

**DÉVELOPPEMENT D'UN PROBIOTIQUE FONCTIONNALISÉ COMME
ANTIBIOTIQUE HAUTEMENT SPÉCIFIQUE**

par

Kevin Neil

Thèse présentée au Département de biologie en vue
de l'obtention du grade de docteur ès sciences (Ph. D.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, janvier 2021

Le 25 janvier 2021

Le jury a accepté la thèse de Kevin Neil dans sa version finale.

Membres du jury

Professeur Sébastien Rodrigue
Directeur de recherche
Département de biologie
Université de Sherbrooke

Professeur Vincent Burrus
Codirecteur de recherche
Département de biologie
Université de Sherbrooke

Professeur David Bikard
Évaluateur externe
Département de microbiologie
Institut Pasteur

Professeur Alfredo Menendez
Évaluateur interne
Département de microbiologie et d'infectiologie
Université de Sherbrooke

Professeur François Malouin
Président-rapporteur
Département de biologie
Université de Sherbrooke

REMERCIEMENTS

J'aimerais d'abord remercier de tout cœur mon directeur de recherche, le Pr. Sébastien Rodrigue, pour m'avoir permis de m'épanouir lors de mes études tant au niveau de la maîtrise que du doctorat. Un véritable leader par l'exemple, son dévouement à la recherche a été une grande source de motivation tout au long de mes études graduées. Armé de son calme, sa patience, sa rigueur, sa disponibilité et sa confiance, il a su m'encadrer de façon totalement irréprochable et m'offrir la liberté dont j'avais besoin pour me développer en tant que chercheur. Je me considère chanceux d'avoir pu apprendre sous sa direction et garderai d'excellents souvenirs des années passées sous sa supervision.

Je voudrais également remercier mon co-directeur Pr. Vincent Burrus, ainsi que mes conseillers Pr. François Malouin et Pr. Alfredo Menendez pour leur soutien. Par leur vaste expertise, ils ont su mettre au défi mes idées et apporter des solutions aux difficultés rencontrées. Nos nombreuses rencontres et discussions ont définitivement contribué positivement au succès de mon projet doctoral. Je tiens également à remercier Pr. David Bikard pour l'évaluation de ma thèse.

Je veux également remercier les membres du laboratoire : Joelle, Vincent, Dominick, Frédéric, Jean-Christophe, Nancy, Antoine², Catherine, Nicolas, Patricia, David, Olivier, Anthony, Julien et Maire-Pier pour leur camaraderie et leur support tout au long du projet. J'ai eu un plaisir fou à travailler avec chacun d'eux. J'aimerais remercier mon partenaire d'affaire Jean-François Millau pour son dévouement au projet et sa confiance. Je remercie spécialement ma conjointe Kassandra Mole pour son support durant mes études. Malgré mes longues heures au laboratoire, elle a su être présente pour m'épauler et m'aider à surmonter les nombreux défis que représente le doctorat et embellir mes journées. Je tiens également à remercier ma famille, mon père Robert Neil, ma mère Martine Jallon, mon frère Philippe Neil et ma sœur Sophie Neil qui ont toujours cru en moi.

“The greatest gift life has to offer is the opportunity to work hard at work worth doing.”

-Theodore Roosevelt

SOMMAIRE

La biologie synthétique est une discipline appliquant les principes d'ingénierie à la conception de circuits génétiques complexes ayant pour but d'exercer de nouvelles fonctions non retrouvées dans la nature. Ces approches ont rapidement suscité un grand intérêt dans le domaine médical où bon nombre de systèmes furent créés afin de répondre à des besoins non comblés par des approches classiques. Un des défis médicaux d'importance des prochaines années concerne le traitement des infections bactériennes résistantes aux antibiotiques. L'accumulation rapide de gènes de résistance, combinée au ralentissement des efforts de développement de nouveaux antibiotiques par les compagnies pharmaceutiques menacent de créer une grande crise sanitaire d'ici 2050. Afin de répondre à ce défi, l'utilisation de CRISPR-Cas9 pour éliminer spécifiquement des bactéries résistantes aux antibiotiques fut proposée. En utilisant des approches de biologie synthétique et de génomique, il est maintenant possible d'utiliser Cas9 pour discriminer des souches bactériennes selon leurs séquences d'ADN. En ciblant des séquences uniquement retrouvées dans des souches problématiques (par exemple des gènes de résistance aux antibiotiques, des gènes de virulences, des toxines), il est possible d'éliminer ces bactéries d'un mélange complexe sans affecter les autres membres de la communauté bactérienne. Le principal défi de ce type d'approche concerne la méthode de livraison de CRISPR-*cas9*. L'efficacité globale du système sera dépendante de la capacité de livraison du complexe CRISPR-*cas9* aux cellules cibles.

Durant mon doctorat, j'ai travaillé au développement d'un outil de livraison de CRISPR-*cas9* hautement performant dans le microbiote entérique. Cet outil exploite la conjugaison bactérienne comme véhicule de livraison des gènes encodant le système CRISPR-Cas9 à des bactéries directement dans l'intestin. Le système CRISPR-*cas9* est alors exprimé dans la cellule réceptrice, interroge son génome pour une séquence cible programmable et la clive, menant ainsi à la perte d'intégrité du génome et en la mort de la cellule cible. Afin de développer cet outil, j'ai d'abord effectué le criblage de 13 systèmes conjugatifs pour leur efficacité de transfert

d'ADN directement dans l'intestin d'un modèle de souris. Ces criblages m'ont permis d'identifier TP114, un plasmide de la famille d'incompatibilité (Inc) I₂ pouvant se transférer à virtuellement toutes les *Escherichia coli* réceptrices sondés dans nos tests. L'étude de TP114 m'a ensuite permis d'identifier un mécanisme essentiel à la conjugaison bactérienne dans l'intestin, soit la stabilisation de la paire de conjugaison. Chez TP114, ce mécanisme est exercé par un pilus de type IV qui est essentiel pour la conjugaison *in situ*, mais pas pour la conjugaison sur milieu de culture en agar.

J'ai ensuite travaillé sur l'insertion du système CRISPR-*cas9* directement dans TP114. Pour ce faire, il me fallait manipuler deux molécules d'ADN pour forcer leur fusion de façon dirigée et sans stimuler la recombinaison homologue afin d'éviter que des systèmes composés de plusieurs ARN guide (ARNg) recombinent sur eux-mêmes et perdent des sections de séquences. Pour ce faire, j'ai mis au point une technique d'assemblage d'ADN *in vivo* appelé DROID qui utilise la recombinaison site-spécifique dépendante de Bxb1 afin de médier la fusion des deux molécules d'ADN. Ensuite, des sites FRT adéquatement placés permettaient la délétion de marqueurs de résistance aux antibiotiques et autres éléments génétiques superflus. La méthode s'est montrée très efficace et robuste avec des taux de succès d'insertion de >99 %.

Finalement, j'ai été en mesure de développer le système COP (*CO*njugative *PR*obiotic), un probiotique exploitant la grande efficacité conjugative de TP114 *in situ* dans l'intestin pour mobiliser un système CRISPR-*cas9*. Après avoir conçu un ARNg ciblant le gène de résistance au chloramphénicol, j'ai ensuite effectué une preuve de concept où un mélange de bactéries cibles et non-cibles était introduit en quantité égale dans le système digestif de la souris. Ce premier système exploitant les capacités naturelles de TP114 a été en mesure d'éliminer 98 % des bactéries cibles en 3 jours. Afin d'améliorer l'efficacité et la dynamique de conjugaison de TP114, j'ai ensuite eu recours à des techniques d'évolution accélérée afin d'obtenir un mutant de TP114 présentant des performances accrues. Ce système COP amélioré a ensuite atteint une efficacité d'élimination spécifique de 99,5 % en moins de 24 heures avec une seule dose.

L'ensemble de ces résultats démontrent le potentiel de la technologie COP comme alternative aux antibiotiques pour le traitement d'infections entériques. Les expériences de preuve de concept ont d'ailleurs permis l'application pour un brevet maintenant en phases nationales et la création d'une entreprise du secteur des biotechnologies. Les données accumulées sur TP114 permettront, dans les travaux futurs, de raffiner le système en lui ajoutant des mesures de bioconfinement et en continuant d'améliorer ses propriétés.

Mots-clés : biologie synthétique, génie génétique, conjugaison bactérienne, biothérapeutiques vivants, probiotiques, *Escherichia coli*.

TABLE DES MATIÈRES

CHAPITRE 1 INTRODUCTION	1
1.1 La manipulation génétique au service de l'humanité.....	1
1.1.1. L'évolution rapide des outils de manipulation génétique.....	1
1.1.2. Les endonucléases programmables : la révolution CRISPR	4
1.2 L'application de la biologie synthétique au domaine médical.....	7
1.2.1. La biologie synthétique appliquée à la production de molécules pharmaceutiques	9
1.2.2. La biologie synthétique comme outil diagnostique	11
1.2.3. La biologie synthétique appliquée à la médecine régénérative	15
1.2.4. La biologie synthétique comme outil prophylactique ou thérapeutique.....	16
1.3 Le microbiome	19
1.3.1. Plus qu'une population bactérienne	22
1.3.2. Fonctions du microbiome	23
1.3.3. Technologies de modification de la composition du microbiome.....	24
1.3.3.1 Les organismes sans germes	25
1.3.3.2 Les souris sans pathogènes spécifiques (SPS)	26
1.3.3.3 Les souris gnotobiotiques.....	26
1.3.3.4 Les traitements antibiotiques.....	27
1.3.3.5 Les transplantations fécales.....	28
1.3.3.6 L'ajout de probiotiques	29
1.3.3.7 Les bactériophages comme outil de modification du microbiote.	30
1.3.3.8 La résistance aux antibiotiques et le microbiome entérique.	32
1.4 La conjugaison bactérienne.....	35
1.4.1. Le spectre d'hôte des plasmides conjugatifs.....	36

1.4.2.	La conjugaison bactérienne dans le microbiome entérique	39
1.5	Objectifs et hypothèses du projet de recherche	41
CHAPITRE 2 Highly efficient gene transfer in the mouse gut microbiota is enabled by the IncI₂ conjugative plasmid TP114		45
2.1	Présentation de l'article et contributions	45
2.2	Title Page.....	47
2.3	Abstract:	48
2.4	Introduction:	49
2.5	Results:	50
2.5.1.	<i>In vitro</i> transfer rates quantification.....	50
2.5.2.	Establishment of a mouse model.	51
2.5.3.	Quantification of <i>in situ</i> DNA transfer rates.....	52
2.5.4.	Highly effective transfer of TP114 <i>in situ</i>	53
2.5.5.	High-density transposon mutagenesis (HDTM) of TP114.	54
2.5.6.	Identification of essential genes for <i>in situ</i> conjugation of TP114.	55
2.5.7.	Comparative genomics in IncI plasmids.....	56
2.5.8.	Activation of TP114 under anaerobic conditions.	57
2.6	Discussion:	58
2.7	Methods:.....	61
2.7.1.	Strains, plasmids and growth conditions.	61
2.7.2.	DNA manipulations.	61
2.7.3.	DNA purification.	62
2.7.4.	DNA transformation into <i>E. coli</i> by electroporation.	62
2.7.5.	DNA transformation into <i>E. coli</i> by heat-shock.	63
2.7.6.	Recombineering.	63
2.7.7.	Generation of modified EcN strains.	63
2.7.8.	<i>In vitro</i> conjugation assays.	64
2.7.9.	Mouse Model.	65
2.7.10.	<i>In situ</i> conjugation assays.	66

2.7.11.	DNA sequencing.....	66
2.7.12.	Analysis of plasmid gene function and generation of phylogenetic trees.	66
2.7.13.	High-density transposon mutagenesis.	67
2.7.14.	Mouse model for <i>in situ</i> HDTM library conjugation.....	67
2.7.15.	HDTM libraries analysis.....	68
2.7.16.	Comparative genomics using BRIGG.	69
2.7.17.	Statistics and reproducibility.	69
2.7.18.	Data availability.....	69
2.8	References:.....	70
2.9	End notes:.....	74
2.9.1.	Acknowledgments:	74
2.9.2.	Author contributions:.....	74
2.9.3.	Competing interests:	74
2.10	Supporting information	75
2.10.1.	Page titre	75
2.10.2.	Supplementary Methods	76
2.10.2.1	Construction of <i>E. coli</i> KN01 Δ dapA.	76
2.10.2.2	Generation of rifampicin and nalidixic resistant mutants of EcN.....	76
2.10.2.3	Mice inoculum preparation.	77
2.10.2.4	Feces sampling.	77
2.10.2.5	Mice dissection and EcN colonization pattern assessment.	78
2.10.2.6	TP114 sequencing.	78
2.10.2.7	Sequencing and annotation of other conjugative plasmids.	79
2.10.2.8	HDTM libraries sequencing.....	79
2.10.2.9	Deletion of pilS in TP114.....	80
2.10.2.10	Induction of pilS expression.....	80
2.10.3.	Supplementary References:	100

CHAPITRE 3 Assembly of Large Mobilizable Genetic Cargo by Double Recombinase Operated Insertion of DNA (DROID)101

3.1	Présentation de l'article et contributions.....	101
3.2	Page titre.....	103
3.3	Abstract:	104
3.4	Introduction	106
3.5	Material and Methods.....	108
3.5.1.	Bacterial strains, plasmids and growth conditions.....	108
3.5.2.	DNA manipulations	108
3.5.3.	Construction of the loading docks pREC1 and pBXB1	109
3.5.4.	Construction of the CRISPR-Cas9 modules pKill1 and pKill3.....	110
3.5.5.	Loading CRISPR- <i>cas9</i> into TP114 by DROID	111
3.5.6.	Plasmid pT copy number determination.....	112
3.5.7.	CRISPR-Cas9 module activity test.....	112
3.5.8.	Statistical analysis.....	113
3.6	Results and discussion.....	113
3.6.1.	DROID method rationale.....	113
3.6.2.	Insertion of the loading dock in TP114	115
3.6.3.	Engineering of two functional CRISPR-Cas9 systems.....	117
3.6.4.	Loading the CRISPR-cas9 systems on TP114:: <i>tetB</i>	118
3.6.5.	FLP mediated deletion of antibiotics markers	122
3.6.6.	TP114 can perform in-cis mobilization of CRISPR-Cas9 systems	123
3.6.7.	DROID is a potent DNA assembly method that is readily adaptable to multiple applications.....	125
3.7	Concluding remarks	129
3.8	Acknowledgments.....	129
3.9	References	129
3.10	Supplementary Data.....	133

CHAPITRE 4 High-efficiency CRISPR-Cas9 Delivery by Engineered Bacteria for Precise Elimination of Antibiotic Resistant *Enterobacteriaceae*.....136

4.1	Présentation de l'article et contributions.....	136
-----	---	-----

4.2	Page titre.....	138
4.3	Abstract:	139
4.4	Main Text:	139
4.5	Methods:.....	148
4.5.1.	Strains, plasmids and growth conditions.	148
4.5.2.	DNA manipulations.	148
4.5.3.	DNA transformation into <i>E. coli</i> by electroporation.	149
4.5.4.	<i>In vitro</i> conjugation assays.	149
4.5.5.	Mouse Model.	150
4.5.6.	Accelerated Laboratory Evolution.....	151
4.5.7.	16S ribotyping.....	152
4.5.8.	Statistics and reproducibility.	153
4.5.9.	Data and materials availability.....	153
4.6	Author contributions:	153
4.7	Acknowledgments:.....	154
4.8	Competing interests:.....	154
4.9	References:	154
4.10	Extended Data Document.....	157
4.11	Supplementary references:	169
CHAPITRE 5 DISCUSSION GÉNÉRALE ET CONCLUSION.....		170
5.1	Discussion et conclusion générale.....	170
5.2	Comprendre TP114 pour mieux le contrôler.....	171
5.3	Bioconfinement de la souche COP.....	174
5.4	Facteurs limitant l'efficacité d'élimination du système COP	176
5.5	Perspective	180
BIBLIOGRAPHIE		182

LISTE DES ABRÉVIATIONS

°C	Degrés Celsius
5'RACE	<i>Rapid Amplification of cDNA 5' Ends</i>
ADN	Acide Désoxyribonucléique
ADNr	ADN ribosomal
antiSMASH	<i>Secondary Metabolite Analysis SHell</i>
ApE	<i>A plasmid Editor</i>
Ap ^R	Résistance à l'ampiciline
ARN	Acide Ribonucléique
ARNcr	ARN CRISPR
ARNg	ARN guide
ARNsi	Petit ARN interférant
ARNt	ARN de transfert
ARNtracr	ARN transactivant CRISPR
BAC	<i>Bacterial Artificial Chromosome</i> – Chromosome Bactérien Artificiel
BLAST	<i>Basic Local Alignment Search Tool</i>
Bp	<i>Base Pair</i>
BRIG	BLAST Ring Image Generator
CARD	<i>Comprehensive Antibiotic Resistance Database</i>
CAT#	Numéro de catalogue
CDsearch	<i>Conserved Domain Search</i>
CFU	<i>Colony Forming Unit</i> – Unité Formatrice de Colonies

ChIP	<i>Chromatin Immuno-Precipitation</i> – Immuno-Précipitation de la Chromatine
CIHR	<i>Canadian Institute of Health Research</i> – Institut de Recherche sur la Santé Canadien
Cm ^R	Résistance au chloramphénicol
COP	<i>Conjugative Probiotic</i> – Probiotique Conjugatif
CRISPR	<i>Clustered Regularly Interspaced Short Palindromic Region</i>
DAP	<i>Diaminopimelic acid</i> – Acide Diaminopimelique
DROID	<i>Double Recombinase Operated Insertion of DNA</i>
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	<i>Exempli gratia</i>
EcN	<i>E. coli</i> Nissle1917
Fig.	Figure
FRQNT	Fond de Recherche Québécois – Nature et Technologies
FRQS	Fond de Recherche Québécois – Santé
gDNA	ADN génomique
HDTM	<i>High Density Transposon Mutagenesis</i>
HGT	<i>Horizontal Gene Transfer</i> – Transfert horizontal de gène
ICE	<i>Integrative Conjugative Element</i> – Élément conjugative et intégratif
iGEM	<i>International Genetically Engineered Machines</i>
Inc	<i>Incompability</i> - Incompatibilité
Kbp	<i>Kilobase pair</i>
Km ^R	Résistance à la kanamycine
LB	<i>Luria Broth</i>
LBA	<i>Luria Broth Agar</i>

MetaHIT	<i>Metagenomics of the Human Intestinal Tract</i>
MOB	<i>Mobility</i> – Mobilité
MPF	<i>Mating Pair Formation</i> – Formation de la paire de conjugaison
MPS	<i>Mating Pair Stabilization</i> – Stabilisation de la paire de conjugaison
n	Nombre de réplicat
NGS	<i>Next Generation Sequencing</i> – Séquençage de nouvelle génération
NHEJ	<i>Non-Homologous End Joining</i> – Jonction des extrémités non-homologues
NSERC	<i>National Science and Engineering Research Council of Canada</i>
OD _{600nm}	Optical Density at 600 nm – Densité optique à 600 nm
ORF	Open Reading Frame – Cadre de lecture
<i>oriT</i>	<i>Origin of transfer (or origin of transference)</i> – Origine de transfert
<i>oriV</i>	<i>Origin of vegetative replication</i> – Origine de réplication végétative
p.ex.	Par exemple
PAM	<i>Protospacer Adjacent Motif</i> – Motif adjacent au protospacer
pb	Paire de base
PBS	<i>Phosphate Buffered Saline</i> – Saline tamponné au phosphate
PCR	<i>Polymerase Chain Reaction</i> – Réaction en chaîne de polymérase
pNT	<i>Plasmid Non Target</i> – Plasmide non ciblé
pT	<i>Plasmid Target</i> – Plasmid ciblé
QC	Québec
RAST	<i>Rapid Annotation using Subsystems Technology</i>
Sm ^R	Résistance à la Streptomycine
SPS	Souris sans pathogène spécifique

T4P	<i>Type IV pilus</i> – Pilus de type IV
T4SS	<i>Type IV Secretion System</i> – Système de sécrétion de type IV
TALEN	<i>Transcription Activator-Like Effector Nuclease</i>
Tc ^R	Résistance à la tétracycline
USER	<i>Uracil Specific Excision Reagent</i>
ZFN	<i>Zinc Finger Nuclease</i> – Nucléase à doigt de zinc

