin. Genet. Dev.*, vol. 15, no. 6, pp. 589–594, Dec. 2005, doi: 10.1016/J.GDE.2005.09.006.
- [250] H. Hendrickson, “Order and disorder during *Escherichia coli* divergence,” *PLoS Genet.*, vol. 5, no. 1, Jan. 2009, doi: 10.1371/JOURNAL.PGEN.1000335.
- [251] O. Lukjancenko, T. M. Wassenaar, and D. W. Ussery, “Comparison of 61 sequenced *Escherichia coli* genomes,” *Microb. Ecol.*, vol. 60, no. 4, pp. 708–720, Nov. 2010, doi: 10.1007/S00248-010-9717-3.
- [252] A. Mira, A. B. Martín-Cuadrado, G. D’Auria, and F. Rodríguez-Valera, “The bacterial pan-genome: a new paradigm in microbiology,” *Int. Microbiol.*, vol. 13, no. 2, pp. 45–

- 57, 2010, doi: 10.2436/20.1501.01.110.
- [253] J. G. Lawrence, H. Ochman, and M. A. Ragan, “Reconciling the many faces of lateral gene transfer,” *Trends Microbiol.*, vol. 10, no. 1, pp. 1–4, Jan. 2002, doi: 10.1016/S0966-842X(01)02282-X.
- [254] H. Schmidt and M. Hensel, “Pathogenicity islands in bacterial pathogenesis,” *Clin. Microbiol. Rev.*, vol. 17, no. 1, pp. 14–56, Jan. 2004, doi: 10.1128/CMR.17.1.14-56.2004.
- [255] M. Touchon *et al.*, “Organised Genome Dynamics in the Escherichia coli Species Results in Highly Diverse Adaptive Paths,” *PLoS Genet.*, vol. 5, no. 1, p. 1000344, 2009, doi: 10.1371/JOURNAL.PGEN.1000344.
- [256] E. W. Alm, S. T. Walk, and D. M. Gordon, “The Niche of Escherichia coli,” *Popul. Genet. Bact.*, pp. 67–89, Apr. 2014, doi: 10.1128/9781555817114.CH6.
- [257] D. M. Gordon, C. L. O’Brien, and P. Pavli, “Escherichia coli diversity in the lower intestinal tract of humans,” *Environ. Microbiol. Rep.*, vol. 7, no. 4, pp. 642–648, Aug. 2015, doi: 10.1111/1758-2229.12300.
- [258] L. Zhang, B. Foxman, and C. Marrs, “Both urinary and rectal Escherichia coli isolates are dominated by strains of phylogenetic group B2,” *J. Clin. Microbiol.*, vol. 40, no. 11, pp. 3951–3955, Nov. 2002, doi: 10.1128/JCM.40.11.3951-3955.2002.
- [259] M. Obata-Yasuoka, W. Ba-Thein, T. Tsukamoto, H. Yoshikawa, and H. Hayashi, “Vaginal Escherichia coli share common virulence factor profiles, serotypes and phylogeny with other extraintestinal E. coli,” *Microbiology*, vol. 148, no. Pt 9, pp. 2745–2752, 2002, doi: 10.1099/00221287-148-9-2745.
- [260] S. Watt, P. Lanotte, L. Mereghetti, M. Moulin-Schouleur, B. Picard, and R. Quentin, “Escherichia coli strains from pregnant women and neonates: intraspecies genetic distribution and prevalence of virulence factors,” *J. Clin. Microbiol.*, vol. 41, no. 5, pp. 1929–1935, May 2003, doi: 10.1128/JCM.41.5.1929-1935.2003.
- [261] F. L. Nowrouzian, A. E. Wold, and I. Adlerberth, “Escherichia coli strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants,” *J. Infect. Dis.*, vol. 191, no. 7, pp. 1078–1083, Apr. 2005, doi: 10.1086/427996.
- [262] D. M. Gordon, S. E. Stern, and P. J. Collignon, “Influence of the age and sex of human hosts on the distribution of Escherichia coli ECOR groups and virulence traits,” *Microbiology*, vol. 151, no. Pt 1, pp. 15–23, Jan. 2005, doi: 10.1099/MIC.0.27425-0.
- [263] D. Skurnik *et al.*, “Characteristics of human intestinal Escherichia coli with changing environments,” *Environ. Microbiol.*, vol. 10, no. 8, pp. 2132–2137, 2008, doi:

- 10.1111/J.1462-2920.2008.01636.X.
- [264] J. N. V. Martinson and S. T. Walk, “Escherichia coli residency in the gut of healthy human adults,” *EcoSal Plus*, vol. 9, no. 1, Mar. 2020, doi: 10.1128/ECOSALPLUS.ESP-0003-2020.
- [265] J. Tomas *et al.*, “Early colonizing Escherichia coli elicits remodeling of rat colonic epithelium shifting toward a new homeostatic state,” *ISME J.*, vol. 9, no. 1, pp. 46–58, Jan. 2015, doi: 10.1038/ISMEJ.2014.111.
- [266] N. T. Mueller, E. Bakacs, J. Combellick, Z. Grigoryan, and M. G. Dominguez-Bello, “The infant microbiome development: mom matters,” *Trends Mol. Med.*, vol. 21, no. 2, pp. 109–117, Feb. 2015, doi: 10.1016/J.MOLMED.2014.12.002.
- [267] M. Jiang, Y. Cao, Z. F. Guo, M. Chen, X. Chen, and Z. Guo, “Menaquinone biosynthesis in Escherichia coli: identification of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate as a novel intermediate and re-evaluation of MenD activity,” *Biochemistry*, vol. 46, no. 38, pp. 10979–10989, Sep. 2007, doi: 10.1021/BI700810X.
- [268] J. P. Karl *et al.*, “Fecal concentrations of bacterially derived vitamin K forms are associated with gut microbiota composition but not plasma or fecal cytokine concentrations in healthy adults,” *Am. J. Clin. Nutr.*, vol. 106, no. 4, pp. 1052–1061, Oct. 2017, doi: 10.3945/AJCN.117.155424.
- [269] H. Fang, J. Kang, and D. Zhang, “Microbial production of vitamin B12: A review and future perspectives,” *Microb. Cell Fact.*, vol. 16, no. 1, p. 15, Jan. 2017, doi: 10.1186/s12934-017-0631-y.
- [270] P. H. Degnan, M. E. Taga, and A. L. Goodman, “Vitamin B12 as a modulator of gut microbial ecology,” *Cell Metab.*, vol. 20, no. 5, pp. 769–778, Nov. 2014, doi: 10.1016/J.CMET.2014.10.002.
- [271] Q. R. Ducarmon, R. D. Zwitter, B. V. H. Hornung, W. van Schaik, V. B. Young, and E. J. Kuijper, “Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection,” *Microbiol. Mol. Biol. Rev.*, vol. 83, no. 3, Aug. 2019, doi: 10.1128/MMBR.00007-19.
- [272] S. Hudault, J. Guignot, and A. L. Servin, “Escherichia coli strains colonising the gastrointestinal tract protect germfree mice against Salmonella typhimurium infection,” *Gut*, vol. 49, no. 1, pp. 47–55, 2001, doi: 10.1136/GUT.49.1.47.
- [273] M. P. Leatham, S. Banerjee, S. M. Autieri, R. Mercado-Lubo, T. Conway, and P. S. Cohen, “Precolonized human commensal Escherichia coli strains serve as a barrier to E. coli O157:H7 growth in the streptomycin-treated mouse intestine,” *Infect. Immun.*, vol.