LISTE DES TABLEAUX

Tableau 1.1	Diversité des études sur le microbiome	20
Tableau 1.2	Spectre d'hôte des familles de systèmes conjugatifs	38
Supplementary Table 2.1	Conjugative plasmids used in this study	88
Supplementary Table 2.2	High Density Transposon Mutagenesis Tnseq results	89
Supplementary Table 2.3	List of strains, plasmids and primers used in this study	91
Supplementary Data 2.1	TP114 <i>in silico</i> analysis of gene function, essentiality and conservation	95
Table 3.1	DROID protocol summary	128
Table S3.1	List of strains and plasmid used in this study	133
Table S3.2	Oligonucleotides used in this study	134
Extended Data Table 4.1	List of strains and plasmids used in this study	168
Extended Data 4.1	Position and putative impact of mutations found by sequencing accelerated evolution mutant libraries and clones	169

LISTE DES FIGURES

Figure 1.1	Classification des différents types de CRISPR-Cas	6
Figure 1.2	Cycle d'optimisation typique de la biologie synthétique	8
Figure 1.3	Production de molécules bioactives chez des organismes modifiés par approche de biologie synthétique	10
Figure 1.4	Exemples de modules de biologie synthétique comme outil diagnostique	14
Figure 1.5	Définition possible d'un microbiome sain	22
Figure 2.1	Comparison of the conjugation rates of 13 plasmids under different conditions	51
Figure 2.2	Evaluation of mouse model parameters for <i>in situ</i> conjugation	53
Figure 2.3	Mating Pair Stabilization (MPS) is essential for <i>in situ</i> conjugation	55
Figure 2.4	TP114 transfer is activated under hypoxic conditions	57
Supplementary Figure 2.1	Specific features of strains used in the conjugation mouse model	82
Supplementary Figure 2.2	EcN colonization optimization	83
Supplementary Figure 2.3	Transfer rates of 13 plasmids in the mouse intestinal tract	84
Supplementary Figure 2.4	Analysis of TP114 gene essentiality by high-density transposon mutagenesis (HDTM)	85
Supplementary Figure 2.5	Sequence homology between TP114 and other plasmids of the IncI family	86
Supplementary Figure 2.6	TP114 gene essentiality and conservation among IncI2 plasmid family members	87
Figure 3.1	Plasmid maps and design	110

Figure 3.2	Key steps of the DROID method	114
Figure 3.3	Loading dock insertion have no negative effect on TP114 conjugation	116
Figure 3.4	The CRISPR-Cas9 systems plasmids pKill1 and pKill3 specifically cleave their target sequence	118
Figure 3.5	Bxb1 allows efficient manipulation of large DNA molecules	120
Figure 3.6	The FRT recombination by the FLP flipase removes any unnecessary genes from the final construction	122
Figure 3.7	TP114 can deliver functional CRISPR-Cas9 systems in cis to specifically cure an antibiotic resistance plasmid	124
Figure 3.8	Marker-free insertion of large payloads in the chromosome	126
Figure 4.1	Conjugative probiotics (COPs) can transfer CRISPR- <i>cas9</i> at high-efficiency to specifically eliminate a target strain from the gut microbiota	142
Figure 4.2	Accelerated Laboratory Evolution can boost TP114 transfer rates	144
Figure 4.3	The eB-COP system can reshape bacterial populations in the gut as followed in feces	147
Extended Data Figure 4.1	COPs can selectively eliminate a target strain from a mixture <i>in vitro</i>	158
Extended Data Figure 4.2	COPs can protect the microbiota from colonization of a specific strain	159
Extended Data Figure 4.3	COP treatment has no major effect on microbiota composition	160
Extended Data Figure 4.4	Accelerated laboratory evolution selected for mutations in specific regions of TP114	161
Extended Data Figure 4.5	Isolated clones for mutagenesis rounds on agar present enriched mutations in specific regions of TP114	163
Extended Data Figure 4.6	Isolated clones for mutagenesis rounds in broth present enriched mutations in specific regions of TP114	164

Extended Data Figure 4.7	Correlation between mutation mapping and transfer rates of TP114 evolved clones	165
Extended Data Figure 4.8	The remaining fraction of target bacteria after eB-COP do not display inheritable resistance to the system	167

CHAPITRE 1

INTRODUCTION

1.1 La manipulation génétique au service de l'humanité

Bien avant l'invention des outils de manipulation génétique, l'humain maniait déjà la génétique des êtres vivants qui l'entourent. Par croisements sélectifs, nous avons su influencer la génétique des plantes que nous cultivons et des animaux que nous élevons pour répondre à nos besoins. En appliquant ces pratiques, l'humanité a su transformer la téosinte en maïs (Kellogg, 1997) et le loup sauvage en chien fidèle (Frantz *et al*, 2016). Par contre, la manipulation génétique par croisement est un travail de longue haleine qui dépend de la sélection de mutations spontanées à chaque génération (Hill & Caballero, 1992). C'est avec les travaux de Paul Berg en 1972 que nous avons obtenu les premiers outils nécessaires à la manipulation directe de l'ADN. En effet, Paul Berg a réussi le premier clonage en insérant de l'ADN bactérien dans le génome du virus SV40, créant ainsi le premier ADN recombinant (Jackson *et al*, 1972). À peine savait-on manipuler l'ADN que Stanley Cohen et Herbert Boyer entreprirent en 1974 le premier clonage d'ADN eucaryote chez *E. coli*, un premier pas important dans leur quête du clonage du gène de l'insuline (Morrow *et al*, 1974). L'approvisionnement en insuline, anciennement extraite du pancréas de porc de l'industrie alimentaire, a su être facilité par la production à grande échelle d'insuline en bioréacteur grâce à leurs efforts. Ces travaux pionniers ont su démontrer le potentiel de la biotechnologie, une discipline qui vise l'application de la science et de la technologie à des organismes vivants afin de produire des connaissances, des biens et des services. Depuis, les outils de manipulations génétiques ont beaucoup évolué, permettant l'assemblage de circuits géniques beaucoup plus complexes.

1.1.1. L'évolution rapide des outils de manipulation génétique

Les premiers clonages dirigés utilisaient des endonucléases appelés enzymes de restriction (Cohen *et al*, 1973). Ces enzymes reconnaissent et clivent des séquences d'ADN spécifiques et produisent différentes extrémités d'ADN aux suites de leur coupure, lesquelles peuvent être classées en

extrémités franches ou cohésives (Chan *et al*, 2011). Les extrémités de molécules d'ADN compatibles peuvent alors être liées ensemble en utilisant une ligase, formant ainsi de nouvelles molécules d'ADN (Modrich & Lehman, 1973). Avec l'invention de la PCR qui suivit quelques années après les premiers clonages par enzymes de restriction, il devint possible de combiner les deux techniques en amplifiant des fragments de génomes par PCR, puis en les digérant pour les assembler (Saiki *et al*, 1985, 1987). Cependant, ce type de technique devient rapidement limitant lorsque les constructions désirées nécessitent de l'assemblage de précision. En effet, les sites de restrictions ne sont pas toujours bien distribués pour permettre le clonage et malgré la production de vecteurs de clonage, nos capacités d'assemblages demeuraient limitées.

C'est pour répondre à ces besoins que des techniques plus poussées d'assemblage d'ADN furent développées. Avec l'arrivée du « USER cloning » (Geu-flores *et al*, 2007), du « Golden Gate » (Engler *et al*, 2008) et de l'assemblage de Gibson (Gibson *et al*, 2009), le clonage n'est alors plus limité par la présence de site de restrictions, permettant l'assemblage de morceaux d'ADN à la limite de notre imagination. Parmi ces techniques, l'assemblage de Gibson a gagné beaucoup de popularité par sa simplicité pour l'utilisateur (Gibson *et al*, 2009). Avec cette technique, il suffit de prévoir des régions d'homologies entre les différents fragments d'ADN pour permettre l'assemblage de jusqu'à 12 morceaux d'ADN à la fois (comme stipulé par le manufacturier NEB #E2621L). L'assemblage de Gibson contient trois enzymes qui travaillent en simultanément à la même température (Gibson *et al*, 2009). D'abord, une 5' exonucléase dégrade l'extrémité des brins d'ADN, ce qui permet aux régions d'homologie de s'hybrider. Ensuite, une polymérase remplit les trous laissés par l'exonucléase et une ligase lie les molécules d'ADN ensemble. La simplicité et l'efficacité de l'assemblage de Gibson permettent maintenant de construire de nouvelles molécules d'ADN rapidement et sans compromis.

Des outils de modification génomiques ont également vu le jour. La plupart de ces outils exploitent la recombinaison afin d'introduire des changements dans la séquence d'ADN endogène de la cellule bactérienne (Lee *et al*, 2009; Datsenko & Wanner, 2000; Datta *et al*, 2006; Reisch & Prather, 2015). Une des techniques les plus utilisées est le « Recombineering », qui exploite la machinerie du phage λ (Datsenko & Wanner, 2000) ou du phage Rac (Youming *et al*, 1998) afin

d'insérer des fragments d'ADN dans la fourche de réplication de manière similaire aux fragments d'Okazaki (Pines *et al*, 2015). Brièvement, un gène de sélection est d'abord amplifié par des amorces portant de l'homologie pour la région génomique à modifier. Le produit PCR est ensuite transformé dans une cellule exprimant les trois gènes de la machinerie du phage λ nécessaire au « Recombineering ». Le premier produit de gène Exo, est une exonucléase qui dégrade les fragments d'ADN introduits dans la cellule par les extrémités, générant ainsi de l'ADN simple brin, ensuite Bet vient lier l'ADN simple brin et facilite l'insertion dans la fourche de réplication, Gam, inhibe RecBCD pour réduire les chances que l'ADN introduit soit dégradé (Sharan *et al*, 2009). La plupart des autres techniques de modification génomiques sont dérivées du « Recombineering » et ajoutent des méthodes de contre-sélection ou des moyens de retirer le marqueur de sélection après le « Recombineering » à l'aide de flipase, d'endonucléases et de gènes toxiques (Datta *et al*, 2006; Reisch & Prather, 2015; Lee *et al*, 2009; Bird *et al*, 2012; St-Pierre *et al*, 2013). Ces techniques permettent la génération rapide de mutants de délétions et l'insertion de fragments d'ADN dans le génome. Des projets d'envergures ont d'ailleurs été réalisés grâce à ces techniques, comme la génération d'une banque de mutant de délétion de chaque gène non essentiel chez *E. coli* appelé la collection Keio (Baba *et al*, 2006) et des souches de *E. coli* ayant un génome réduit (Pósfai *et al*, 2006; Hirokawa *et al*, 2013). D'autres variants du « recombineering » permettant de faire plusieurs modifications de bases uniques à différents endroits dans le génome comme le MAGE (Multiplex Automated Genome Engineering) ou même plusieurs modifications sur une large région du génome comme le CAGE (Conjugative Assembly Genome Engineering) ont également été développés (Wang *et al*, 2009; Isaacs *et al*, 2011). Grâce à ces techniques, il est maintenant possible de recoder des génomes entiers. Cet exploit a d'ailleurs été réalisé chez *E. coli* en 2011, où tous les codons TAG ont été éliminés de son génome (Isaacs *et al*, 2011) pour ensuite être réattribués à un acide aminé non-naturel (Mandell *et al*, 2015).

Ces grandes avancées technologiques dans la manipulation de génomes découlent de notre habileté à lire et à écrire l'ADN. Dès les premières images de la molécule d'ADN par Watson et Crick en 1953 (Watson & Crick, 1953), la communauté scientifique s'est rapidement intéressée à déchiffrer le code génétique. Avec la naissance du séquençage de Sanger en 1977, il nous était maintenant possible de lire les séquences d'ADN, facilitant ainsi sa manipulation (Sanger *et al*, 1977). Par contre, les coûts de séquençage étant faramineux, peu de laboratoires disposaient des ressources

nécessaires à l'utilisation de ces techniques. Suite au séquençage du génome humain en 2001 (Lander *et al*, 2001; Venter *et al*, 2001) et devant son impressionnante facture, le *National Human Genome Research Institute* lança un projet de réduction des coûts de séquençage avec pour but la réduction du coût de séquençage à 1000 \$/génome humain d'ici 10 ans (Schloss, 2008). Cet effort a su stimuler la communauté scientifique à développer une nouvelle génération de technologies de séquençage (NGS) capable de séquencer des génomes entiers à des coûts plusieurs ordres de grandeur plus bas. Les technologies NGS comme le séquençage 454 (Margulies *et al*, 2005), Illumina (Bennett, 2004) et Ion Torrent (Rothberg *et al*, 2011), ont permis la démocratisation des technologies de séquençage d'ADN et le développement des techniques d'analyse globale des processus cellulaires (souvent appelés techniques « omiques »), permettant ainsi l'accumulation rapide de données biologiques (Manzoni *et al*, 2018). Parallèlement, le coût de synthèse d'ADN a également beaucoup diminué, rendant envisageable la synthèse de gros fragment d'ADN et récemment la synthèse de génomes entiers (Hughes & Ellington, 2017).

1.1.2. Les endonucléases programmables : la révolution CRISPR

Bien que plusieurs outils moléculaires aient déjà été développés chez *E. coli*, notre capacité à modifier le génome d'autres organismes, comme les cellules végétales ou animales, demeure limitée. En étudiant les mécanismes de réparation de l'ADN, il fut noté qu'un bris double brin de l'ADN pouvait être réparé de deux façons, soit par recombinaison homologue, ou par jonction des extrémités non homologues (NHEJ) (Featherstone & Jackson, 1999). Puisque le NHEJ est souvent imprécis et peut mener à la délétion ou l'incorporation de nucléotides dans la séquence réparée, la recombinaison homologue est souvent priorisée (Featherstone & Jackson, 1999). C'est devant cette observation que les techniques de modification de l'ADN par utilisation d'endonucléase ont vu le jour. Ces techniques reposent sur le même principe fondamental : l'utilisation d'une endonucléase pour induire une coupure double brin à un endroit souhaité du génome et l'introduction (ou non) d'une matrice d'ADN de réparation. Différentes technologies d'endonucléases, programmables ou non, se sont rapidement succédées. On note par exemple l'utilisation de méganucléases, de ZFN et de TALEN (Genetically Engineered Crops: Experiences and Prospects, 2016). Chacune de ces nouvelles technologies améliorerait successivement la facilité de programmation du site reconnu et la fidélité de l'endonucléase. Par contre, ces techniques ont

grandement diminuées en popularité lorsque les systèmes CRISPR-Cas furent appliqués à l'ingénierie génétique, devant leur facilité d'utilisation et leur haute spécificité.

Les systèmes CRISPR-Cas se divisent en deux grandes classes (I et II) et en 6 types selon leur phylogénie et l'organisation des gènes impliqués (Figure 1.1) (Koonin & Makarova, 2019). Ces complexes protéines-ARN sont utilisés par les bactéries comme système immunitaire adaptatif contre l'ADN exogène. Globalement, les systèmes CRISPR-Cas médient l'immunité en trois grandes étapes, soit l'adaptation, l'expression et l'interférence. À des fins de simplification, le système Cas le plus utilisé dans les technologies d'édition de génome (classe II, type II) sera utilisé en exemple (Mohanraju *et al*, 2016). Dans les systèmes de type II spécifiquement, la phase d'adaptation implique Cas9, Cas1 et Cas2 (avec l'aide additionnelle de Csn2 chez le type II-A (Heler *et al*, 2015) ou de Cas4 chez le type II-B (Li *et al*, 2014)). Cas9 reconnaît une séquence spécifique appelée « protospacer adjacent motif » (PAM) sur l'ADN exogène et recrute Cas1 et Cas2. La discrimination de l'ADN endogène et exogène dans ce procédé dépend de la réplication ainsi que de l'activité des systèmes de défenses innées de la bactérie comme les enzymes de restriction (Levy *et al*, 2015; Dupuis *et al*, 2013). Le complexe Cas1 et Cas2 médie alors l'excision de la séquence adjacente au PAM nommée « protospacer » et l'insert dans une séquence répétée du locus CRISPR via un mécanisme de recombinaison ressemblant aux intégrases/transposases rétrovirales (Nuñez *et al*, 2015). Durant la phase de l'expression, le locus CRISPR est exprimé par l'entremise d'un seul ou de plusieurs promoteurs, permettant la production d'un ARN précurseur contenant les protospacers maintenant appelés « spacer » (Zhang *et al*, 2013). Un autre ARN, appelé ARNtracr est également exprimé et s'hybride aux répétitions se trouvant entre chaque spacer sur le pré-ARNcr (Jinek *et al*, 2012; Deltcheva *et al*, 2011; Charpentier *et al*, 2015). L'ARN double brin formé via l'association entre l'ARNtracr et le pré-ARNcr est reconnu par la RNase III et clivé (Jinek *et al*, 2012; Deltcheva *et al*, 2011; Charpentier *et al*, 2015). Parallèlement, les ARNtracr et ARNcr liés sont également reconnus et liés par Cas9 qui stabilise le complexe d'ARN et permet sa maturation par dégradation des extrémités non protégées par des nucléases endogènes (Gasiunas *et al*, 2012). Le complexe CRISPR-Cas9 est alors mature et prêt à débiter la phase d'interférence. Dans cette dernière phase, Cas9 utilise son partenaire à ARN pour cribler les séquences d'ADN par homologie à son spacer et via la présence d'une séquence spécifique PAM (Garneau *et al*, 2010). Si une séquence d'ADN possédant un PAM montre également une bonne homologie par

rapport au spacer, Cas9 clive l'ADN cible (Garneau *et al*, 2010). Ceci permet alors à la bactérie d'éviter une seconde infection par un même parasite. Par contre, l'efficacité de systèmes CRISPR-Cas dans les bactéries peut être compromise par des systèmes anti-CRISPR (Bondy-Denomy *et al*, 2013). Ces protéines interagissent directement avec les systèmes CRISPR-Cas pour inhiber leur action d'interférence via divers mécanismes (Hwang & Maxwell, 2019). Les mécanismes d'inhibition les plus courants impliquent le blocage de la reconnaissance de l'ADN ciblé ou le blocage de la capacité à cliver l'ADN (Hwang & Maxwell, 2019).

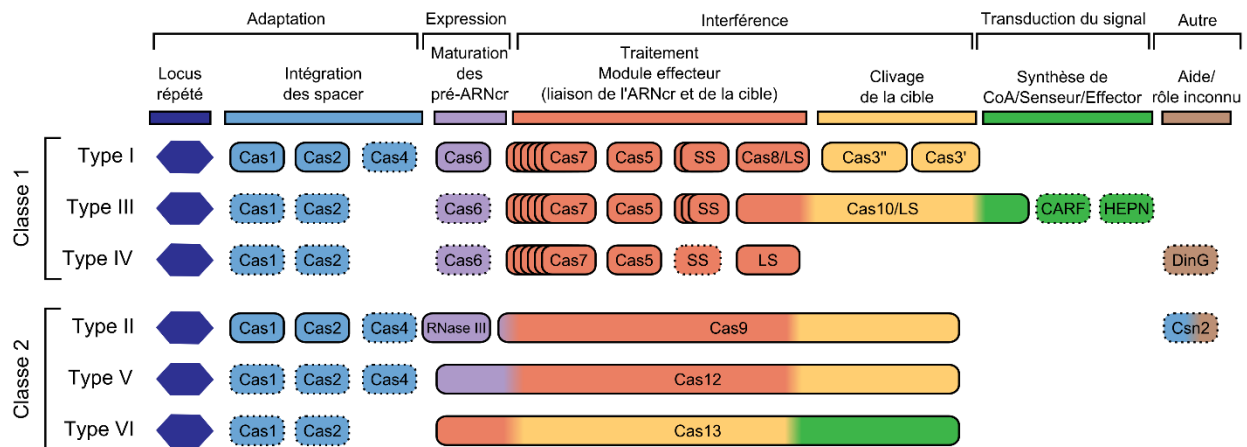


Figure 1.1. Classification des différents types de CRISPR-Cas. Les systèmes CRISPR-Cas partagent des fonctions similaires distribuées sur différentes combinaisons de gènes. Ces systèmes sont d'abord divisés en deux classes selon la présence de plusieurs (classe 1) ou une seule (classe 2) protéine dans leur module effecteur. Ensuite l'organisation des gènes, leur homologie de séquence, les mécanismes de maturation de l'ARNcr et la méthode de détermination du site cible départagent les différents types. Traduit de (Koonin & Makarova, 2019).

L'application des systèmes CRISPR-Cas à l'édition génomique a révolutionné la biologie, valant à ses découvreurs un prix Nobel de chimie cette année 2020 (Jinek *et al*, 2012). Contrairement aux technologies précédentes (méganuléases, ZFN, TALEN) qui requièrent de l'ingénierie au niveau des domaines protéiques de reconnaissance de l'ADN afin de changer leur séquence cible, les CRISPR-Cas permettent une programmation facile via la modification de la séquence du spacer

de l'ARN partenaire (Yu *et al*, 2016). Cette facilité d'utilisation jumelée à sa fidélité et sa facilité d'adaptation dans un large éventail d'espèces ont contribué largement à la démocratisation de son utilisation en génie génétique. L'utilisation de CRISPR-Cas9 a d'ailleurs été facilitée davantage par la fusion de son ARNtracr et de son ARNcr en un seul ARN synthétique, l'ARN guide (ARNg) (Mali *et al*, 2013). En plus de son fort potentiel en thérapie génique, CRISPR-Cas a également des applications très intéressantes dans le contrôle des populations bactériennes et comme antibiotique de précision (Bikard *et al*, 2014; Citorik *et al*, 2014; Yosef *et al*, 2015).

1.2 L'application de la biologie synthétique au domaine médical

Ces nouveaux outils, combinés aux avancés dans les technologies d'écriture et de lecture d'ADN, ont ensemble permis l'émergence de la biologie synthétique (ou biologie de synthèse), une discipline biologique alliant des approches de biologie et d'ingénierie afin de concevoir, construire et tester des systèmes biologiques (Cameron *et al*, 2014). De tels systèmes sont basés sur des circuits géniques qui sont eux-mêmes composés de gènes (Cameron *et al*, 2014). La biologie synthétique se distingue en trois points de la biotechnologie et du génie génétique, soit par l'utilisation d'approches propres à l'ingénierie, par la construction de systèmes biologiques modulaires plus complexes souvent à large échelle et par le recours à de la synthèse d'ADN permettant le clonage de séquences d'ADN non retrouvées dans la nature (Purnick & Weiss, 2009). En biologie synthétique, on fait souvent référence au cycle de conception, construction et test lorsqu'on parle du développement d'un circuit génétique (Figure 1.2). À défaut d'avoir accès à des organismes entièrement compris et prédictibles, les biologistes synthétiques devront souvent faire plus d'un cycle d'optimisation afin d'obtenir un circuit génétique fonctionnel. Chacun des cycles d'optimisation nous en apprend davantage sur le système biologique et aide à l'améliorer.

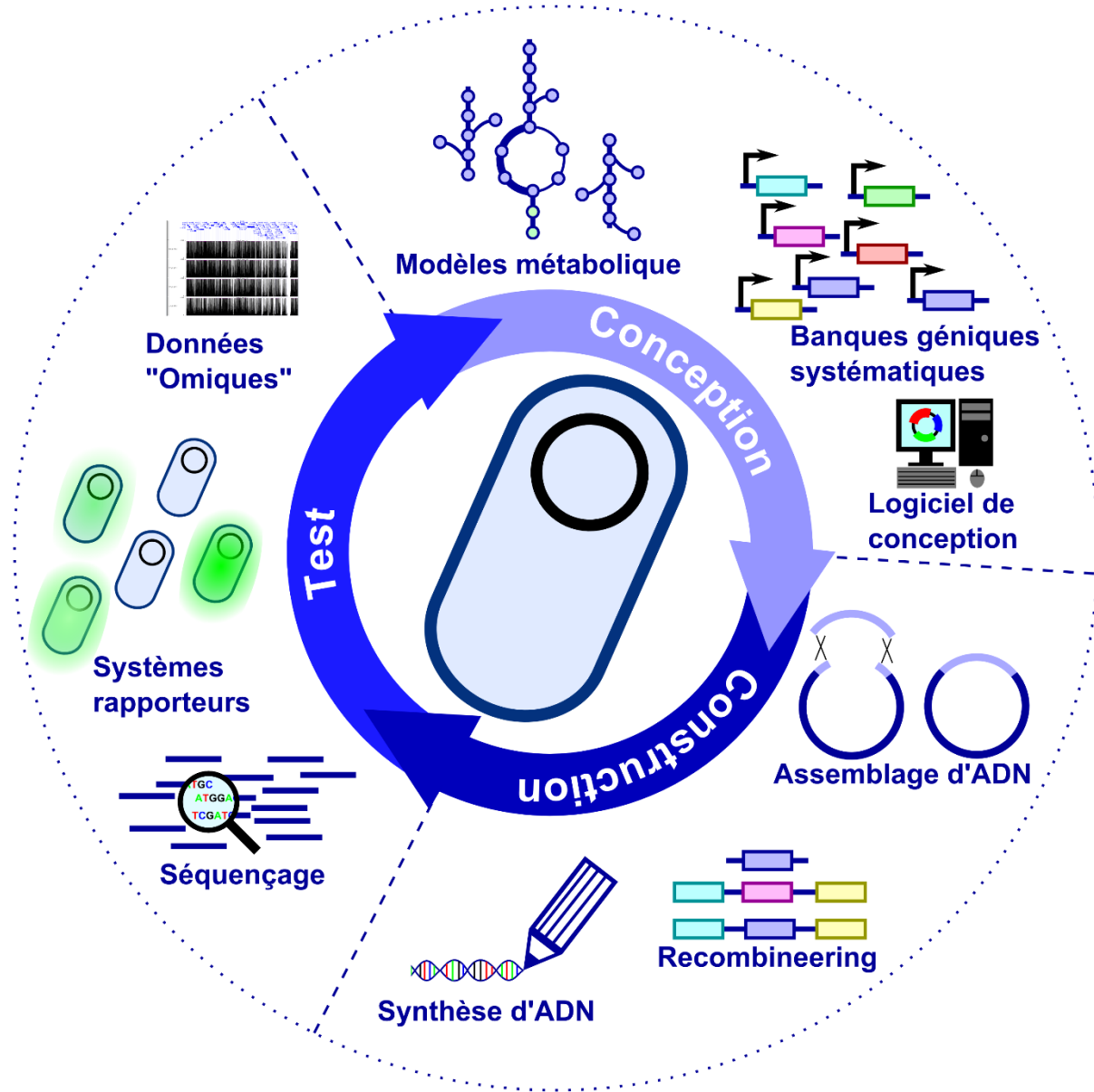


Figure 1.2. Cycle d'optimisation typique de la biologie synthétique. Des outils sont à la disposition des biologistes synthétiques pour leur permettre de franchir chaque étape du cycle. Afin de prévoir l'impact des modifications génétiques sur le métabolisme de notre châssis cellulaire, des modèles informatiques simulant le métabolisme et même l'expression des gènes chez *E. coli* ont été développés (Lloyd *et al.*, 2018). De la même façon, pour faciliter l'assemblage de circuits géniques, des banques de gènes (par exemple le *iGEM parts registry* et *Addgene*) et de promoteurs caractérisés (Mutalik *et al.*, 2013) sont disponibles à la communauté. Le tout peut

ensuite être conçu dans des logiciels tels que ApE (A plasmid Editor) (<https://jorgensen.biology.utah.edu/wayned/apE/>) ou *SnapGene software* (de *Insightful Science*; disponible à snapgene.com). La construction de ces circuits génétiques peut alors être entreprise grâce aux outils moléculaires décrits ci-haut. Ces circuits peuvent ensuite être testés de différentes manières dépendamment de l'objectif du circuit génétique.

Comme lors de la naissance de la biotechnologie avec Cohen et Boyer, les scientifiques pratiquant la biologie synthétique appliquent déjà ces principes d'optimisation à des systèmes biologiques trouvant une multitude d'applications dans le domaine médical dont quelques-unes sont brièvement discutées ci-bas.

1.2.1. La biologie synthétique appliquée à la production de molécules pharmaceutiques

Tout comme la biotechnologie, la biologie synthétique a su trouver sa première application dans la production de molécules naturelles ayant des propriétés pharmaceutiques intéressantes (Martin *et al*, 2003). Par contre, contrairement à la biotechnologie qui fut majoritairement utilisée pour la production de protéines thérapeutiques ou de métabolite simples, la biologie synthétique permet la synthèse de nouvelles molécules via l'assemblage et l'optimisation de voies métaboliques complètes (Martin *et al*, 2003). En effet, l'ajout d'enzymes hétérologues et la délétion/l'inhibition de voies métaboliques endogènes peuvent rediriger le flux métabolique vers la production de la molécule d'intérêt (Figure 1.3.A) (Wohlleben *et al*, 2012). Cette approche a d'ailleurs été démontrée pour la production d'artémisine, une molécule thérapeutique contre la malaria, dans *Saccharomyces cerevisiae* par l'importation des gènes responsables de sa biosynthèse chez *Artemisia annua* (Ro *et al*, 2006). La synthèse de nouvelles voies métaboliques par ajout d'enzymes exogènes a été grandement facilitée par le séquençage de génomes/métagénomes et l'utilisation d'outils bio-informatiques comme antiSMASH (Medema *et al*, 2011), nous donnant ainsi accès à une large banque d'enzymes catalysant des réactions spécifiques. Bien que nous disposions d'outils de découverte d'enzymes aux propriétés intéressantes, un défi considérable est toutefois toujours présent dans l'expression de ces gènes chez une espèce modèle. La biologie synthétique promet de surmonter ce problème grâce à sa modularité et son contrôle d'expression

bien caractérisé, par exemple, via des facteurs de transcription synthétiques (Leavitt & Alper, 2015). Puisque l'assemblage de plusieurs gènes afin de créer une nouvelle voie métabolique demeure complexe, des alternatives plus simples ont également été développées. Plusieurs groupes se sont tournés vers l'expression d'une seule enzyme par souche bactérienne catabolisant la transformation d'un précurseur en composés bioactifs (Figure 1.3.B) (Trosset & Carbonell, 2015). Cette approche est plus simple à mettre au point, mais son efficacité dépend des taux d'importation du précurseur et d'exportation du composé bioactif qui exploitent souvent des transporteurs naturels non adaptés. Bien que des améliorations soient nécessaires dans les taux de production et dans l'optimisation des voies de synthèse de métabolites, la biologie synthétique présente de fortes promesses dans le développement de molécules à intérêt pharmaceutique.

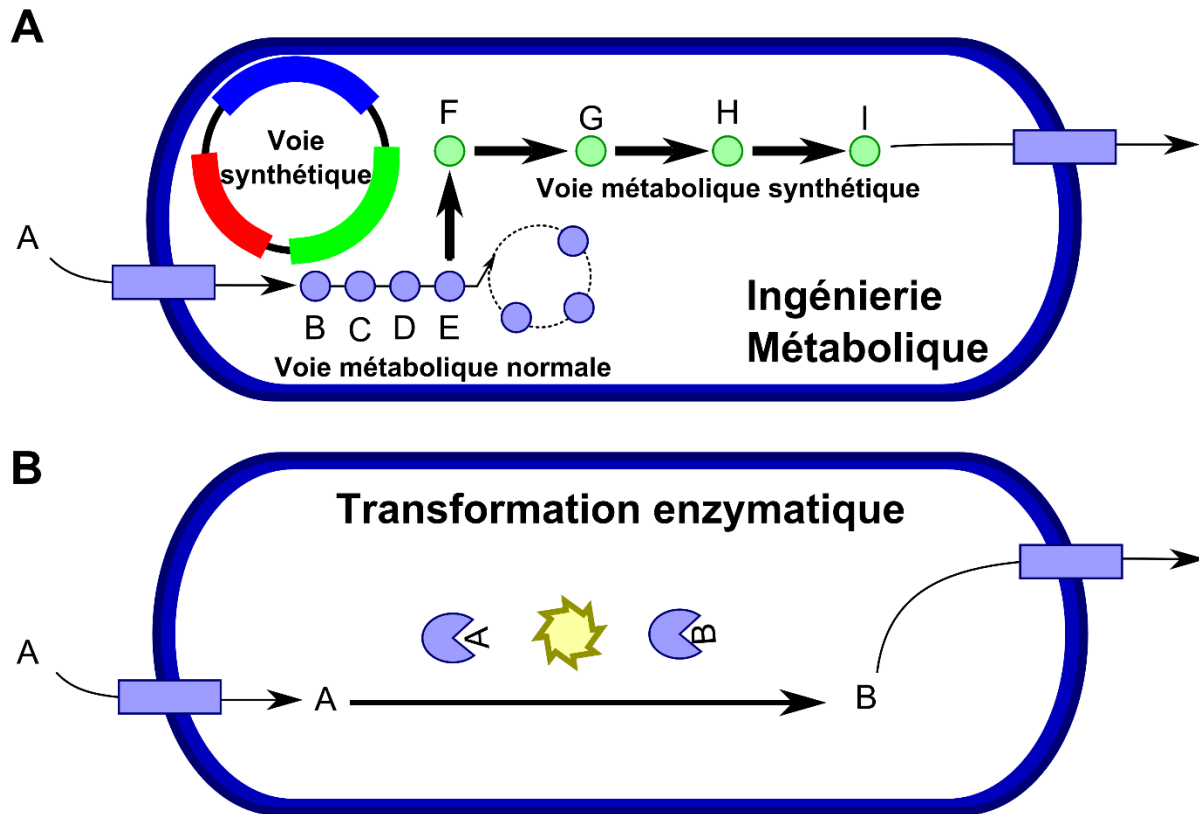


Figure 1.3. Production de molécules bioactives chez des organismes modifiés par approche de biologie synthétique. Les cellules peuvent être modifiées de sorte que le métabolisme cellulaire produise un métabolite secondaire bioactif (A) ou simplement pour modifier une molécule afin de changer sa structure chimique (B). Tiré de (Trosset & Carbonell, 2015).

1.2.2. La biologie synthétique comme outil diagnostique

La biologie synthétique comme outil diagnostique exploite toujours le même principe de base, c'est-à-dire l'utilisation de unités biologiques et/ou de circuits génétiques pour la détection de composés/stimuli émettant un signal mesurable comme de la lumière, de la fluorescence ou même un courant électrique (Slomovic *et al*, 2015). Le domaine s'est d'abord intéressé à utiliser la biologie synthétique afin de développer des machines hybrides, c'est-à-dire des détecteurs spécifiques liés à des systèmes biologiques qui réagissent à la présence de certains composés. Un exemple intéressant de machine hybride est le développement de biosenseurs utilisant des cyanobactéries pour détecter, via leur activité photosynthétique, la présence d'herbicide dans l'eau par la production d'un signal électrique (Rawson *et al*, 1989). Bien que cette approche ne soit pas de la biologie synthétique proprement dite, elle a tout de même été pionnière dans l'utilisation de bactéries à des fins diagnostiques. Une autre étude particulièrement intéressante a utilisé des petits ARN interférant (ARNsi) afin de détecter des lignées de cellules cancéreuses spécifiques ou de mener à leur apoptose (Figure 1.4.A) (Xie *et al*, 2011). Ce genre de technologie pourrait, à terme, être adapté comme outil de typage des cellules tumorales à partir d'une biopsie et pourrait même devenir un outil thérapeutique.

Beaucoup de senseurs sont basés sur les protéines liant le maltose et exploitent la propriété de la protéine native à se replier sur elle-même. En effet, en liant la protéine à deux domaines fluorescents, on peut détecter en temps réel le rapprochement des deux domaines fluorescents de façon similaire au FRET (Neumann & Neumann-Staubitz, 2010). On peut également modifier le site de liaison au ligand par évolution dirigée pour permettre la détection de métabolites secondaires variés (Dwyer & Hellinga, 2004).

On avance également l'utilité de circuits génétiques permettant le suivi d'un état pathologique cellulaire (déstabilisation de la membrane, dommage à l'ADN, stress oxydatif, etc) dans le criblage de bibliothèques de composés chimiques (Trosset & Carbonell, 2015). Un exemple simple d'un tel système serait d'utiliser des bactéries disposant d'un gène rapporteur (p.ex. *sfGFP*) inductible via la réponse SOS. Ce circuit pourrait alors permettre la détection rapide d'effets mutagènes dans une

librairie de molécule d'intérêt pharmacologique. De tels circuits pourraient à terme être adaptés à la détection de différents signaux pour permettre le criblage de grandes bibliothèques de composés pharmacologiques et offrir une détection quasi instantanée des molécules bioactives.

De façon inédite, la biologie synthétique permet le développement d'outils diagnostiques qui sont utilisables *in situ* afin de détecter des anomalies. En utilisant des organismes programmés pour la détection et l'enregistrement de leur contact avec un facteur environnemental précis grâce à des circuits synthétiques mémoires, un patient pourra bientôt bénéficier de nouveaux outils de suivi de son état de santé en continu. Ces circuits sont basés sur l'utilisation de recombinaisons inducibles qui réarrangent l'ADN pour produire un certain phénotype stable en réponse à une exposition à un certain composé (Figure 1.4.B) (Bonnet *et al*, 2012). Par exemple, de tels systèmes ont été utilisés afin de détecter la présence d'anhydrotétracycline (aTc) dans le microbiote entérique murin (Kotula *et al*, 2014). Les bactéries senseurs pouvaient garder leur exposition en mémoire pour au moins 15 jours dans le microbiote murin et pouvaient être détectées directement dans les fèces, permettant ainsi un diagnostic très peu invasif. De plus, en exploitant la capacité de certaines bactéries à proliférer dans les tumeurs dues à des avantages environnementaux (p.ex. basse concentration d'oxygène, baisse de l'activité du système immunitaire), on peut développer des bactéries capables de détecter les tumeurs et d'émettre un signal mesurable (Yu *et al*, 2008), par exemple, en bioluminescence (Figure 1.4.C) (Cronin *et al*, 2012).

Comme les processus cellulaires naturels peuvent interférer avec le biosenseur, plusieurs groupes développent des systèmes d'expression sans cellules qui trouvent des applications diagnostiques intéressantes. Ces systèmes peuvent être imbibés dans du papier et ainsi permettre la détection de composés comme le ferait un papier pH (Pardee *et al*, 2014). Un exemple d'application de cette technologie est l'utilisation de systèmes d'expression contrôlés par « toehold switch » pour l'expression d'un gène rapporteur produisant un pigment visible à l'œil nu qui a été employé, entre autres, dans la détection de 24 souches d'Ebola via la détection d'ARN spécifique à ces souches (Pardee *et al*, 2014). Par ailleurs, les senseurs à ARN de type « toehold switch » peuvent en théorie permettre la détection de virtuellement n'importe quelle molécule d'ARN par complémentarité de séquence (Green *et al*, 2014). Les biosenseurs sur support en papier sont d'ailleurs le cœur de

métier de plusieurs compagnies dont FREDsense Technologies, une compagnie canadienne qui développe des biosenseurs pour l'analyse de contaminant dans l'eau comme l'arsenic (<https://fredsense.com/>). Un des grands défis des systèmes de détection de composés utilisant de l'expression sans cellule réside dans la variété de signaux pouvant être détectés par des facteurs de transcription pour mener à l'expression du gène rapporteur. Pour adresser cette limitation, les groupes des professeurs Faulon et Bonnet ont proposé une stratégie fort intéressante (Voyvodic *et al*, 2019). En utilisant des enzymes de conversion, par exemple CocE qui converti la cocaïne en acide benzoïque, il devient possible d'utiliser le facteur de transcription BenR répondant à la présence de l'acide benzoïque pour détecter la cocaïne dans des milieux complexes. Il sera donc intéressant de voir la variété de biosenseurs qui pourront être développés dans les prochaines années en utilisant cette stratégie.

Les systèmes de diagnostiques utilisant CRISPR afin de détecter des molécules d'ADN spécifiques est un domaine effervescent (Chen *et al*, 2018; Gootenberg *et al*, 2017, 2018). Ces systèmes utilisent la grande spécificité de CRISPR-Cas à cliver des séquences d'ADN programmables afin de générer un signal mesurable. Par exemple, le système « DETECTR » a permis la détection de souches différentes du virus du papillome humain via l'amplification isothermique d'une des sections variables avec une amorce fluorescente et une amorce « quencher » (Chen *et al*, 2018). La détection de la séquence correspondante à une souche particulière par Cas12a permet le clivage de la molécule d'ADN et l'éloignement du groupement fluorescent de son « quencher » ce qui résulte en une augmentation du signal fluorescent. Un autre système particulièrement intéressant appelé « SHERLOCK » permet la détection à la fois d'ADN et d'ARN d'intérêt (Gootenberg *et al*, 2017, 2018). Dans cette méthode, les nucléotides de l'échantillon sont d'abord amplifiés isothermique avec des amorces introduisant un promoteur T7 de sorte de créer une librairie de fragments d'ADN pouvant être transcrit. Ensuite la T7 polymérase transcrit l'ADN en ARN et Cas13a est utilisé pour reconnaître des séquences d'ARN spécifique. Suite à la reconnaissance de sa cible, Cas13a active une fonction RNase non-spécifique qui clive une molécule d'ARN rapportrice et libère un fluorophore de son « quencher ». Le domaine des outils diagnostique CRISPR est vaste, renferme bon nombre de systèmes très intéressants et sera à suivre de près dans les prochaines années.