- 77, no. 7, pp. 2876–2886, Jul. 2009, doi: 10.1128/IAI.00059-09.
- [274] G. Rizzatti, L. R. Lopetuso, G. Gibiino, C. Binda, and A. Gasbarrini, “Proteobacteria: A common factor in human diseases,” *Biomed Res. Int.*, vol. 2017, 2017, doi: 10.1155/2017/9351507.
- [275] P. Seksik *et al.*, “Alterations of the dominant faecal bacterial groups in patients with Crohn’s disease of the colon,” *Gut*, vol. 52, no. 2, pp. 237–242, Feb. 2003, doi: 10.1136/GUT.52.2.237.
- [276] T. Wang *et al.*, “Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers,” *ISME J.*, vol. 6, no. 2, pp. 320–329, Feb. 2012, doi: 10.1038/ISMEJ.2011.109.
- [277] L. Krogius-Kurikka *et al.*, “Microbial community analysis reveals high level phylogenetic alterations in the overall gastrointestinal microbiota of diarrhoea-predominant irritable bowel syndrome sufferers,” *BMC Gastroenterol.*, vol. 9, Dec. 2009, doi: 10.1186/1471-230X-9-95.
- [278] N. Fei and L. Zhao, “An opportunistic pathogen isolated from the gut of an obese human causes obesity in germfree mice,” *ISME J.*, vol. 7, no. 4, pp. 880–884, Apr. 2013, doi: 10.1038/ISMEJ.2012.153.
- [279] L. Rigottier-Gois, “Dysbiosis in inflammatory bowel diseases: the oxygen hypothesis,” *ISME J.*, vol. 7, no. 7, pp. 1256–1261, Jul. 2013, doi: 10.1038/ISMEJ.2013.80.
- [280] M. A. Henson and P. Phalak, “Microbiota dysbiosis in inflammatory bowel diseases: in silico investigation of the oxygen hypothesis,” *BMC Syst. Biol.*, vol. 11, no. 1, Dec. 2017, doi: 10.1186/S12918-017-0522-1.
- [281] E. R. Hughes *et al.*, “Microbial Respiration and Formate Oxidation as Metabolic Signatures of Inflammation-Associated Dysbiosis,” *Cell Host Microbe*, vol. 21, no. 2, pp. 208–219, Feb. 2017, doi: 10.1016/J.CHOM.2017.01.005.
- [282] B. Stecher, “The Roles of Inflammation, Nutrient Availability and the Commensal Microbiota in Enteric Pathogen Infection,” *Microbiol. Spectr.*, vol. 3, no. 3, Jun. 2015, doi: 10.1128/MICROBIOLSPEC.MBP-0008-2014.
- [283] F. Rivera-Chávez, C. A. Lopez, and A. J. Bäumler, “Oxygen as a driver of gut dysbiosis,” *Free Radic. Biol. Med.*, vol. 105, pp. 93–101, Apr. 2017, doi: 10.1016/J.FREERADBIOMED.2016.09.022.
- [284] Y. Litvak, M. X. Byndloss, R. M. Tsohis, and A. J. Bäumler, “Dysbiotic Proteobacteria expansion: a microbial signature of epithelial dysfunction,” *Curr. Opin. Microbiol.*, vol. 39, pp. 1–6, Oct. 2017, doi: 10.1016/J.MIB.2017.07.003.

- [285] S. E. Winter *et al.*, “Gut inflammation provides a respiratory electron acceptor for Salmonella,” *Nature*, vol. 467, no. 7314, pp. 426–429, Sep. 2010, doi: 10.1038/NATURE09415.
- [286] S. E. Winter, C. A. Lopez, and A. J. Bäumlner, “The dynamics of gut-associated microbial communities during inflammation,” *EMBO Rep.*, vol. 14, no. 4, pp. 319–327, Apr. 2013, doi: 10.1038/EMBOR.2013.27.
- [287] S. E. Winter *et al.*, “Host-derived nitrate boosts growth of E. coli in the inflamed gut,” *Science*, vol. 339, no. 6120, pp. 708–711, Feb. 2013, doi: 10.1126/SCIENCE.1232467.
- [288] P. H. Bradley and K. S. Pollard, “Proteobacteria explain significant functional variability in the human gut microbiome,” *Microbiome*, vol. 5, no. 1, 2017, doi: 10.1186/S40168-017-0244-Z.
- [289] M. Wouthuyzen-Bakker, F. A. J. A. Bodewes, and H. J. Verkade, “Persistent fat malabsorption in cystic fibrosis; lessons from patients and mice,” *J. Cyst. Fibros.*, vol. 10, no. 3, pp. 150–158, May 2011, doi: 10.1016/J.JCF.2011.03.008.
- [290] L. R. Hoffman *et al.*, “Escherichia coli dysbiosis correlates with gastrointestinal dysfunction in children with cystic fibrosis,” *Clin. Infect. Dis.*, vol. 58, no. 3, pp. 396–399, Feb. 2014, doi: 10.1093/CID/CIT715.
- [291] S. Matamouros *et al.*, “Adaptation of commensal proliferating Escherichia coli to the intestinal tract of young children with cystic fibrosis,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 7, pp. 1605–1610, Feb. 2018, doi: 10.1073/PNAS.1714373115.
- [292] D. A. Garsin, “Ethanolamine utilization in bacterial pathogens: Roles and regulation,” *Nature Reviews Microbiology*, vol. 8, no. 4. NIH Public Access, pp. 290–295, Apr. 2010, doi: 10.1038/nrmicro2334.
- [293] Y. L. Huang, C. Chassard, M. Hausmann, M. Von Itzstein, and T. Hennet, “Sialic acid catabolism drives intestinal inflammation and microbial dysbiosis in mice,” *Nat. Commun.*, vol. 6, Aug. 2015, doi: 10.1038/NCOMMS9141.
- [294] S. Kitamoto *et al.*, “Dietary L-serine confers a competitive fitness advantage to Enterobacteriaceae in the inflamed gut,” *Nat. Microbiol.*, vol. 5, no. 1, p. 116, Jan. 2020, doi: 10.1038/S41564-019-0591-6.
- [295] L. I. PIZER and M. L. POTOCHNY, “NUTRITIONAL AND REGULATORY ASPECTS OF SERINE METABOLISM IN ESCHERICHIA COLI,” *J. Bacteriol.*, vol. 88, no. 3, pp. 611–619, 1964, doi: 10.1128/JB.88.3.611-619.1964.
- [296] D. M. Roof and J. R. Roth, “Ethanolamine utilization in Salmonella typhimurium,” *J. Bacteriol.*, vol. 170, no. 9, pp. 3855–3863, 1988, doi: 10.1128/JB.170.9.3855-