L'utilisation de bactériophages à des fins diagnostiques a également suscité beaucoup d'intérêt au niveau de la détection de pathogènes spécifiques. En effet, les bactériophages ont une tendance naturelle à infecter un spectre d'hôte limité, ce qui peut être utile pour la détection et le typage de pathogènes. Afin de transformer un bactériophage en outil diagnostique, on relocalise sa séquence d'encapsidation sur un phagemide codant pour une protéine rapporteuse (Schofield *et al*, 2012). Par exemple, un groupe a démontré qu'il était possible de détecter *Listeria* spp. en l'infectant avec un phage A511 exprimant une luciférase lorsqu'il atteint le cytoplasme de sa cible (Loessner *et al*, 1996).

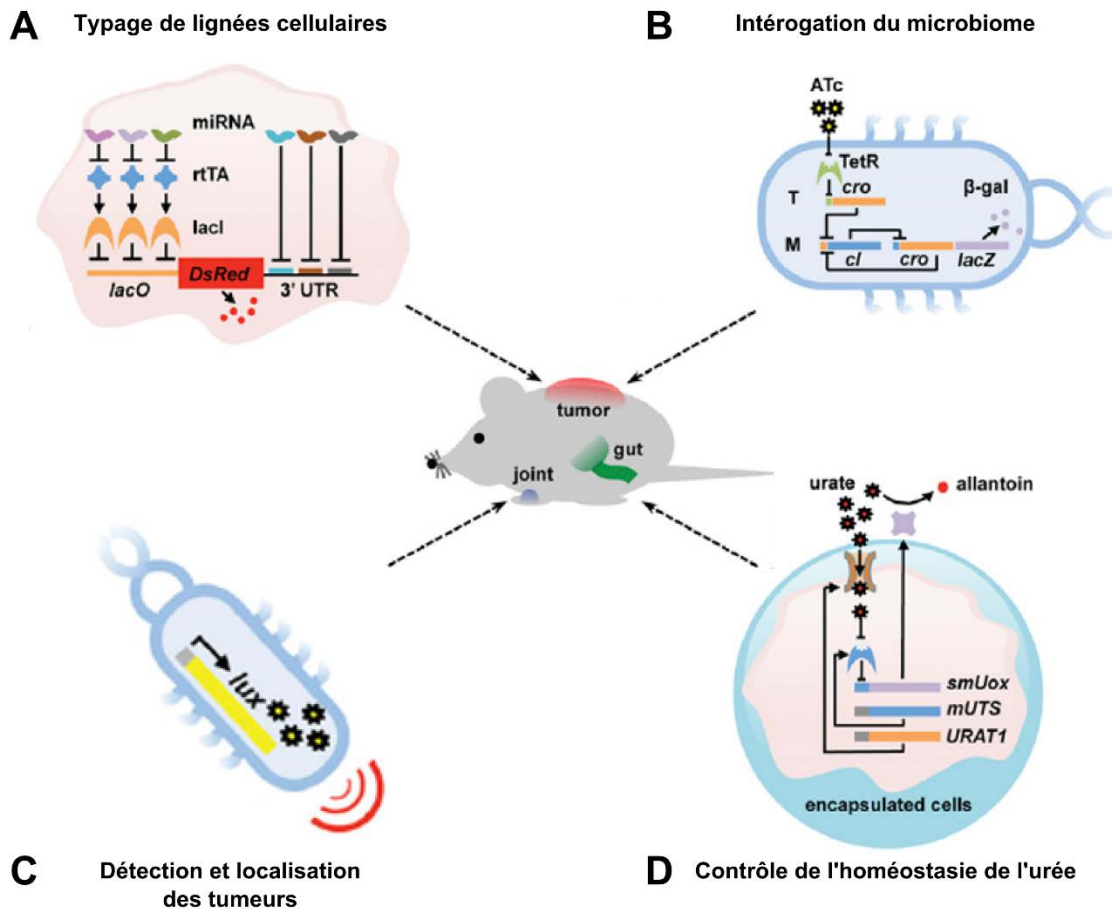


Figure 1.4. Exemples de modules de biologie synthétique comme outil diagnostique. (A) Le *Caspase Tracker*, une technologie qui détecte le clivage de Gal4 par les caspases des cellules apoptotiques et qui est alors transloqué au noyau pour médier l'expression de RFP (rapporteur pour l'état apoptotique), mais également de FLP, une flipase qui recombine un locus de sorte que

GFP se retrouve exprimé. Ceci garde en mémoire l'état apoptotique de la cellule même chez sa progéniture. (B) Exemple de bactérie senseur de cancer exprimant la luciférase afin de signaler sa prolifération à l'intérieur de la tumeur. (C) Circuit génique mesurant les niveaux d'expression de trois miRNA dans le but de détecter son type cellulaire. (D) Biosenseur des niveaux d'urée sanguin rétablissant l'homéostasie en exprimant une uréase. (E) Système de détection d'anomalies au niveau du microbiome intestinal induisant la mémoire cellulaire lorsqu'il y a détection du composé cible (tiré de (Slomovic *et al*, 2015)).

1.2.3. La biologie synthétique appliquée à la médecine régénérative

Bien qu'encore à une étape conceptuelle, la biologie synthétique peut trouver des applications intéressantes en médecine régénérative. En effet, il est actuellement souvent impossible pour la médecine régénérative de créer des tissus ayant des formes non naturelles pour le traitement de patients ayant un corps grandement altéré par des accidents ou maladies congénitales. La biologie synthétique promet de répondre à cette problématique (Cachat & Davies, 2011). Par exemple, l'utilisation de circuits logiques Boolean (AND, OR, NOR, etc.) (Kramer *et al*, 2004a) et de circuits mémoires (Kramer *et al*, 2004b) (plus tard testés chez la souris (Greber *et al*, 2008)) pourraient être intéressants afin de contrôler précisément le relâchement de molécules spécifiques (par exemple un facteur de croissance) qui contrôlent précisément la prolifération et la spécialisation cellulaire. Une équipe a, par exemple, utilisé un système d'induction à la lumière afin de déclencher la formation de lamellipode (Levskaya *et al*, 2009), une technologie intéressante dans la production de tissus par exposition à la lumière. De plus, la biologie synthétique offre la possibilité de directement programmer des cellules chez les individus pour remplacer ou bonifier certaines fonctions existantes. Par exemple, un circuit capable de sentir les niveaux d'urée sanguins et de les réguler a déjà été développé (Figure 1.4.D) (Kemmer *et al*, 2010). Aussi, l'introduction de circuits géniques activables par des petites molécules chimiques permettrait l'expression de facteurs impliqués dans la guérison au site de la blessure (Gitzinger *et al*, 2009). Une autre application possible de la biologie synthétique dans la médecine régénérative se trouve dans la production d'hydrogels intelligents, lesquels ont déjà démontré de l'importance dans l'organisation tridimensionnelle des tissus générés *in vitro* (Carpenedo *et al*, 2009). Un exemple d'une telle technologie repose sur des hydrogels d'ADN-acrylamide. En effet, l'ADN peut être

arrangé en matrice formant un hydrogel. En utilisant un complexe CRISPR-Cas, il est possible de modifier la structure de cet hydrogel en ciblant des séquences précises contenues dans certains liens, permettant ainsi, par exemple, le relâchement calculé de cellules en culture (English *et al*, 2019). Somme toute, l'application de la biologie synthétique en médecine régénérative est un domaine nouveau qui demandera des efforts considérables, mais renferme de grandes promesses.

1.2.4. La biologie synthétique comme outil prophylactique ou thérapeutique

Les traitements prophylactiques exploitant la biologie synthétique visent surtout la production de souches vaccins. Par exemple, certains groupes ont développé des vaccins basés sur des pathogènes atténués entéro-invasifs comme *Salmonella typhimurium*, *Listeria monocytogenes* et *Shigella flexneri* modifiés afin de leur permettre de transférer les cellules humaines avec un vaccin à ADN (Schoen *et al*, 2004). Par contre, l'utilisation de bactéries pathogènes atténuées comporte son lot de risques étant donné la possibilité de réversion du pathogène vers son état virulent. Une stratégie alternative se base sur *Lactococcus lactis* rendu envahissant par expression de FnBPA+ (*fibronectin-binding protein A*) (Vaudoaux *et al*, 2001) provenant de *Staphylococcus aureus*. Cette souche peut être utilisée comme vecteur de transfection pour un plasmide codant un antigène spécifique (p. ex. ESAT-6 de *Mycobacterium tuberculosis*) (Pereira *et al*, 2014).

Le dernier domaine appliqué de la biologie synthétique au domaine médical vise la production d'organismes ayant une action thérapeutique directe (Charbonneau *et al*, 2020; Pinero-Lambea *et al*, 2015). En effet, en appliquant les principes de la biologie synthétique, nous pouvons maintenant programmer des organismes afin de leur faire exercer diverses activités thérapeutiques. Ces organismes seraient alors capables de produire le composé bioactif ou l'activité thérapeutique directement sur place et permettraient un contrôle temporel de son relâchement, ce qui nous donnerait accès à des courbes pharmacodynamiques inédites (Charbonneau *et al*, 2020). Les efforts de développement d'organismes médicaments se concentrent surtout dans trois champs d'applications principaux :

Le premier champ d'application vise à implémenter des organismes médicaments dans le microbiome du patient afin d'y avoir un effet thérapeutique durable. En effet, un des premiers exemples d'organismes médicament était un *L. lactis* programmé pour relâcher de l'IL-17 lorsqu'introduit dans le système digestif comme traitement potentiel contre les maladies inflammatoires de l'intestin (Braat *et al*, 2006; Steidler *et al*, 2003). De tels organismes ont également beaucoup de potentiel pour le contrôle de l'homéostasie. Par exemple, afin de s'attaquer au diabète de type II, une bactérie commensale de l'intestin a été programmée pour relâcher GLP-1, un peptide stimulant la sécrétion de l'insuline en répondant au niveau de glucose sanguin (Duan *et al*, 2008). Aussi, l'implémentation de bactéries encodant des fonctions métaboliques précises peut être envisagée comme traitement pour des désordres métaboliques. Un organisme médicament qui dégrade la phénylalanine est d'ailleurs présentement en essai clinique pour le traitement de la phénylcétonurie, soit l'accumulation de la phénylalanine dans le sang (Isabella *et al*, 2018).

Un deuxième champ d'application prometteur des organismes médicaments vise le traitement de cancers. En effet, il fut démontré pour plusieurs espèces bactériennes, dont *E. coli*, colonisent préférentiellement les tumeurs solides lorsqu'administrées de façon intraveineuse dans les souris (Zhou *et al*, 2018). Cette colonisation préférentielle est due à la nature hypoxique et peu immunogène des tissus tumoraux qui protège les bactéries du système immunitaire (Zhou *et al*, 2018). Partant de ce principe, plusieurs modules de biologie synthétique ont été développés pour traiter des cancers (Forbes, 2010). Un premier exemple élégant de stratégie antitumorale est l'expression d'inhibiteur des points de contrôle immunitaires, comme l'utilisation d'un « nanobody » se liant à CD47, un récepteur cellulaire souvent surexprimé chez les cellules tumorales et qui inhibe la réponse immunitaire (Chowdhury *et al*, 2019). Dans ce système, la bactérie est confinée par un système de « quorum sensing » qui provoque la lyse de la cellule ainsi que le relâchement du nanobody. Le blocage du CD47 mène alors à l'élimination de la cellule tumorale par le système immunitaire. Un autre exemple particulièrement intéressant est une bactérie sécrétant de l'AMP cyclique (Leventhal *et al*, 2020). L'AMP cyclique est reconnue par les cellules présentatrices d'antigènes lors de la phagocytose des bactéries au site de la tumeur, ce qui déclenche la voie immunitaire STING et mène à l'élimination de la tumeur.

Le dernier champ d'application d'importance vise l'élimination de bactéries problématiques à l'aide de probiotiques fonctionnalisés. Par exemple, le probiotique *E. coli* Nissle 1917 (EcN) a été modifié afin d'exprimer les molécules de quorum utilisées par *Vibrio cholerae*, améliorant ainsi significativement le taux de survie de souris infectées en inhibant la croissance de ce pathogène (Duan & March, 2010). D'autres groupes ont opté pour des circuits synthétiques « recherche et tue », où les probiotiques étaient modifiés pour agiter davantage leur flagelle en réponse à une molécule produite par *Pseudomonas aeruginosa* avant de sécréter une DNase permettant de dégrader les biofilms ainsi que la microcine S pour tuer *P. aeruginosa* (Hwang *et al*, 2014). Plusieurs équipes de recherche ont plutôt porté leur intérêt sur la sécrétion de phages modifiés permettant l'élimination de bactéries spécifiques. Les exemples les plus marquants sont l'utilisation de phages pour envoyer des signaux anti-quorum (Pei & Lamas-Samanamud, 2014) et pour dissoudre des biofilms (Lu & Collins, 2007). Une stratégie particulièrement intéressante est de livrer des systèmes CRISPR-Cas9 pour l'élimination spécifique de pathogènes (Citorik *et al*, 2014; Bikard *et al*, 2014; Yosef *et al*, 2015). Il fut d'ailleurs démontré chez *E. coli* qu'une coupure médiée par CRISPR-Cas9 dans son chromosome peut mener à une perte d'intégrité du génome et à la mort de la cellule hôte (Cui & Bikard, 2016). La livraison de CRISPR-Cas9 à des bactéries cibles est donc une méthode d'élimination de bactéries précises qui discrimine les souches à éliminer selon leur contenu en gènes. Différents groupes ont développé des bactériophages pouvant livrer CRISPR-Cas9. Par exemple, l'équipe du Dr Lu a produit des phagemides basés sur le bactériophage M13 afin d'éliminer spécifiquement les souches portant le gène de résistance au β -lactame *bla* chez *E. coli* (Citorik *et al*, 2014). Cette équipe a d'ailleurs démontré l'efficacité de son système dans un modèle d'infection chez *Galleria mellonella*. L'équipe de Marrafini (Bikard *et al*, 2014) a quant à elle utilisé des phagemides basés sur le bactériophage ϕ NM1 pour infecter et éliminer spécifiquement une souche de *S. aureus* résistante à la méthicilline. Bikard et collaborateurs ont aussi effectué des tests chez la souris dans un modèle d'infection sous-cutané où ils ont observé une diminution de moitié de la densité de la souche ciblée par rapport à une souche non ciblée. Finalement, l'équipe de Quimron (Yosef *et al*, 2015) a démontré qu'il était possible d'utiliser des phagemides portant CRISPR-Cas9 afin de contre-sélectionner les résistances aux antibiotiques, déplaçant ainsi les pressions sélectives vers la sensibilité aux antibiotiques. Bien que les bactériophages soient un véhicule efficace pour la livraison de CRISPR-Cas9, d'autres groupes ont porté leur intérêt sur les plasmides conjugatifs (López-Igual *et al*, 2019; Ronda *et al*, 2019).

En effet, les plasmides conjugatifs pour la livraison de matériel génétique comportent certains avantages, comme un spectre d'hôte souvent plus large, leur permettant d'adresser un plus large éventail de bactéries retrouvées dans le microbiote humain (Suzuki *et al*, 2010). La livraison par plasmide conjugatif a été décrite pour plusieurs circuits génétique. Entre autres, elle fut employée pour livrer des complexes toxine-intéine activés seulement chez *V. cholerae* pour son élimination spécifique (López-Igual *et al*, 2019). D'autres groupes visent plutôt des plasmides portant des gènes de résistance aux antibiotiques. À l'aide d'un module d'interférence à la réplication (composé d'une origine de réplication et d'un gène encodant l'antitoxine), l'équipe du Pr Thomas a démontré que de conjuguer leur vecteur vers une souche portant un plasmide de résistance de la famille ciblée permettait de perdre ce plasmide de résistance (Kamruzzaman *et al*, 2017; Lazdins *et al*, 2020). Plutôt que de retirer certains gènes ou bactéries d'une population, d'autres ont proposé d'utiliser les plasmides conjugatifs afin d'ajouter des gènes chez des bactéries endogènes du microbiome (Ronda *et al*, 2019). Le transfert d'ADN par conjugaison bactérienne est donc une méthode qui pourrait permettre le développement d'outils de modification de populations bactérienne dans des environnements complexes tel le microbiome entérique. L'étude du microbiome suscite d'ailleurs beaucoup d'intérêt grâce à plusieurs recherches pionnières qui ont su démontrer son rôle dans une panoplie de fonctions chez l'humain (Lloyd-Price *et al*, 2016; Turnbaugh *et al*, 2007; Laukens *et al*, 2015). Pouvoir altérer précisément la composition du microbiote offre donc un grand potentiel thérapeutique.

1.3 Le microbiome

Le microbiote peut être défini comme l'ensemble des micro-organismes vivants dans un certain environnement. Le microbiome quant à lui consiste en l'ensemble de tous les génomes de ces micro-organismes (Turnbaugh *et al*, 2007). En raison de son importance médicale, le microbiome humain et celui de certains mammifères ont été particulièrement bien étudiés (Huttenhower *et al*, 2012).

L'étude du microbiome a été popularisée au début des années 2000 avec la démocratisation des méthodes « indépendantes de la culture » pour la caractérisation de communautés microbiennes

(« *Whole Genome Shotgun* », séquençage métagénomique de l'ADNr 16S) (Weinstock, 2012). Dans les mêmes années, avec l'assemblage du premier génome humain (Venter *et al*, 2001), la communauté scientifique put constater que le génome humain encode beaucoup moins d'information qu'initialement anticipé. Comment toutes les fonctions importantes à l'humain peuvent-elles être encodées dans ce peu de données ? Cette question fit naître l'idée de l'humain en tant que supra-organisme, c'est-à-dire l'idée que le génome humain est supplémenté par le contenu génique des bactéries qui vivent dans son microbiome (Gill *et al*, 2006).

Tableau 1.1 Diversité des études sur le microbiome (tiré de (Lloyd-Price *et al*, 2016))

Région du corps	Publications	
	Tout	2011-2016
Colon/intestin	17,546	10,707
Bouche	4843	2089
Méat urinaire/Vagin/Pénis	1477	706
Peau	1372	754
Poumons	764	524
Placenta	702	426
Yeux	152	82

Depuis lors, la composition du microbiome a fasciné plusieurs groupes de recherche (Tableau 1.1) qui ont caractérisé différents environnements du corps humain, dont la bouche (Takeshita *et al*, 2016), le nez (Bassis *et al*, 2014), l'œsophage (Pei *et al*, 2004), la peau (Gao *et al*, 2007), l'estomac (Bik *et al*, 2006), l'intestin (Eckburg *et al*, 2005), le vagin (Hyman *et al*, 2005), etc. Des compositions variables de micro-organismes ont été retrouvées dans ces différents environnements. Par contre, le nombre d'individus échantillonnés pour ces études restait très faible et les données médicales concernant ces individus quasi indisponibles. Un consortium établi en 2007, « *The Human Microbiome Project* », a rapidement entrepris la caractérisation du microbiome humain sain, cette fois avec un nombre de sujets beaucoup plus important et des standards d'échantillonnages précis (Turnbaugh *et al*, 2007). La variabilité du microbiome entre

les individus étant déjà bien connue, ils émirent l'hypothèse que le microbiome d'une personne saine devrait avoir une constance au niveau des fonctions de gènes présents (Turnbaugh *et al*, 2007). En d'autres mots, des gènes de rôles similaires devraient prédominer dans le microbiome de chaque individu sain. En ce sens, une différence de composition des espèces du microbiome serait expliquée par le fait que différents organismes peuvent contenir des gènes encodant la même fonction. Leur hypothèse fut enfin confirmée lorsqu'en 2012, le consortium «*The Human Microbiome Project*» publia les résultats de l'analyse de toutes les grandes régions du corps (excepté l'estomac et l'œsophage) chez près de 300 personnes saines (Huttenhower *et al*, 2012). En parallèle, le consortium «*MetaHIT*» (Dusko Ehrlich, 2010), s'est concentré sur la caractérisation du microbiome entérique chez 154 personnes. Le consortium *MetaHIT* est arrivé à la même conclusion que l'équipe de «*The Human Microbiome Project*» en plus de contribuer à cataloguer les fonctions du microbiome (Qin *et al*, 2010). Ensemble, ces études nous ont permis d'avancer grandement dans la caractérisation du microbiome sain (Figure 1.5). La constance des gènes retrouvés dans le microbiote malgré la variabilité de la composition en organismes entre individus peut être partiellement expliquée par la théorie des îles en écologie. En effet, tout comme une île, les premières espèces à être introduites sur le corps humain ont plus de chances de s'établir que celles qui arriveraient plus tard. Ceci est dû à des effets de compétition avec les résidents du microbiote pour une même niche. La dépendance maternelle du microbiote chez la souris (Ley *et al*, 2005) supporte bien ce modèle.

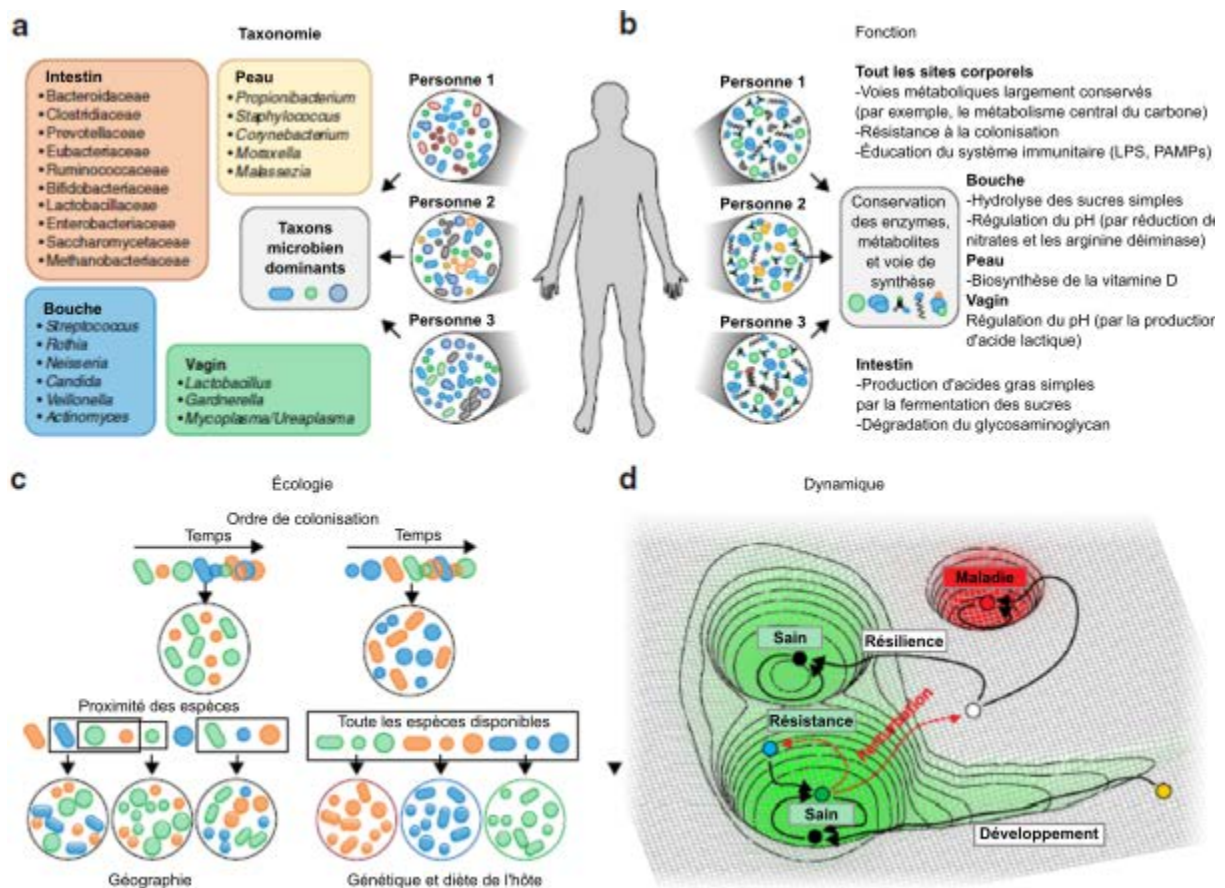


Figure 1.5 Définition possible d'un microbiome sain. (A) Familles ou genres bactériens particulièrement conservés et abondants chez la plupart des individus sains à différents sites corporels. (B) Principales fonctions conservées du microbiome sain à différents endroits du corps humain. (C) Hypothèses pouvant expliquer la variance entre les microbiomes de différents individus. Les bactéries sont introduites dans un ordre aléatoire qui peut être important pour leur établissement. L'effet de proximité pourrait également jouer un rôle pour les bactéries ayant des interactions mutualistes. Finalement, les facteurs génétiques de l'hôte peuvent influencer la composition finale du microbiome. (D) Exemple de dynamique pouvant mener à une dysbiose, soit un microbiome anormal associé à un état pathologique. Tiré de (Lloyd-Price *et al*, 2016).

1.3.1. Plus qu'une population bactérienne

En plus des bactéries, le microbiote normal est composé de mycètes (mycobiome), d'archéobactéries, de bactériophages (phageome) et même de protozoaires. Par contre, ces derniers

sont beaucoup moins représentés dans le microbiome. Par exemple, les mycètes représentent 0,1 % de séquences d'ADN retrouvées par séquençage de type « *shotgun* » (Underhill & Iliev, 2014). On retrouve 50 genres de mycètes dans le microbiome entérique normal (Hoffmann *et al*, 2013), contre 75 au niveau buccal (Ghannoum *et al*, 2010), 40 à 80 au niveau de la peau (Findley *et al*, 2013) et 11 à 20 au niveau vaginal (Drell *et al*, 2013). La composition du mycobioime semble avoir un effet sur le développement du système immunitaire et sur la composition du microbiome bactérien, mais beaucoup de travail reste à faire pour élucider sa composition, entre autres dans le séquençage de génome complet de mycètes (Underhill & Iliev, 2014). Les archéobactéries présentes dans le microbiome entérique sont souvent méthanogènes et transforment les produits de fermentation de bactéries en méthane (Hoffmann *et al*, 2013). Pour ce qui est du phageome, il pourrait contribuer au contrôle de la composition bactérienne du microbiome et potentiellement aider à combattre des pathogènes (Manrique *et al*, 2016). Finalement, le rôle des protozoaires dans le microbiome a été très peu investigué. L'hypothèse la plus probable est qu'ils aident à moduler la composition bactérienne du microbiome (Chabé *et al*, 2017).

1.3.2. Fonctions du microbiome

Le microbiome humain peut être vu comme un organe ayant comme fonction de compléter le génome humain. Louis Pasteur a déclaré en 1885 qu'il croyait que le microbiome était d'une telle importance pour l'animal qu'il devait être essentiel (Wostmann, 1981). Cette idée a par contre rapidement été démentie par un groupe de recherche lorsqu'ils ont créé le premier cobaye sans germes en 1895, bien qu'il ne vécut que 13 jours dû à une diète non adaptée aux animaux sans germes (Wostmann, 1981). Bien que le microbiome ne soit pas essentiel à la survie de l'animal, il assure plusieurs fonctions que le génome de son hôte ne peut pas accomplir.

La fonction la plus évidente du microbiome est d'aider à la digestion. En effet, il a été démontré que le microbiome entérique aide à digérer des xénobiotiques (Gill *et al*, 2006), soit des substances non métabolisables par le corps de l'hôte. Le microbiome peut également aider à rendre des sucres et des lipides plus biodisponibles (Ley *et al*, 2005). D'ailleurs, il a été démontré chez la souris que la composition du microbiome change en fonction de l'alimentation, particulièrement chez les

individus obèses (Ley *et al*, 2006; Turnbaugh *et al*, 2006), provoquant ainsi un enrichissement des fonctions d'assimilation d'énergie (digestion des sucres et des lipides) chez les souris (Ley *et al*, 2005) et les humains (Ley *et al*, 2006; Turnbaugh *et al*, 2006). En plus de ses rôles sur la digestion, le microbiome agit également dans la régulation du métabolisme de son hôte (Bäckhed *et al*, 2007).

Le microbiote a également un rôle important à jouer pour l'éducation du système immunitaire, l'expression de peptides antibactériens naturels (Cash *et al*, 2006) et la prévention du développement de réponses auto-immunes comme l'asthme (Braun-Fahrlander *et al*, 2002; Kozyrskyj *et al*, 2007).

Quelques études suggèrent que le microbiome pourrait avoir une influence sur le développement postnatal. Des souris sans germes ont, par exemple, des cœurs plus petits que leurs équivalents colonisés (Turnbaugh *et al*, 2007), ainsi que des cæcums élargis (Wostmann & Bruckner-Kardoss, 1959). Une réduction de la surface totale de la lumière intestinale est également observée chez les souris sans germes (Meslin *et al*, 1973).

Plus récemment, plusieurs études ont démontré un rôle du microbiome dans le développement cérébral (Luczynski *et al*, 2016), connu sous l'appellation axe microbio-entéro-cérébral (Rhee, Sang H, Pothoulakis, C and Mayer, 2009). Entre autres, il a été établi que les souris sans germes sont plus actives physiquement que les souris ayant un microbiome (Bäckhed *et al*, 2007). Ce domaine a également permis la découverte des psychobiotiques (p. ex. *Bifidobacterium infantis*), soit des probiotiques ayant un effet bénéfique sur le cerveau (Dinan *et al*, 2013). La plupart des études sont toutefois à prendre avec circonspection puisque dans la plupart des cas, les résultats sont variables d'une souche de souris à une autre ou d'une espèce à une autre (Luczynski *et al*, 2016).

1.3.3. Technologies de modification de la composition du microbiome

Le microbiome ayant des fonctions très diversifiées, il serait intéressant de développer des outils afin de pouvoir le modifier. De tels outils permettraient à la fois de mieux comprendre le

microbiome et de traiter les états pathologiques du microbiome (dysbioses). Dans cette section, les quelques outils disponibles pour la modification du microbiome et son étude seront décrits.

1.3.3.1 Les organismes sans germes

Une des façons d'étudier le microbiome est de complètement le retirer. Les organismes sans germes ont été convoités depuis l'époque de Pasteur (Wostmann, 1981). Depuis, la gnotobiologie, c'est-à-dire la biologie visant l'étude d'un environnement présentant uniquement des formes de vie définies, a grandement évolué (Williams, 2014). La procédure à suivre afin de générer une lignée animale sans germes est assez similaire, peu importe l'animal choisi. Pour un mammifère, les bébés sont mis au monde par césarienne en prenant soin d'éviter tout contact du bébé avec les parois vaginales ou la peau de sa mère. Les bébés sont ensuite déposés dans des environnements stériles et nourris avec de la nourriture stérile pour le reste de leur vie (Luczynski *et al*, 2016). Une méthode alternative impliquant une transplantation d'embryons au stade de deux cellules dans une souris pseudo-enceinte sans germes a également été développée (Smith *et al*, 2007). Pour les poissons, le sperme et les œufs doivent être extraits par dissection pour éviter la contamination par l'eau ambiante. Des étapes additionnelles de désinfection sont requises avant que les œufs soient fécondés et déposés dans un bocal stérile (Rawls *et al*, 2004). Pour les oiseaux, les œufs sont simplement désinfectés et placés dans un environnement stérile. Le bébé est extrait de l'œuf en prenant soin qu'il ne touche pas aux parois externes et il est ensuite maintenu dans un environnement stérile (Phillips *et al*, 1962).

Une fois la colonie établie, des tests routiniers assurent l'absence de micro-organismes dans les fèces et la cage. Une reproduction naturelle est alors possible. Bien que les souris et les rats sans germes soient beaucoup plus utilisés (Al-Asmakh & Zadjali, 2015), des cobayes, poules, cochons, poissons et chiens sans germe ont, entre autres, également été développés (Wostmann, 1981).

La technologie des animaux sans germes a permis l'étude des fonctions du microbiome et est couramment utilisée de pair avec d'autres technologies comme la transplantation fécale et la colonisation de souches pures. Ces animaux nous ont entre autres permis d'en apprendre beaucoup

sur la nutrition (Wostmann, 1981), le cerveau (Luczynski *et al*, 2016) et le système immunitaire (Rooks & Garrett, 2016). Plus tard, les souris sans germes ont permis l'établissement de souris certifiées sans pathogènes spécifiques et de souris gnotobiotiques pour lesquelles tous les organismes composant son microbiote sont connus (Bibiloni, 2012).

1.3.3.2 *Les souris sans pathogènes spécifiques (SPS)*

La technologie d'animaux sans germes a également été utilisée afin de développer des méthodes de certification de l'état de santé des animaux de laboratoire. Pour développer des souris SPS, des colonies de souris sans germes sont sorties de leur environnement stérile, inoculées avec quelques isolats de fèces d'une souris saine et nourries avec de la nourriture conventionnelle (Schaedler *et al*, 1965). Des tests routiniers utilisés pour la certification de ces souris permettent la détection de bactéries pathogènes et de bactéries opportunistes (Nicklas *et al*, 2002). Ces souris sont intéressantes afin d'isoler les interactions immunologiques entre l'hôte et le microbiome étant donné l'absence de bactéries pathogènes. Par contre, la normalité de la composition du microbiome de ces souris est toujours à définir et donc les résultats quant à l'interaction entre les micro-organismes et l'hôte peuvent ne pas être fidèle à la réalité (Bibiloni, 2012).

1.3.3.3 *Les souris gnotobiotiques*

Plusieurs études ont pour but l'introduction de souches dans le microbiome sauvage ou affaibli de la souris. Quelques-unes de ces expériences sont discutées brièvement plus bas. Cette section se concentre sur la génération de souches de souris ayant été colonisées de façon stable par des monocultures ou des mélanges de bactéries commensales. Des souris colonisées en monocultures sont particulièrement intéressantes dans l'étude des interactions directes entre la bactérie et l'hôte. Par exemple, des souris colonisées en monoculture ont permis de déterminer que *Bacteroides thetaiotaomicron* utilisait le fucose comme source d'énergie et était capable de signaler à son hôte de produire plus de glycan-fucose lorsque les concentrations en nutriments dans l'intestin deviennent trop faibles (Bry *et al*, 1996). Des études avec des souris bicolonisées ont permis d'explorer les interactions entre deux bactéries commensales dans leur environnement naturel. Par

exemple, une bicolonisation de *B. thetaiotaomicron* et de *Eubacterium rectale* nous a montré que chaque micro-organisme réagissait à sa façon à la compétition et s'adaptait à la présence d'autres organismes. Dans cet exemple, *B. thetaiotaomicron* changeait son métabolisme pour assimiler une plus grande variété de sources de carbone, et *E. rectale* se spécialisait dans l'utilisation de seulement quelques sources de carbone (Mahowald *et al*, 2009). Par contre, cette adaptation diffère selon la paire de micro-organismes utilisée (Samuel & Gordon, 2006).

Enfin, les souris sans germes ont été utilisées afin de générer des souris au microbiome simplifié et défini. Des expériences de colonisation avec 10 souches bactériennes ont montré que l'introduction simultanée du mélange de bactéries menait à une instabilité de colonisation. Il était préférable de d'abord monocoloniser des souris avec chaque souche, puis de mélanger ces souris dans une même cage afin d'exploiter leur nature saprophyte. Les bactéries ainsi introduites présentaient des adaptations favorisant leur colonisation, ce qui aboutissait en de meilleurs taux de colonisation à long terme (Rezzonico *et al*, 2011). Certaines études ont également montré que la colonisation de sept souches d'origine humaine complétait les fonctions métaboliques manquantes chez les rats sans germes (Becker *et al*, 2011). Certains microbiomes définis ont d'ailleurs été commercialisés, comme le microbiome altéré de *Schaedler*, populaire dans les années 1970-1980 (Brand *et al*, 2015).

1.3.3.4 Les traitements antibiotiques

Bien que les antibiotiques soient beaucoup plus utilisés dans le traitement des maladies infectieuses, leur spectre d'action est bien caractérisé, ce qui permet de modifier la composition du microbiome (Langdon *et al*, 2016). Aussi, leur utilisation pour la génération de souris sans germes a été récemment proposée (Lundberg *et al*, 2016). En effet, les souris sans germes ne peuvent répondre à certaines questions, comme l'effet qu'a un changement dans la composition du microbiome à certaines étapes de la vie. Aussi, l'utilisation de souris sans germes amène son lot de problématiques, puisque de nombreux procédés physiologiques sont affectés durant le développement du corps de l'animal sans microbiome. L'utilisation d'antibiotiques afin de dépléter le microbiome à différents stades de développement a donc été proposée afin de répondre

à ces problèmes (Lundberg *et al*, 2016). Cette technique est régulièrement utilisée conjointement avec la transplantation fécale (Ellekilde *et al*, 2015) ou des essais de colonisation (Sidjabat *et al*, 2009) afin d'induire, à des moments précis, un changement dans le microbiome. On utilise souvent des antibiotiques à large spectre seuls (Ellekilde *et al*, 2015; Sidjabat *et al*, 2009) ou en combinaison (Hansen, 1995; Hu *et al*, 2015) pour éliminer ou faire diminuer l'abondance de la plupart des bactéries du microbiome (Hu *et al*, 2015), ou de certaines populations précises (Hansen *et al*, 2012). Certaines études exploitant les antibiotiques afin de modifier le microbiome ont déjà fait l'objet de critiques. Par exemple, il fut démontré que les antibiotiques peuvent avoir des influences directement sur l'expression des gènes chez l'animal traité (Cox *et al*, 2014). Aussi, l'utilisation d'antibiotiques sur de longues périodes entraîne la sélection de résistances et constitue un biorisque important (Modi *et al*, 2013). L'utilisation d'antibiotiques est donc un moyen viable, quoiqu'imprécis, de manipuler le microbiome.

1.3.3.5 Les transplantations fécales

La transplantation fécale permet le transfert de contenu entérique microbien d'un individu à un autre. Le transfert fécal peut être utilisé afin d'étudier l'effet d'un microbiome altéré sur son hôte en transférant ce microbiome dans un individu récepteur sans germes (Turnbaugh *et al*, 2006). Chez la souris, la transplantation fécale se fait en sacrifiant la souris donatrice et en effectuant l'extraction de son caecum en anaérobiose. Le contenu caecal est ensuite extrait et dilué dans une solution tamponnée. Il est alors possible de gaver la souris réceptrice avec le contenu caecal en limitant le plus possible l'exposition de l'inoculum à l'environnement aérobie (Ellekilde *et al*, 2015). La technique nous a permis de vérifier les liens entre le microbiome et certains phénotypes comme l'obésité (Turnbaugh *et al*, 2006), et d'investiguer les pressions sélectives entrant en jeu lors de l'établissement du microbiome (Rawls *et al*, 2006).

La transplantation fécale est également utilisée afin d'humaniser le microbiome d'animaux de laboratoire. Ce procédé est particulièrement utile pour procéder à des expériences sur le microbiome humain sans la variabilité génétique retrouvée entre les individus, ou pour faire des expériences qu'il ne serait pas éthique de faire chez l'humain. Des animaux sans germes sont

régulièrement utilisés comme récepteurs dans ces expériences pour faciliter la colonisation du microbiome humain (Bibiloni, 2012). Par contre, même avec des souris sans germes, il est légitime de se demander si leur utilisation comme modèle d'étude du microbiome humain est adéquate (Silley, 2009). En effet, de récentes études suggèrent que le microbiome transplanté d'une espèce à une autre change en composition suite à des différences au niveau des pressions sélectives sur les populations (Rawls *et al*, 2006). Entre autres, les différences entre la diète murine et humaine sont des causes possibles d'une modification de la composition du microbiome de la souris humanisée par rapport à son inoculum de départ (Bibiloni, 2012). Toutefois, certaines études suggèrent que la transplantation fécale humaine vers les souris sans germes produise des souris ayant un microbiome dont la composition est très similaire à celle d'un humain tant au niveau fonctionnel (Goodman *et al*, 2011) que descriptif (Turnbaugh *et al*, 2009). La fiabilité de cette méthode dans l'étude du microbiome humain est donc encore à débattre à ce jour. Il n'en reste pas moins que la transplantation fécale constitue une thérapie émergente pour le traitement des dysbioses et des patients qui souffrent d'infection récurrente à *Clostridium difficile* (Casals-Pascual *et al*, 2018). Son utilisation en médecine pose par contre certains défis dans la sélection du donneur afin d'éviter la transplantation fécale d'un microbiote contenant des pathogènes, pathobiontes ou des organismes résistants aux antibiotiques (Casals-Pascual *et al*, 2018).