3863.1988.

- [297] M. M. Kendall, C. C. Gruber, C. T. Parker, and V. Sperandio, “Ethanolamine controls expression of genes encoding components involved in interkingdom signaling and virulence in enterohemorrhagic *Escherichia coli* O157:H7,” *MBio*, vol. 3, no. 3, 2012, doi: 10.1128/MBIO.00050-12.
- [298] C. A. Rowley, C. J. Anderson, and M. M. Kendall, “Ethanolamine Influences Human Commensal *Escherichia coli* Growth, Gene Expression, and Competition with Enterohemorrhagic *E. coli* O157:H7,” 2018, doi: 10.1128/mBio.
- [299] O. Tsoy, D. Ravcheev, and A. Mushegian, “Comparative genomics of ethanolamine utilization,” *J. Bacteriol.*, vol. 191, no. 23, pp. 7157–7164, Dec. 2009, doi: 10.1128/JB.00838-09.
- [300] K. G. Kaval and D. A. Garsin, “Ethanolamine Utilization in Bacteria.,” *MBio*, vol. 9, no. 1, 2018, doi: 10.1128/mBio.00066-18.
- [301] D. M. Roof and J. R. Roth, “Autogenous regulation of ethanolamine utilization by a transcriptional activator of the eut operon in *Salmonella typhimurium*,” *J. Bacteriol.*, vol. 174, no. 20, pp. 6634–43, Oct. 1992, doi: 10.1128/jb.174.20.6634-6643.1992.
- [302] D. H. Luzader, D. E. Clark, L. A. Gonyar, and M. M. Kendall, “EutR is a direct regulator of genes that contribute to metabolism and virulence in enterohemorrhagic *Escherichia coli* O157:H7.,” *J. Bacteriol.*, vol. 195, no. 21, pp. 4947–53, Nov. 2013, doi: 10.1128/JB.00937-13.
- [303] J. T. Penrod, C. C. Mace, and J. R. Roth, “A pH-sensitive function and phenotype: evidence that EutH facilitates diffusion of uncharged ethanolamine in *Salmonella enterica*,” *J. Bacteriol.*, vol. 186, no. 20, pp. 6885–6890, Oct. 2004, doi: 10.1128/JB.186.20.6885-6890.2004.
- [304] D. Patel and S. N. Witt, “Ethanolamine and Phosphatidylethanolamine: Partners in Health and Disease,” *Oxidative Medicine and Cellular Longevity*, vol. 2017. Hindawi Limited, 2017, doi: 10.1155/2017/4829180.
- [305] T. A. Bobik, B. P. Lehman, and T. O. Yeates, “Bacterial microcompartments: widespread prokaryotic organelles for isolation and optimization of metabolic pathways,” *Mol. Microbiol.*, vol. 98, no. 2, pp. 193–207, Oct. 2015, doi: 10.1111/MMI.13117.
- [306] S. Tanaka, M. R. Sawaya, and T. O. Yeates, “Structure and mechanisms of a protein-based organelle in *Escherichia coli*,” *Science*, vol. 327, no. 5961, pp. 81–84, Jan. 2010, doi: 10.1126/SCIENCE.1179513.