1.3.3.6 L'ajout de probiotiques

La littérature ayant pour objet l'introduction de souches dans un microbiome, afin de le modifier, est très vaste. Seulement quelques exemples plus en lien avec mon projet seront donc donnés. Lors des essais de colonisation, la bactérie introduite est souvent pathogène et l'essai sert à étudier la pathogénèse de cette souche. De très bons modèles d'infection ont été développés pour certains pathogènes entériques comme *Citrobacter rodentium* (Crepin *et al*, 2016), *Salmonella enterica* sérovar *typhimurium* (Mathur *et al*, 2012), *V. cholerae* (Sawasvirojwong *et al*, 2013) et *Clostridium difficile* (Hutton *et al*, 2014). D'autres équipes ont concentré leurs efforts sur la caractérisation d'espèces commensales ou probiotiques. Par contre, les essais de colonisation se révèlent souvent infructueux dans des animaux dont le microbiote n'est pas déplété ou absent, dû à la résistance du microbiome intestinal à la colonisation. Ce phénomène s'applique également à

la résistance au pathogène, qui est grandement diminuée à la suite d'un traitement aux antibiotiques (Lawley & Walker, 2012). Bien que bénéfique sur plusieurs points, la résistance du microbiome à la colonisation rend plus fastidieux le traitement de pathologies par des probiotiques, comme le syndrome du côlon irritable (Grimm *et al*, 2015). Des études sur l'amélioration du potentiel de colonisation des probiotiques ont donc été entamées. Entre autres, une équipe a étudié l'effet de transformer *E. coli* avec une banque d'ADN provenant de *B. thetaiotaomicron*. Les gènes de *B. thetaiotaomicron* viennent alors compléter les gènes d'*E. coli* et l'aident à coloniser la souris. Dans cette étude, les gènes en relation avec le métabolisme du galactose et du sucrose étaient particulièrement enrichis (Yaung *et al*, 2015), montrant ainsi que le métabolisme de ces sucres avait un rôle important dans la colonisation de la souris. Par contre, de tels résultats ne peuvent être transposés sur l'humain puisqu'ils sont dépendants de la diète des souris qui est très différente de celle des humains. Une stratégie intéressante permettant de favoriser l'implémentation des souches probiotiques dans le microbiome est d'utiliser un prébiotique seulement métabolisable par le probiotique introduit. Par exemple, la porphyrane, un polysaccharide dérivé d'une algue rouge, fut utilisée comme prébiotique afin d'améliorer l'implémentation d'une souche de *Bacteroides ovatus*, modifiée génétiquement pour métaboliser ce sucre (Shepherd *et al*, 2018). L'ajout de porphyrane pouvait alors augmenter les taux de colonisation du probiotique de plusieurs ordres de grandeur. L'ajout de probiotiques seuls ou en combinaison est d'ailleurs le cœur de métier de plusieurs compagnies, dont Seres Thérapeutique, qui développe, entre autres, une alternative plus sécuritaire à la transplantation fécale composée d'un mélange connu de souches bactériennes (<https://www.serestherapeutics.com/our-programs/>).

1.3.3.7 Les bactériophages comme outil de modification du microbiote.

La thérapie par les bactériophages se base sur l'utilisation d'un ou de plusieurs phages lytiques, normalement très spécifiques pour un pathogène à éliminer, pour traiter une maladie infectieuse. L'utilisation de bactériophages dans le traitement d'infections, bien que peu courantes en Occident, est une pratique toujours actuelle dans plusieurs pays dont la Russie (Abedon *et al*, 2011). Par leur spécificité et leur grande efficacité d'infection, les bactériophages constituent un outil intéressant dans la manipulation des microbiomes. Par exemple, des bactériophages lytiques

pourraient être employés pour diminuer l'abondance de souches spécifiques causant une dysbiose, ou des phages tempérés pourraient être induits dans ces mêmes souches pour déclencher un cycle d'infection (Ganeshan & Hosseinidou, 2019). L'utilisation de bactériophages lytiques pour le contrôle du microbiote demeure toutefois difficile due au spectre d'hôte étroit et mal défini de la plupart des phages ainsi que de la tendance des bactéries à devenir rapidement résistantes à des souches de bactériophage (Ganeshan & Hosseinidou, 2019). Les bactériophages tempérés sont induits par des signaux encore mal compris, leur utilisation pour ce genre d'application impliquerait donc de trouver des signaux déclenchant leur cycle lytique qui soit spécifique au bactériophage ciblé (Ganeshan & Hosseinidou, 2019). Bien que l'utilisation des bactériophages pour l'édition du microbiote soit encore un champ de recherche relativement peu développé, il n'en demeure pas moins prometteur. Quelques compagnies, comme BiomX, se sont d'ailleurs lancés dans des applications médicales de modification du microbiome par les bactériophages naturels. Parmi leurs produits, BiomX est présentement en phase clinique I pour un mélange de phages ciblant *Cutibacterium acnes* dans le traitement de l'acné (<https://www.biomx.com/pipeline/#pipeline-acne>, comme présenté lors du 5th *Microbiome Movement – Drug Development Summit*, 2020). Aussi, combiner les bactériophages et CRISPR-Cas9 présenterait l'avantage de contrôler les populations bactériennes composant le microbiote selon leur contenu en gène (Bikard *et al*, 2014; Citorik *et al*, 2014; Yosef *et al*, 2015). Plusieurs compagnies hautement innovantes comme Eligo Bioscience, Locus Bioscience, SNIPR Biome et Nemesis Bioscience, ont d'ailleurs entrepris le développement et la commercialisation de médicaments basés sur des bactériophages mobilisant des modules CRISPR.

Les technologies décrites aux sections précédentes permettent de changer la composition du microbiote entérique en ajoutant ou retirant des populations de bactéries. Par contre, comme démontré par le « Human Microbiome Project », la composition bactérienne du microbiome varie considérablement d'un individu à l'autre, seules des fonctions de gènes contenus dans le microbiome sont conservées (Huttenhower *et al*, 2012). Un outil de manipulation du microbiome idéal devrait donc être capable de modifier son contenu sur la base des gènes présents, plutôt que sur la base de la présence de bactéries spécifiques. En éliminant des gènes encodant une fonction précise, l'étude de leur importance pour l'hôte deviendrait plus facile. Aussi, un tel outil pourrait permettre d'éliminer des gènes qui sont clairement nocifs chez les bactéries résidentes du

microbiote. Par exemple, des gènes encodant des facteurs de virulence comme des toxines ou des gènes de résistances aux antibiotiques pourraient être éliminés spécifiquement du microbiome, diminuant les chances de complications lors d'infections bactériennes ou de procédures opératoires (Buelow *et al*, 2014; Vincent, 2003). L'élimination de gènes de résistance aux antibiotiques dans le microbiome entérique est d'une importance particulière, puisque cet environnement est un réservoir important pour ces gènes.

1.3.3.8 La résistance aux antibiotiques et le microbiome entérique.

La résistance aux antibiotiques est largement considérée comme une menace importante à notre capacité de traiter des infections bactériennes. D'alarmantes projections montrent que les organismes résistants aux antibiotiques pourraient être la cause d'une grande crise sanitaire d'ici 2050 (Antimicrobial resistance: global report on surveillance 2014. Available online at <https://www.who.int/drugresistance/documents/surveillancereport/en/>, 2014). Bien qu'un problème grandissant pour la médecine moderne, les phénomènes de résistances aux antibiotiques ne datent pas d'hier (Davies & Davies, 2010). Les antibiotiques étant largement dérivés de molécules retrouvées dans la nature, il n'est pas étonnant que les bactéries aient évoluées pour résister à ces attaques (Schaik, 2015). On peut d'ailleurs estimer l'émergence du gène de résistance aux β -lactames (serine β -lactamase) à plus de 2 milliards d'années avant aujourd'hui (Hall & Barlow, 2004). Les phénomènes d'acquisition de résistances peuvent être très rapides. Certains gènes comme le 2'-*N*-acetyltransferase provenant de *Providencia stuartii* qui a pour fonction première la modification du peptidoglycane chez son hôte d'origine procurent une résistance aux aminoglycosides lorsque transférés à d'autres bactéries dues à la similarité structurale entre les deux substrats (Macinga & Rather, 1999). La résistance aux antibiotiques peut donc être acquise rapidement par transfert de gènes permettant ainsi l'émergence de pathogènes résistants aux antibiotiques (Tamminen *et al*, 2012). L'accumulation récente de ces gènes de résistance, autant chez les bactéries commensales que pathogènes, peut-être attribuée à l'industrialisation de leur usage depuis la fin de la Deuxième Guerre mondiale (Ventola, 2015). En effet, l'utilisation des antibiotiques pour la prévention ou le traitement des infections bactériennes tant chez l'humain

que les animaux d'élevage sélectionnent les organismes résistants en leur conférant un avantage de croissance clair (Casals-Pascual *et al*, 2018).

Une stratégie intéressante pour permettre l'utilisation au long terme des molécules antibiotique est de limiter l'exposition aux antibiotiques des organismes non pathogènes (Schaik, 2015). En identifiant les environnements agissants comme réservoir d'organismes résistants aux antibiotiques ainsi que les phyla de ces organismes, nous pourrions prévenir la sélection des organismes résistants aux antibiotiques. Par exemple, des antibiotiques à spectre d'action plus restreints ou des combinaisons d'antibiotiques pourraient être utilisés dans ces environnements pour éviter l'accumulation des gènes de résistance aux antibiotiques (Pérez-Cobas *et al*, 2013). C'est sous cette philosophie qu'est née l'étude du résistome, soit le sondage des populations bactériennes dans un environnement donné dans le but de définir la prévalence des gènes de résistance (D'Costa *et al*, 2006). Il existe plusieurs méthodes d'analyse du résistome qui génèrent des données complémentaires. Une première façon d'étudier le résistome est par la culture de bactéries provenant de l'environnement en laboratoire (Schaik, 2015). Des échantillons environnementaux sont collectés, puis étalés sur boîte de pétri contenant différents antibiotiques pour évaluer le profil de résistance. Malgré de récentes avancés dans le domaine, la culture d'une grande partie des micro-organismes isolés de l'environnement demeure difficile (Rettedal *et al*, 2014). Afin de sonder une plus grande proportion des organismes composant le microbiote local, d'autres techniques comme le séquençage de type « *shotgun* » de l'ADN total sont également employées (Forslund *et al*, 2014). Pour cette technique, l'ADN total est extrait de l'échantillon environnemental, puis des librairies d'ADN de nouvelle génération sont préparées à partir de l'ADN mélangé de tout ce qui se trouve dans l'échantillon. Les séquences sont ensuite assemblées *de novo* puis analysées à l'aide de base de données de gènes de résistance comme CARD pour identifier les gènes présents (McArthur *et al*, 2013). Bien que cette technique soit plus complète, elle demeure dépendante des banques de données et ne peut pas identifier de nouveaux gènes de résistances encore inconnus. Pour ce faire, d'autres groupes ont plutôt opté pour des techniques d'analyse fonctionnelle. Dans de tels techniques, l'ADN extrait de l'échantillon environnemental est cloné dans un vecteur sous forme de petits inserts de quelques kilobases (Sommer & Dantas, 2011) ou sous forme d'inserts d'environ 40 kilobases (Allen *et al*, 2009). Cette technique permet de découvrir de nouveaux gènes de résistance chez *E. coli* et simule un transfert horizontal de gène.

L'étude des résistomes a rapidement identifié le microbiote entérique comme un important réservoir de gènes de résistance aux antibiotiques (Schaik, 2015; Casals-Pascual *et al*, 2018). En effet, les traitements antibiotiques à large spectre affectent autant le pathogène ciblé que les résidents du microbiote entérique, provoquant une importante mortalité et parfois un déséquilibre des populations bactériennes (Pilmis *et al*, 2020). Dans un tel contexte, les bactéries résidentes du microbiote entérique ont un avantage sélectif important à posséder un gène de résistance aux antibiotiques. Il fut d'ailleurs démontré que les bactéries anaérobies strictes du phylum *Firmicute* et *Bacteroidetes* sont d'importants réservoirs de gènes de résistances aux antibiotiques dans le microbiote entérique (Schaik, 2015). De façon alarmante, d'importantes études chez plus de 300 patients provenant de la Chine, l'Italie, la France, du Malawi, du Japon l'Espagne, des États-Unis et du Danemark ont révélées la présence ubiquitaire de gènes de résistances aux antibiotiques dans le microbiote entérique avec en moyenne 21 gènes de résistance différents par échantillon (Forslund *et al*, 2013; Hu *et al*, 2013). Par ailleurs, tous les microbiotes entériques sondés dans ces études contenaient des gènes de résistance à la tétracycline et à la bacitracine.

Bien que souvent retrouvée dans des bactéries non pathogènes, l'accumulation de gènes de résistance dans le microbiote entérique demeure alarmante. En effet, plusieurs mécanismes de transfert horizontal de gènes peuvent contribuer à la mobilité des gènes de résistance entre différentes espèces bactériennes. Les trois principaux mécanismes de transfert horizontal de gènes sont la transformation, la transduction et la conjugaison (Furuya & Lowy, 2006). La transformation est dépendante de la compétence bactérienne et consiste en l'acquisition active de brins d'ADN nus dans l'environnement entourant la bactérie. La contribution de ce mécanisme de transfert horizontal de gènes a rapidement été identifiée comme étant beaucoup moins importante que les deux autres mécanismes (Nordgård *et al*, 2012). La transduction contribue significativement à la mobilité des gènes de résistance aux antibiotiques (Modi *et al*, 2013). Ce phénomène repose sur le transfert de gènes par l'intermédiaire d'un bactériophage. Les gènes de résistance peuvent être intégrés dans le génome du phage, ou peuvent se trouver dans le chromosome de l'hôte bactérien et être encapsidés par erreur pour les phages lytique ou lysogène. Bien qu'un important véhicule de gènes entre différentes bactéries, les bactériophages ont un spectre d'hôte souvent réduit (Ross *et al*, 2016). Le transfert horizontal de gènes de résistances aux antibiotiques par transduction est

donc surtout réservé à des transferts entre bactéries apparentés. Le dernier mécanisme de transfert horizontal de gènes, soit la conjugaison bactérienne, implique une machinerie de sécrétion spécialisée qui forme un pore entre la cellule portant l'élément conjugatif et une bactérie réceptrice (Arutyunov & Frost, 2013). Ce mécanisme spécialisé contribue largement à la mobilisation des gènes dans le microbiote entérique parfois entre espèces relativement éloignées (Tamminen *et al*, 2012). Par contre, les règles régissant le transfert conjugatif dans le microbiote entérique sont toujours mal comprises.

1.4 La conjugaison bactérienne

La conjugaison bactérienne est un mécanisme de transfert horizontal de gènes qui a été découvert par Lederberg et Tatum en 1947 (Tatum & Lederberg, 1947). Depuis, bon nombre d'études tant mécanistiques qu'épidémiologiques sont parues dans la littérature. Ces études nous ont fait comprendre quelques grands mécanismes sous-jacents à la conjugaison bactérienne. Ainsi, la conjugaison peut être décrite comme le transfert d'un élément génétique mobile d'une bactérie à une autre en contact direct via la formation d'un pore intermembranaire (Llosa *et al*, 2002). L'élément génétique mobile peut être un plasmide, un élément intégratif conjugatif (ICE), un transposon conjugatif ou un îlot génomique (Juhas *et al*, 2009). L'élément peut être mobilisé par sa propre machinerie de conjugaison, ou mobilisé en *trans* par la machinerie de conjugaison d'un autre élément conjugatif (Hochhut *et al*, 2000). La mobilisation des plasmides conjugatifs est largement dépendante de la présence d'une séquence d'origine de transfert (*oriT*) qui est reconnue par une relaxase (Fu *et al*, 1991). Le mécanisme exact par lequel le plasmide est transféré d'une cellule à une autre varie entre chaque système conjugatif. Le cas du plasmide F, un des plasmides conjugatifs modèles les plus étudiés, sera présenté ci-dessous. Aux fins de simplification, le brin du plasmide F qui est transféré à la cellule réceptrice sera nommé ici le brin « transféré », tandis que le brin « matrice » fera référence au brin servant à la préservation et la réplication du plasmide dans la cellule donneuse.

Un complexe réunissant plusieurs protéines se lie d'abord à la séquence *oriT* du plasmide, formant ainsi le relaxosome. La nickase comprise dans le relaxosome induit alors une coupure simple brin

à son site spécifique sur le brin transféré. À la suite de la coupure simple brin de l'*oriT*, la nickase (TraI) reste liée de façon covalente à l'extrémité 5' du brin transféré. Puis, le relaxosome sépare les deux brins d'ADN (Howard *et al*, 1995). L'extrémité 3' du brin d'ADN encore attaché au brin matrice sert alors d'amorce pour la réplication du brin transféré. Une seconde nickase empêche que le brin transféré ne contienne plus d'une fois la séquence du plasmide en coupant une deuxième fois le brin transféré lorsque la séquence *oriT* est dupliquée. Le relaxosome se lie alors à la protéine de couplage (TraD) du système de sécrétion de type 4 (T4SS), ce qui induit le transfert du brin vers la cellule réceptrice par le T4SS. Lors du transfert, l'extrémité 5' du brin est gardée dans la cellule donneuse par un dimère de TraM. À la suite du transfert de la majorité du plasmide, l'extrémité 3' du brin (également liée par un dimère de TraM) se trouve à proximité du T4SS. Les deux dimères de TraM forment alors un tétramère et catalysent la circularisation du brin transféré (Firth *et al*, 1996). Par contre, la formation du brin complémentaire d'ADN dépend de procédés encore inconnus (Arutyunov & Frost, 2013).

La conjugaison peut atteindre des efficacités très élevées (Bradley *et al*, 1980) et ce, entre une bactérie donneuse et réceptrice parfois éloignées phylogénétiquement (Tamminen *et al*, 2012). Le spectre d'hôte des plasmides conjugatifs, soit l'ensemble des espèces bactériennes pouvant porter un élément conjugatif donnée, est relativement peu étudié. Par contre, la connaissance du spectre d'hôte des différents éléments conjugatifs est cruciale afin d'évaluer leur contribution à l'émergence de pathogènes résistants aux antibiotiques.

1.4.1. Le spectre d'hôte des plasmides conjugatifs

La littérature sur le spectre d'hôte des grandes familles de plasmides conjugatifs est grandement dépendante d'événements anecdotiques rapportés dans la littérature. En ce sens, le spectre d'hôte de chaque système est souvent dépendant d'études épidémiologiques ou d'expériences en laboratoire à petite échelle. Les spectres d'hôte mentionnés dans cette section sont donc limités à ce qui est connu pour chaque système conjugatif à ce jour. Une étude de criblage à haut débit du spectre d'hôte de divers plasmides conjugatifs serait particulièrement intéressante.

Avant de décrire le spectre d'hôte des plasmides conjugatifs, il faut d'abord comprendre qu'il existe plusieurs types de plasmides conjugatifs qui sont regroupés en familles. La première classification systématique des plasmides conjugatifs a d'abord été en groupes d'incompatibilité (Couturier *et al*, 1988). L'incompatibilité (Inc) peut être définie comme l'incapacité de deux plasmides à être propagés dans une même cellule. La base de cette classification était que deux plasmides incompatibles devaient posséder des systèmes de réplication ou de partitionnement similaire et donc être apparentés. Jusqu'à ce jour, l'incompatibilité a été classifiée dans trois groupes bactériens différents, soit 27 Inc chez les *Enterobacteriaceae*, 17 Inc chez *Staphylococcus* et 14 Inc chez *Pseudomonas* (Shintani *et al*, 2015). Cette classification concerne à la fois les plasmides conjugatifs, mobilisables et non mobiles. Par contre, nous savons aujourd'hui que l'incapacité de deux plasmides à co-exister peut être due à d'autres phénomènes, par exemple à l'exclusion plasmidique (Garcillán-Barcia & De la Cruz, 2008). Aussi, bien que l'incompatibilité des systèmes de réplication soit une bonne façon d'évaluer rapidement la distance phylogénétique entre les éléments, elle n'est pas très indicative de la proximité phylogénique de la machinerie de transfert conjugatif. De ce fait, le groupe de Laura Frost a plus récemment proposé de classifier les éléments génétiques mobiles par homologie de séquence de leur nickase (Garcillán-Barcia *et al*, 2009) ou des protéines formant le Système de Sécrétion de type 4 (T4SS) (Smillie *et al*, 2010). Ceci a permis de classer les plasmides selon deux nouvelles nomenclatures, soit par groupe de mobilité (MOB) formant plus de 24 classes (divisées en six grandes familles), ou par groupe T4SS (MPF) formant 8 grandes familles. La classification par élément MOB est particulièrement intéressante puisqu'elle permet la classification de tous les éléments génétiques mobiles alors que la classification MPF se limite aux éléments conjugatifs encodant leur propre machinerie de transfert. En effet, la nickase est le seul élément essentiel à la mobilité du plasmide, étant donné qu'elle est responsable de la reconnaissance de l'*oriT*. Un élément mobile peut parasiter les systèmes de transfert d'autres éléments conjugatifs (Núñez & De la Cruz, 2001).