- [307] J. T. Penrod and J. R. Roth, “Conserving a volatile metabolite: a role for carboxysome-like organelles in *Salmonella enterica*,” *J. Bacteriol.*, vol. 188, no. 8, pp. 2865–2874, Apr. 2006, doi: 10.1128/JB.188.8.2865-2874.2006.
- [308] S. R. Brinsmade, T. Paldon, and J. C. Escalante-Semerena, “Minimal Functions and Physiological Conditions Required for Growth of *Salmonella enterica* on Ethanolamine in the Absence of the Metabolosome,” *J. Bacteriol.*, vol. 187, no. 23, p. 8039, Dec. 2005, doi: 10.1128/JB.187.23.8039-8046.2005.
- [309] J. Burisch, T. Jess, M. Martinato, and P. L. Lakatos, “The burden of inflammatory bowel disease in Europe,” *J. Crohns. Colitis*, vol. 7, no. 4, pp. 322–337, May 2013, doi: 10.1016/J.CROHNS.2013.01.010.
- [310] J. Galloway *et al.*, “The impact of disease severity and duration on cost, early retirement and ability to work in rheumatoid arthritis in Europe: an economic modelling study,” *Rheumatol. Adv. Pract.*, vol. 4, no. 2, 2020, doi: 10.1093/RAP/RKAA041.
- [311] U. Gophna, K. Sommerfeld, S. Gophna, W. F. Doolittle, and S. J. O. Veldhuyzen Van Zanten, “Differences between tissue-associated intestinal microfloras of patients with Crohn’s disease and ulcerative colitis,” *J. Clin. Microbiol.*, vol. 44, no. 11, pp. 4136–4141, Nov. 2006, doi: 10.1128/JCM.01004-06.
- [312] G. S. Jutley *et al.*, “Relationship Between Inflammation and Metabolism in Patients With Newly Presenting Rheumatoid Arthritis,” *Front. Immunol.*, vol. 12, Sep. 2021, doi: 10.3389/FIMMU.2021.676105.
- [313] Y. Wang *et al.*, “Plasma Metabolic Profiling Analysis of Gout Party on Acute Gout Arthritis Rats Based on UHPLC-Q-TOF/MS Combined with Multivariate Statistical Analysis,” *Int. J. Mol. Sci.*, vol. 20, no. 22, Nov. 2019, doi: 10.3390/IJMS20225753.
- [314] V. Malmström, A. I. Catrina, and L. Klareskog, “The immunopathogenesis of seropositive rheumatoid arthritis: from triggering to targeting,” *Nat. Rev. Immunol.*, vol. 17, no. 1, pp. 60–75, Jan. 2017, doi: 10.1038/NRI.2016.124.
- [315] A. Kalinkovich and G. Livshits, “A cross talk between dysbiosis and gut-associated immune system governs the development of inflammatory arthropathies,” *Semin. Arthritis Rheum.*, vol. 49, no. 3, pp. 474–484, Dec. 2019, doi: 10.1016/J.SEMARTHRT.2019.05.007.
- [316] F. H. Epstein and E. D. Harris, “Rheumatoid arthritis. Pathophysiology and implications for therapy,” *N. Engl. J. Med.*, vol. 322, no. 18, pp. 1277–1289, May 1990, doi: 10.1056/NEJM199005033221805.
- [317] R. Nissinen *et al.*, “Immune activation in the small intestine in patients with rheumatoid

- arthritis,” *Ann. Rheum. Dis.*, vol. 63, no. 10, pp. 1327–1330, Oct. 2004, doi: 10.1136/ARD.2003.011304.
- [318] C. Heßlinger, S. A. Fairhurst, and G. Sawers, “Novel keto acid formate-lyase and propionate kinase enzymes are components of an anaerobic pathway in *Escherichia coli* that degrades L-threonine to propionate,” *Mol. Microbiol.*, vol. 27, no. 2, pp. 477–492, 1998, doi: 10.1046/J.1365-2958.1998.00696.X.
- [319] N. O’Boyle, J. P. R. Connolly, N. P. Tucker, and A. J. Roe, “Genomic plasticity of pathogenic *Escherichia coli* mediates d-serine tolerance via multiple adaptive mechanisms,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 36, p. 22484, Sep. 2020, doi: 10.1073/PNAS.2004977117.
- [320] R. L. Moritz and R. A. Welch, “The *Escherichia coli* argW-dsdCXA genetic island is highly variable, and *E. coli* K1 strains commonly possess two copies of dsdCXA,” *J. Clin. Microbiol.*, vol. 44, no. 11, pp. 4038–4048, Nov. 2006, doi: 10.1128/JCM.01172-06.
- [321] K. Jahreis *et al.*, “Adaptation of Sucrose Metabolism in the *Escherichia coli* Wild-Type Strain EC3132†,” *J. Bacteriol.*, vol. 184, no. 19, p. 5307, Oct. 2002, doi: 10.1128/JB.184.19.5307-5316.2002.
- [322] G. L. Marcone, E. Rosini, E. Crespi, and L. Pollegioni, “D-amino acids in foods,” *Appl. Microbiol. Biotechnol.*, vol. 104, no. 2, pp. 555–574, Jan. 2020, doi: 10.1007/S00253-019-10264-9.
- [323] T. Asakawa *et al.*, “Oral administration of D-serine prevents the onset and progression of colitis in mice,” *J. Gastroenterol.*, vol. 56, no. 8, pp. 732–745, Aug. 2021, doi: 10.1007/S00535-021-01792-1.
- [324] J. P. R. Connolly *et al.*, “The host metabolite D-serine contributes to bacterial niche specificity through gene selection,” *ISME J.*, vol. 9, no. 4, p. 1052, Mar. 2015, doi: 10.1038/ISMEJ.2015.17.
- [325] Y. Bertin *et al.*, “Enterohaemorrhagic *Escherichia coli* gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content,” *Environ. Microbiol.*, vol. 13, no. 2, pp. 365–377, Feb. 2011, doi: 10.1111/J.1462-2920.2010.02334.X.
- [326] M. J. Ormsby *et al.*, “Inflammation associated ethanolamine facilitates infection by Crohn’s disease-linked adherent-invasive *Escherichia coli*,” *EBioMedicine*, vol. 43, p. 325, May 2019, doi: 10.1016/J.EBIOM.2019.03.071.
- [327] C. J. Anderson, J. Satkovich, V. K. Köseoğlu, H. Agaisse, and M. M. Kendall, “The

Ethanolamine Permease EutH Promotes Vacuole Adaptation of *Salmonella enterica* and *Listeria monocytogenes* during Macrophage Infection,” *Infect. Immun.*, vol. 86, no. 5, May 2018, doi: 10.1128/IAI.00172-18.

- [328] J. Shan *et al.*, “Integrated Serum and Fecal Metabolomics Study of Collagen-Induced Arthritis Rats and the Therapeutic Effects of the Zushima Tablet,” *Front. Pharmacol.*, vol. 9, no. AUG, Aug. 2018, doi: 10.3389/FPHAR.2018.00891.

ABSTRACT

An imbalance of the gut microbiota called dysbiosis has been linked to several inflammatory diseases such as Crohn's disease (CD) and Rheumatoid arthritis (RA). Dysbiosis is often characterized by an expansion of *Enterobacteriaceae*, a bacterial family that normally represents only a few percentage of the gut microbiota in healthy individuals (HI). To explain proliferation of this bacterial family during dysbiosis, one hypothesis is based on the metabolic adaptation of *Enterobacteriaceae* to the disturbed gut environment. The global aim of my PhD work was to screen for nutritional substrates that could give a metabolic advantage to *Enterobacteriaceae* for expansion in the dysbiotic gut of patients suffering from CD or RA. Using a phenotype microarray strategy, the utilization of 190 different substrates as a carbon source was compared between *Enterobacteriaceae* strains isolated from fecal samples of either HI or from patients suffering from CD or RA. Among the identified substrates, we observed an opposite ability of HI and CD strains to metabolize D-serine and sucrose, a characteristic arising from genetic rearrangements in the *argW* locus of *Escherichia coli* chromosome. The D-serine⁺ phenotype of CD strains was shown to be dependent of the presence of *dsdCXA* genes and to confer a nutritional advantage over HI strains *in vitro*. We also evaluated the ability of our strain collections to catabolize ethanolamine (EA), a small organic molecule naturally present in the human gut with an increased concentration in case of gut inflammation. Whereas no specific difference was observed between the 3 strain collections, the ability to consume this substrate was highly diverse between isolates. Strains unable to use EA present either mutations in the *eut* locus responsible for EA degradation or a defect in the gene transcription process. Isolates consuming EA can use the substrate as a nitrogen source only or as nitrogen and carbon sources. Using a mouse model of gut colonization and a new developed gene reporter tool, we also demonstrated that *eut* genes are expressed in the gut and that EA utilization is essential for an optimal fitness of commensal *E. coli* *in vivo*. With the example of EA, my work highlights the high diversity of phenotypes occurring between gut isolates of the species *E. coli*, illustrating how *Enterobacteriaceae* are a major source of variable genes in the gut microbiome. Importantly, our study provides new clues on potential substrates that could promote the expansion of *Enterobacteriaceae* in dysbiotic conditions and could open new doors on therapeutic approaches in order to restore gut homeostasis and health in patients.