Les groupes MPF (T4SS) sont les groupes de plasmides les plus généraux et classent les plasmides par archétype de T4SS. En classant les plasmides par groupe MPF, on peut grouper les groupes MOB ensemble, qui classent ensuite les plasmides d'un même groupe Inc ensemble (Garcillán-Barcia *et al*, 2011). Le spectre d'hôte pour chacun des groupes Inc, MOB et MPF, ainsi qu'un représentant de chacun de ces groupes, sont montrés dans le tableau 1.2 basé principalement sur

trois revues de littérature (Garcillán-Barcia *et al*, 2009; Smillie *et al*, 2010; Garcillán-Barcia *et al*, 2011). Somme toute, beaucoup de travail reste nécessaire afin de caractériser en profondeur les spectres d'hôte des systèmes conjugatifs de façon précise.

Tableau 1.2. Spectre d'hôte des familles de systèmes conjugatifs.

Classe MPF	Classe MOB	Classe Inc	Spectre d'hôte	Exemple
T	MOB _{F11}	IncW + N	<i>Proteobacteria</i>	R388
F	MOB _{F12}	IncF ₁ + F ₂	<i>Proteobacteria</i>	F plasmid
F	MOB _{H11}	IncHI ₁ + HI ₂	<i>γ-Proteobacteria</i>	R27
F	MOB _{H12}	IncA + C + ICE SXT/R391 + T + P ₇	<i>γ-Proteobacteria</i>	pVCR94
G	MOB _{H2}		<i>γ-Proteobacteria</i> ; <i>β-Proteobacteria</i>	pKLC102
-	MOB _{C1}	(mobilisable)	<i>γ-Proteobacteria</i>	ClodF13
****	MOB _{C2}	IncHI ₃	<i>Firmicutes</i> ; <i>Tenericutes</i>	pAD1
-	MOB _{Q1}	IncQ ₁ + P ₄	<i>Proteobacteria</i>	RSF1010
****	MOB _{Q2}	****	<i>Proteobacteria</i>	p42d
****	MOB _{Q3}	Inc18	<i>Firmicutes</i>	pIP501
T	MOB _{P11}	IncP ₁	<i>Proteobacteria</i> ***	RP4
I	MOB _{P12}	IncI ₁ + K	<i>Proteobacteria</i>	R64
I	MOB _{P13}	IncL/M	<i>Proteobacteria</i>	pCTX-M3
-	MOB _{P14}	IncQ ₂ + G + P ₆	<i>Proteobacteria</i>	pTF-FC2
****	MOB _{P2}	****	<i>α-Proteobacteria</i>	pTiC58
T	MOB _{P3}	IncX ₁ + X ₂	<i>γ-Proteobacteria</i>	R6K
T	MOB _{P4}	IncU	<i>Proteobacteria</i>	pFBAOT6
-	MOB _{P5}	ColE1	<i>Proteobacteria</i>	ColE1
T	MOB _{P6}	IncI ₂	<i>Proteobacteria</i>	TP114
****	MOB _{P7}	Inc4 + 10	<i>Firmicutes</i> ; <i>Fusobacteria</i> ; <i>ε-Proteobacteria</i> ; <i>δ-Proteobacteria</i>	pC221

Tableau 1.2 : (suite)

****	MOB _{V1}	Inc11 + 13	Firmicutes	pE194
-	MOB _{V2}	****	<i>Cyanobacteria</i>	pBBR1
****	MOB _{V3}	Tn**	<i>Bacteroidetes</i>	Tn5520
****	MOB _{V4}	Tn**	<i>Bacteroidetes ; Firmicutes</i>	Tn4555
****	Non classé	Tn**	<i>Firmicutes</i>	Tn916

(-) : absence de T4SS

*GI : îlot génomique

**Tn : transposon

***RP4 a déjà été reporté pour son transfert chez *Clostridium difficile* (*Firmicutes*)

****Non déterminé

1.4.2. La conjugaison bactérienne dans le microbiome entérique

Bien que plusieurs évidences tendent à montrer que le microbiome entérique est un lieu particulièrement actif pour le transfert horizontal de gènes, très peu d'études ont pour objet l'étude de l'efficacité de transfert dans ce contexte naturel. La plupart de nos connaissances nous viennent d'études épidémiologiques (Chen *et al*, 2013; Sun *et al*, 2016) ou d'études sur leur impact dans l'évolution des espèces endogènes entériques (Soucy *et al*, 2015; Norman *et al*, 2009). Il serait donc primordial de bien caractériser la fréquence de transfert des grandes familles de systèmes conjuguatifs, ne serait-ce que pour pouvoir identifier les systèmes les plus à risque de se propager lors de traitements aux antibiotiques.

Dans cette optique, quelques groupes ont quantifié les taux de transfert conjuguatifs dans l'intestin de souris. Les fréquences de transfert mesurées étaient les suivantes : pAMβ1 <1 x 10⁻⁷ (Igimi *et al*, 1996), pAT191 1 x 10⁻⁹ (Doucet-Populaire *et al*, 1992), Tn1545 1,1 x 10⁻⁸ (Doucet-Populaire *et al*, 1991), pHUSEC41-1 sans quantification (Gumpert *et al*, 2017), p2 à 100 % (Stecher *et al*, 2012). Certaines de ces études utilisaient des souris sans germes (Doucet-Populaire *et al*, 1992, 1991), des souris gnotobiotiques (Stecher *et al*, 2012) ou des souris traitées aux antibiotiques (Licht *et al*, 1999; Gumpert *et al*, 2017) afin de maximiser la colonisation des souches donneuses et réceptrices. Par contre, aucune étude n'a quantifié les taux de transfert de plusieurs plasmides dans un même contexte expérimental. Comme il a récemment été démontré que la complexité du

microbiote entérique ainsi que le contexte génétique de la souche de souris utilisée influencent les taux de transfert de pCVM29188_146 (Ott *et al*, 2020), il apparaît difficile de comparer les résultats obtenus dans les différentes études. De plus, les conclusions de certaines de ces études peuvent être biaisées puisque les cellules donneuses et réceptrices étaient mélangées avant le gavage chez la souris. C'est entre autres le cas de l'étude sur le plasmide p2 ayant rapporté 100 % d'efficacité de transfert (Stecher *et al*, 2012). Puisque la conjugaison est un processus très rapide, des transconjugants auraient déjà dû être détectables avant même l'introduction du mélange dans la souris. Il est donc primordial de développer de meilleures méthodes et de quantifier dans un même contexte les différentes familles de plasmides conjugatifs dans le microbiome entérique.

Le transfert inter-espèce des plasmides conjugatifs vers des bactéries endogènes du microbiome sans toutefois être quantifié a tout de même été démontré, entre autres avec pES1 (Aviv *et al*, 2016). Voulant caractériser davantage le comportement de leur plasmide dans le microbiote entérique, ce même groupe a ensuite cherché à recréer ces conditions en laboratoire. Les taux de conjugaison de pES1 en condition microaérobie, à 37 °C et en présence de sels biliaries, trois caractéristiques retrouvées dans le microbiome entérique, permettaient au plasmide de se mobiliser à des taux similaires à ceux observés chez la souris. Par contre, ces résultats ne sont pas en accord avec ceux obtenus avec le plasmide R1drd19, où son taux de conjugaison dans l'intestin a été comparé à celui sur agar, en milieu liquide et en biofermenteur en faisant varier de multiples paramètres (Licht *et al*, 1999). Ces résultats montraient que l'efficacité conjugative enregistrée *in vivo* ($\approx 1 \times 10^{-4}$) était plus ressemblante à l'efficacité de leur plasmide en biofilm (sur agar) qu'en bouillon. Ceci suggère que la majorité du transfert de gènes par conjugaison se fait dans le mucus intestinal lorsque les bactéries y forment un biofilm. Par contre, ce groupe a noté une diminution de la fréquence de transfert *in vivo* par rapport aux fréquences observées en biofilm. Aussi, ce groupe n'a pas considéré les changements au niveau de la régulation du plasmide qui pourrait survenir lors de son introduction dans le système digestif. Outre les facteurs environnementaux, les cellules de l'hôte semblent pouvoir avoir un effet sur la conjugaison. En effet, des expériences menées avec un plasmide non identifié montrent l'inhibition de son transfert conjugal en présence de cellules Caco-1 humaines (Machado & Sommer, 2014). Ce mécanisme pourrait expliquer en partie la baisse d'efficacité conjugative *in vivo* notée avec R1drd19. Par contre, ce phénomène pourrait être spécifique à chaque plasmide puisque le plasmide RP4 est en mesure de conjuguer à

haute fréquence à l'intérieur de cellules humaines en culture (Yin *et al*, 2008). Pour trouver les facteurs environnementaux qui affectent le transfert du plasmide pSLT *in vivo*, l'impact de différentes conditions sur le transfert de pSLT a été évalué. La présence de sels biliaires, de proprionate ou d'extraits de fèces semble inhiber le transfert de pSLT alors qu'un milieu 0,3 M NaCl en condition microaérobie rétablissait l'efficacité de transfert trouvé *in vivo* (1×10^{-2}) (García-Quintanilla *et al*, 2008). La relation entre l'hôte et la conjugaison entre bactéries est donc un peu plus complexe qu'une simple répression du taux de transfert. Somme toute, une caractérisation systématique d'un grand nombre de plasmides conjugatifs serait nécessaire pour comprendre comment le transfert horizontal de gènes par conjugaison affecte la dissémination des résistances aux antibiotiques dans le microbiote entérique.

1.5 Objectifs et hypothèses du projet de recherche

Le microbiome entérique joue un rôle important pour la santé humaine (Lloyd-Price *et al*, 2016). De par sa grande contribution à différents processus biologiques, on peut voir le microbiome entérique comme un organe du corps humain (Laukens *et al*, 2015). Comme tout autre organe, le microbiote entérique peut développer des pathologies qui lui sont propres, actuellement regroupées sous le terme "dysbiose" (Belizário & Napolitano, 2015). La découverte de l'importance du microbiote entérique étant encore récente, peu d'outils sont à notre disposition pour traiter les dysbioses.

Mon projet de doctorat porte sur le développement d'un outil d'édition du microbiome capable d'éliminer spécifiquement des bactéries sur la base de leur contenu génique. L'outil que j'ai développé, nommé "Conjugative Probiotic" (COP), utilise la conjugaison bactérienne pour mobiliser un module CRISPR-*cas9* vers des bactéries résidentes du microbiote entérique. Une fois transféré à la bactérie réceptrice, le système CRISPR-Cas9 peut ensuite éliminer sa cellule hôte sur la base de la présence de gènes spécifiques. Bien que des technologies utilisant les bactériophages comme véhicules pour CRISPR-Cas9 ont déjà été rapportées dans la littérature (Bikard *et al*, 2014; Citorik *et al*, 2014; Yosef *et al*, 2015), nous pensons qu'un outil utilisant plutôt la conjugaison bactérienne comme véhicule pourrait être une approche avantageuse dans le

microbiote entérique. En effet, le microbiote entérique est un assemblage d'espèces bactériennes variées qui change d'un individu à l'autre. Les plasmides conjugatifs ayant souvent des spectres d'hôte plus large que les bactériophages, une technologie de livraison de CRISPR par conjugaison bactérienne offrirait une plus grande versatilité dans ce contexte. Aussi, plusieurs probiotiques sont disponibles commercialement, peuvent être administrés par voie orale chez les humains et résistent bien aux conditions rencontrées dans le tube digestif (pH, protéases, sels biliaires, etc.) (Nobrega *et al*, 2016; Sonnenborn & Schulze, 2009). Contrairement aux antibiotiques, le probiotique lui-même apporte une compétition aux bactéries partageant sa niche écologique au sein du microbiote entérique contribuant ainsi à la résistance à la colonisation (Pilmis *et al*, 2020). Finalement, un probiotique livrant CRISPR par conjugaison bactérienne peut également exercer d'autres activités thérapeutiques en parallèle. La livraison de CRISPR-Cas9 par conjugaison amène par contre plusieurs grands défis. En effet, l'efficacité des systèmes conjugatifs dans le microbiote entérique est encore mal caractérisée (Schaik, 2015). Les bactériophages offrent également souvent de meilleures efficacités de transfert d'ADN et un bioconfinement plus facile.

Nous avons sélectionné *Escherichia coli* Nissle1917 comme châssis cellulaire pour développer le COP. *E. coli* Nissle1917 est connue pour ses propriétés antidiarrhéique et a un long historique d'utilisation clinique (Sonnenborn & Schulze, 2009). L'importante application de notre système vise la décontamination de la résistance aux antibiotiques dans le microbiote entérique. Les bactéries *E. coli*, *Klebsiella pneumoniae* et *Enterococcus sp.* sont des cibles d'importance dans le microbiote entérique puisqu'elles sont souvent la cause d'infections nosocomiales suivant un traitement aux antibiotiques (Rice, 2008). Nous avons choisi de développer d'abord l'outil COP pour l'élimination d'*E. coli* résistante aux antibiotiques dans le microbiote entérique.

Afin de développer l'outil COP, je m'étais fixé trois objectifs principaux pour mon doctorat :

1. **Identifier une machinerie de transfert conjugatif hautement efficace pour le transfert d'ADN dans le milieu intestinal.** Cette étape est cruciale puisque l'efficacité d'élimination des bactéries par le système COP est directement corrélée à l'efficacité de transfert conjugatif. Pour ce premier objectif, nous avons obtenu 13 plasmides conjugatif

représentant toutes les familles de plasmides les plus souvent isolées chez les *Enterobacteriaceae* (Rozwandowicz *et al*, 2018). Nous avons ensuite quantifié l'efficacité de transfert de ces 13 plasmides conjugatifs sur agar, en bouillon et dans un modèle de souris standardisé. Ce criblage nous a permis d'identifier TP114 comme un élément conjugatif très efficace capable de se transférer vers virtuellement toutes les bactéries réceptrices sondées et de comprendre les requis moléculaires pour la conjugaison bactérienne *in situ*. Ces travaux ont été publiés dans le nouveau journal *Communication Biology* de la famille Nature et les résultats sont présentés au chapitre 2.

2. Insérer le module CRISPR à un site neutre dans la machinerie de transfert conjugatif.

Maintenant que nous avons identifié TP114 comme une excellente machinerie de transfert conjugatif, la prochaine étape est de lui faire mobiliser un module CRISPR-*cas9*. Afin de maximiser l'efficacité de mobilisation, nous avons opté pour une approche de mobilisation *in cis*. Pour ce faire, le module CRISPR-*cas9* doit être inséré à un endroit neutre sur TP114. Par contre, la taille du module CRISPR-*cas9* est trop importante pour procéder par recombineering de manière efficace. Aussi, le recombineering promeut la recombinaison, ce qui peut être un problème pour exprimer plus d'un ARNg à la fois. Afin de médier la fusion entre TP114 et le module CRISPR, nous avons dû développer une méthode que nous avons nommée Double Recombinase Operated Insertion of DNA (DROID). Les efforts de développement de la méthode DROID ainsi que la fusion entre TP114 et le module CRISPR ont été publiés dans le journal *Plasmid* et font l'objet du chapitre 3.

3. Compléter une preuve de concept chez la souris permettant de quantifier l'efficacité du système COP.

Avec le système complètement assemblé, il était particulièrement important de tester son efficacité. Pour ce faire, nous avons opté pour un modèle permettant à la fois de démontrer l'efficacité du système et sa spécificité. Pour la preuve de concept, les souris recevaient un mélange de bactéries cibles et non-cibles, puis étaient traitées avec le système COP. Le système COP cible ici une résistance au chloramphénicol dans le génome de la bactérie cible. Le système COP a ensuite été modifié par évolution accélérée

afin d'améliorer son efficacité et sa dynamique de conjugaison. Ces travaux sont présentés dans le chapitre 4 de ma thèse et sont soumis à *Nature Medicine*.

L'ensemble de ces efforts de recherche ont permis de faire mûrir la technologie COP jusqu'à l'étape de la preuve de concept. Ces travaux ont d'ailleurs permis le dépôt d'un brevet présentement à l'étape PCT protégeant la technologie COP et permettant ainsi la création d'une entreprise, TATUM Bioscience, ayant pour but la création de médicaments à partir d'organismes vivants.

CHAPITRE 2

HIGHLY EFFICIENT GENE TRANSFER IN THE MOUSE GUT MICROBIOTA IS ENABLED BY THE INC12 CONJUGATIVE PLASMID TP114

2.1 Présentation de l'article et contributions

L'efficacité du système COP présenté ci-haut est directement dépendante de l'efficacité de transfert conjugatif de l'élément génétique mobilisant le système CRISPR-Cas9. La sélection d'un plasmide conjugatif performant dans le tractus gastro-intestinal est donc une étape cruciale afin de produire un effet suffisant sur les résidents du microbiome entérique. Bien que la conjugaison bactérienne soit un phénomène bien caractérisé en laboratoire, peu d'études ont comparé la capacité de transfert conjugatif de différents plasmides dans le système digestif (Schaik, 2015). Un effort de caractérisation des taux de transfert conjugatif de différents plasmides dans un même modèle animal est donc nécessaire afin d'isoler un système performant.

Ce chapitre détaille les efforts de criblage effectués afin d'identifier le plasmide conjugatif le plus performant pour l'utiliser comme machinerie de transfert dans l'outil COP. Pour ce faire, 13 plasmides conjugatifs appartenant à des familles de plasmides variés fréquemment retrouvés chez les *Enterobacteriaceae* ont été sélectionnés. Ces plasmides ont été criblés pour leur efficacité conjugative en laboratoire et dans un modèle de souris que nous avons mis sur pied. De manière inattendue, la moitié des plasmides testés étaient incapables de transférer dans le tractus intestinal. Parmi les plasmides testés, seul TP114 démontrait de hautes efficacités de transfert *in situ*. Des expériences d'analyse de l'essentialité des gènes par insertion de transposon ont ensuite révélé l'essentialité d'un pilus accessoire pour la conjugaison de TP114 dans le milieu intestinal. Ces pili accessoires sont impliquées dans la stabilisation de la paire de conjugaison, un mécanisme que nous avons identifié comme essentiel pour la conjugaison dans l'intestin. Les plasmides incapables de se transférer dans le tractus intestinal étaient également incapables de stabiliser la paire de conjugaison. Par analyse génomique comparative, nous avons ensuite démontré que la majorité des gènes essentiels pour la conjugaison *in situ* chez TP114 étaient partagés chez tous les IncI₂.

Les plasmides de la famille IncI₂ pourraient donc contribuer de façon importante à la mobilité des gènes chez *E. coli* dans le microbiote entérique.

Pour cette étude, j'ai effectué la majorité des expériences et des analyses. J'ai également grandement participé à la rédaction de l'article en collaboration avec Sébastien Rodrigue. Nancy Allard a contribué aux expériences de conjugaison *in situ*, construit et testé le mutant *pilS* en laboratoire et a révisé le manuscrit. Frédéric Grenier a effectué l'analyse bio-informatique de la mutagenèse par transposition et a révisé le manuscrit. Vincent Burrus a aidé à l'acquisition des plasmides conjugatifs, fourni des idées pour les expériences et a révisé le manuscrit. Sébastien Rodrigue a supervisé le projet, aidé à établir le plan des expériences et a participé à la rédaction du manuscrit. J'ai également reçu de l'aide de la part de Dominick Matteau qui m'a aidé à assembler la séquence de TP114.

Référence bibliographique : Neil K., Allard N., Grenier F., Burrus V., Rodrigue S. (2020) Highly efficient gene transfer in the mouse gut microbiota is enabled by the IncI₂ conjugative plasmid TP114. *Commun Biol.* 3(1):523.

2.2 Title Page

Highly efficient gene transfer in the mouse gut microbiota is enabled by the *Incl2* conjugative plasmid TP114

Authors: Kevin Neil¹, Nancy Allard¹, Frédéric Grenier¹, Vincent Burrus¹, and Sébastien Rodrigue^{1*}

Affiliations:

¹ Département de biologie, Université de Sherbrooke, Sherbrooke, Québec, Canada.

*Correspondence to: Sebastien.Rodrigue@USherbrooke.ca.

2.3 Abstract:

The gut microbiota is a suspected hotspot for bacterial conjugation due to its high density and diversity of microorganisms. However, the contribution of different conjugative plasmid families to horizontal gene transfer in this environment remains poorly characterized. Here, we systematically quantified the transfer rates in the mouse intestinal tract for 13 conjugative plasmids encompassing 10 major incompatibility groups. The vast majority of these plasmids were unable to perform conjugation *in situ* or only reached relatively low transfer rates. Surprisingly, IncI₂ conjugative plasmid TP114 was identified as a proficient DNA delivery system in this environment, with the ability to transfer to virtually 100% of the probed recipient bacteria. We also reveal that a type IV pilus present in I-complex conjugative plasmids plays a crucial role for the transfer of TP114 in the mouse intestinal microbiota, most likely by contributing to mating pair stabilization. These results provide new insights on the mobility of genes in the gut microbiota and highlights TP114 as a very efficient DNA delivery system of interest for microbiome editing tools.

2.4 Introduction:

Horizontal gene transfer (HGT) is a major driver of bacterial evolution, allowing several microbial species to exchange genes and adapt to various conditions^{1,2}. One of the most important HGT mechanism is bacterial conjugation, a natural process during which DNA is transferred through a type IV secretion system (T4SS) from a donor to a recipient bacterium in close contact³. This phenomenon has been thoroughly studied *in vitro* using several conjugative elements⁴⁻⁶. However, relatively few studies have investigated the requirements for bacterial conjugation in natural and clinically relevant environments such as the gut microbiota⁷⁻¹⁶. The intestinal tract contains a wide diversity of microbial species that varies throughout the digestive system^{17,18}. Understanding bacterial conjugation in the gut could provide valuable insights on the dissemination of antibiotic resistance genes and the emergence of multi-drug resistant pathogens¹⁹.

Conjugative plasmids have been shown to transfer between different bacterial species residing in the gut microbiota^{7,8} but only a few groups have quantified transfer rates *in situ*^{9-12,20}. Many factors such as NaCl, propionate or butyrate concentrations, the presence of epithelial cells or inflammation have been shown to affect DNA transfer of specific conjugative plasmids in the gut^{8,13-15}. However, these reports only address a single plasmid at a time and used different mouse models (gnotobiotic C3H^{9,20}, streptomycin treated BALB/C^{10,13} or C57 BL/6¹²) and bacterial hosts (*E. coli*¹³, *Salmonella enterica*¹², *Lactococcus lactis*¹⁰, *Enterococcus faecalis*^{9,20}), making comparisons difficult. Furthermore, some of these studies pre-mixed donor and recipient strains before introduction in mice^{12,13}. As conjugation is a relatively fast process, this procedure could introduce artefacts or important biases in the quantification of conjugation rates. A standardized and robust *in situ* conjugation mouse model is therefore needed to systematically study the contribution of different plasmid families to the dissemination of antibiotic resistance genes in the microbiota as well as the mechanisms required for bacterial conjugation *in situ*.

In this study, we investigated 13 plasmids transferring within the *Enterobacteriaceae* group, a phylum infamous for their rapid accumulation of resistance genes to many antibiotics currently on the market^{21,22}. We established a robust mouse model to quantify DNA transfer rates in the gut

microbiota and evaluated the contribution of 13 different conjugative plasmids to horizontal gene transfer in this environment. These experiments identified TP114 as a highly effective machinery for conjugative transfer *in situ*. Transposon insertion mutagenesis allowed the identification of an accessory type IV pilus shared among members of the I-complex family of conjugative plasmids as an essential component for TP114 conjugation in the gut. Taken together, our results suggest that mating pair stabilization is a key mechanism for conjugative transfer in the gut microbiota.

2.5 Results:

2.5.1. *In vitro* transfer rates quantification.

A total of 13 conjugative plasmids isolated from enteric bacteria were selected to represent a large phylogenetic diversity spanning 10 incompatibility (Inc) groups²³. All plasmids were sequenced, annotated and their resistance profiles were assessed (Supplementary Table 2.1). Three derivatives of *Escherichia coli* Nissle 1917 (EcN) with different antibiotic resistance profiles were next generated as standardized hosts for the conjugative plasmids. EcN derivatives were all resistant to streptomycin but additionally displayed resistance to either spectinomycin (KN01), chloramphenicol (KN02) or tetracycline (KN03) to facilitate donor and recipient strain discrimination in conjugation assays (Supplementary Fig. 2.1)²⁴. Each conjugative plasmid was introduced into EcN KN01, which provides a standardized donor strain and limits host related conjugation efficiency bias for all conjugation assays. Transfer rates of all conjugative plasmids were then quantified on agar solid support (Fig. 2.1a) or in broth (Fig. 2.1b) by mixing an equal amount of EcN KN01 donor and EcN KN02 or EcN KN03 recipient bacteria before incubating for 2 hours at 37°C. Transconjugant formation frequencies of the 13 conjugative plasmids varied over 5 orders of magnitude and were in good agreement with published observations, usually with higher transfer rates on solid support^{4,25-27}.

2.5.2. Establishment of a mouse model.

A standardized mouse model was next developed to quantify the transfer rates of conjugative plasmids in the gut microbiota. C57 BL/6 female mice were selected as they were previously used for *in situ* conjugation assays and offer the possibility to access many transgenic strains. Mice were provided with varying concentrations of streptomycin in drinking water to deplete endogenous populations of enterobacteria in their digestive tract and to facilitate the establishment of streptomycin-resistant EcN derivatives (Supplementary Fig. 2.2a-f)²⁸. The addition of 1.0 g/L of streptomycin in mice drinking water reliably increased EcN levels by >100-fold in most sections of the gut as well as in feces (Supplementary Fig. 2.2g). This step was found to be important to achieve reproducible colonization levels by EcN, which are in turn crucial to enable the detection of low-frequency conjugative transfer events. While EcN was detected throughout the entire intestinal tract, colonization levels were higher from the caecum to the colon, with particularly high occurrences in feces for streptomycin treated mice (Supplementary Fig. 2.2g). Stool sampling can thus be used to facilitate manipulation, reduce the number of mice required per experiment, and maximize the sensitivity of conjugation assays.

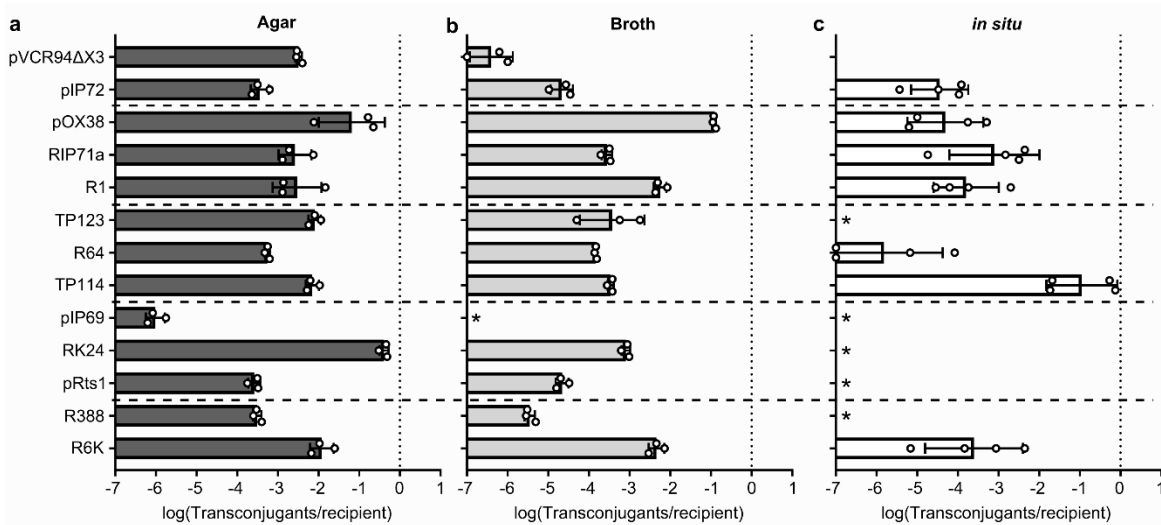


Figure 2.1: Comparison of the conjugation rates of 13 plasmids under different conditions.

A set of 13 conjugative plasmids were tested for their ability to transfer between two EcN cells during 2 hours on agar (n=3, biologically independent samples) (a), in broth (n=3, biologically

independent samples) (b) and for 24 hours *in situ* (n=4, biologically independent mice) (c) in a murine gut conjugation model. The dotted vertical line highlights a conjugation rate that allows all recipient cells to receive a plasmid. Asterisks denote the absence of detectable transconjugant colonies. White circles show individual values for each dataset. Bar graphs represent the average and error bars the standard deviation of the data set.

2.5.3. Quantification of *in situ* DNA transfer rates.

In situ transfer rates of the 13 selected conjugative plasmids were next evaluated in the murine gut using 2 consecutive gavages. The EcN recipient strain (KN02 or KN03) was first administered before allowing mice to rest for 2 hours. A second gavage was then performed with the donor strain EcN KN01 carrying one of the 13 conjugative plasmids. Conjugation events were quantified by counting colony forming units (CFU) per gram of feces collected daily for three days after introduction of the EcN KN01 donor strain (Fig. 2.1c and Supplementary Fig. 2.3). No detectable levels of *in situ* transfer were observed for 6 out of 13 conjugative plasmids. The remaining 7 plasmids showed transfer rates ranging over ~5 orders of magnitude (Fig. 2.1c). No correlation was observed between *in situ* transfer rates and those measured on agar ($R^2=0.0687$) or in broth ($R^2=0.2727$). However, we noticed that the frequency of transconjugants detected at day 4 in the caecum were generally consistent with those obtained from feces ($R^2=0.701$), suggesting that stool sampling provided an appropriate proxy for *in situ* quantification of conjugation (Fig. 2.2a). A few exceptions were noted, for example with R6K, R1 and pOX38 that produced a relatively low number of transconjugants exclusively or predominantly in the caecum. Given that the density of *E. coli* per mg of material is similar between feces and the caecum of streptomycin-treated mice (Supplementary Fig. 2.2g), this phenomenon was likely due to the greater sensitivity of the assay in the caecum from which ~10-fold more material can be analyzed, hence facilitating the detection of low frequency transfer events.

2.5.4. Highly effective transfer of TP114 *in situ*.

Of all tested plasmids, IncI₂ plasmid TP114 clearly displayed the highest transfer rate *in situ*, showing dissemination to a very large proportion of probed EcN recipient bacteria within 24 hours (Fig. 2.1c), which later virtually reached 100% (Fig. 2.2a and 2.2b). Conjugation of TP114 also showed very similar efficiencies between feces and the caecum (Fig. 2.2a). A few parameters of the mouse model were next modified to investigate their potential impact on TP114 transfer rates. The effect of introducing the donor bacterium EcN KN01 12 hours (rather than 2 hours) after the introduction of the recipient strain was first evaluated, without any clear difference on conjugation rates of TP114 (Fig. 2.2b) or colonization levels of EcN KN02 (Fig. 2.2c). The impact of the streptomycin treatment on TP114 transfer rates was also investigated, but no significant difference was observed (p value 0.1681) despite lower colonization levels by the donor and recipient bacteria (Fig. 2.2d). Taken together, these experiments demonstrated the robustness and reproducibility of plasmid TP114 high transfer rates *in situ*.

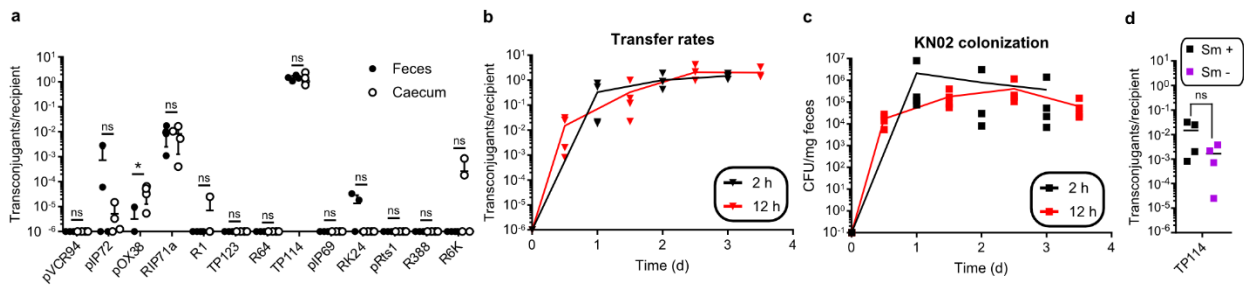


Figure 2.2: Evaluation of mouse model parameters for *in situ* conjugation. **a**, Comparison of the transfer rates observed in the caecum and feces 4 days after the introduction of the donor strain harbouring each of the 13 conjugative plasmids (n=4, biologically independent mice). Error bars span the minimum and maximum observed. Effect of time between gavages on TP114 transfer rates (n=4, biologically independent mice) (**b**) as well as on recipient colonization level (n=4, biologically independent mice) (**c**). **d**, Conjugation rates after 12 hours in mice treated or not with 1000 mg/L of streptomycin (n=4, biologically independent mice). Panels **a** and **d**, One-way ANOVA statistical analysis results are shown as: ns = p>0.05, * = p<0.05. **a-d**, lines represent the average of the data.

2.5.5. High-density transposon mutagenesis (HDTM) of TP114.

Deciphering the molecular mechanisms used by the highly efficient plasmid TP114 could provide valuable insights on the strategies used by conjugative plasmids to disseminate within the gut microbiota. Using a high-density transposon mutagenesis (HDTM) approach, TP114 mutant libraries comprising on average one transposon insertion every ~5 bp were generated (Supplementary Table 2.2). The HDTM library was first analyzed in EcN, in the absence of any recipient strain, to identify the key elements for TP114 maintenance. In this context, transposon insertion sites were found to be uniformly distributed across the TP114 sequence, sparing only the loci encoding replication, maintenance or selection markers such as the replication initiation protein (*repA*) and the kanamycin resistance gene (*aph-III*) (Fig. 2.3a yellow boxes track 1). A conjugation assay on agar solid support was then performed between the initial HDTM library and the EcN KN02 recipient strain (Fig. 2.3a, track 2 and Supplementary Fig. 2.4a). The resulting transconjugants were also used for an additional round of transfer on agar support (Fig. 2.3a, track 3 and Supplementary Fig. 2.4a). HDTM biological triplicates were found to be highly correlated (Supplementary Fig. 2.4b). Gene essentiality after conjugative transfer was investigated using a ratio of the normalized transposon insertion read counts of transconjugants compared to the initial mutant library (Supplementary Fig. 2.4cdef). As expected, the T4SS and relaxosome encoding genes, which were already known to be essential for transfer in conjugative plasmids, were also deemed essential for TP114 to perform conjugation *in vitro* (Fig. 2.3a blue boxes track 2 and 3).

in EcN KN02 obtained after a single *in vitro* transfer, were used as donors for *in situ* transfer assays as described above (Supplementary Fig. 2.4a). The resulting transconjugants were isolated from feces (Fig. 2.3a red boxes track 4 and 6) or from the caecum (Fig. 2.3a red boxes track 5 and 7), and their transposon insertion sites were sequenced. Interestingly, *in situ* conjugation required genes predicted to encode a Type IV Pilus (T4P) in addition to genes that were essential for transfer on agar solid support. T4P generally comprise a machinery responsible for the secretion of a major and a minor pilin, which were previously reported to be involved in Mating Pair Stabilization (MPS)²⁹. The role of the T4P was directly tested in TP114 using a deletion mutant of the predicted major pilin *pilS*. While no difference in conjugative rates was observed when using a solid support, TP114 Δ *pilS* transfer was completely abolished in broth, a phenotype that could be partially complemented using a PilS expressing plasmid (Fig. 2.3b)²⁹. Using the *pilS* mutant for *in situ* conjugation yielded >1,000-fold fewer transconjugants (Fig. 2.3c), with transfer rates ($\leq 10^{-4}$) consistent between the feces and the caecum (Fig. 2.3d). These results confirmed the critical role of the T4P for *in situ* conjugation of TP114 and suggest that MPS is crucial for conjugation in the mouse gut microbiota.

2.5.7. Comparative genomics in IncI plasmids.

Comparative genomics was next performed to investigate if the ability of TP114 to transfer at high rates in the gut microbiota was likely to be shared among other members of the IncI family. Most genes encoded by TP114 (62 out of 92), including those encoding the T4P (10 out of 10), displayed high homology (>99% similarity in nucleic acids and in protein sequences) with 7 selected plasmids of the IncI₂ family (Supplementary Fig. 2.5ab). TP114 genes were also categorized into core, soft-core or accessory gene groups based on their distribution across IncI₂ plasmids (Supplementary Data 2.1). Comparing gene conservation and essentiality in TP114 revealed that 36 out of the 41 genes required for *in situ* conjugation were also part of the core genome of IncI₂ plasmids (Supplementary Fig. 2.6a, Supplementary Data 2.1), supporting the idea that most IncI₂ plasmids are likely to disseminate at high rates in the intestinal tract.

In contrast, TP114 shared very little sequence homology with plasmids of the IncI₁ group although these plasmids also encode a T4P. Sequence similarity between IncI₁ and IncI₂ was in fact limited almost exclusively to the *oriV* region and *pilV* gene, which encodes the minor pilin subunit (Supplementary Fig. 2.5cd). However, while conjugation of the IncI₁ plasmid R64 did not reach high levels (average of 1.54×10^{-6} transconjugants/recipient) in our model, IncI₁ plasmid p2 was shown to transfer at high rates *in situ* from *Salmonella typhimurium* to *E. coli*^{12,15}. Both R64 and p2 were originally isolated from *Salmonella sp.* raising the possibility that host cell compatibility could influence conjugation.

2.5.8. Activation of TP114 under anaerobic conditions.

Of all 13 tested conjugative plasmids, TP114 presented a striking example of higher *in situ* conjugation relative to *in vitro* conditions (Fig. 2.4a), which led to the hypothesis that TP114 was activated by environmental conditions found in the intestinal tract. We sought to identify the potential signal that could lead to increased conjugation in the mouse intestinal tract. Conjugation assays were thus performed at different temperatures (23°C and 30°C), in presence of mouse feces, and under hypoxic conditions and normalized to transfer rates obtained on LB agar at 37°C (Fig. 2.4b). Of these, hypoxic conditions improved the transfer rates of TP114 by ~10-fold, suggesting that TP114 responds to low oxygen concentrations found in the intestinal tract. These results also indicate that the overall performance of conjugative plasmids in the gut depends, not only on MPS, but also on appropriate regulatory responses to activate plasmid transfer in the gut.

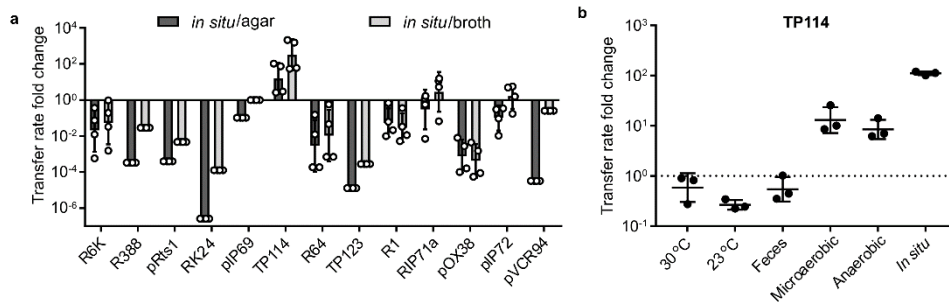


Figure 2.4: TP114 transfer is activated under hypoxic conditions. a, ratio between the transfer rate of the 13 conjugative plasmids *in situ* after 24 hours (n=4, biologically independent mice) and

their transfer efficiency on agar (n=3, biologically independent samples) (dark gray) or in broth (light gray) (n=3). Since R388, pRts1, RK24, pIP69, TP123 and pVCR94 did not produce any transconjugant *in situ*, their transfer rate fold change was calculated using the experimental limit of detection. Individual data points are shown as white circles on graph. **b**, influence of the temperature, presence of feces and hypoxic environment on the transfer rates of TP114 in relation to transfer rates for 2 hours at 37°C on agar (n=3, biologically independent samples). **a and b**, error bars represent the standard deviation of each dataset.

2.6 Discussion:

Conjugative plasmids play a major role in the dissemination of antibiotic resistance. However, few studies have systematically investigated the capacity of different families of conjugative plasmid to transfer within the gut microbiota, where they could contribute to the emergence of multi-drug resistant pathogens. Using a standardized mouse model for *in situ* conjugation assays, our study identified IncI₂ plasmid TP114 as a highly effective conjugation machinery. In addition, transposon mutagenesis revealed that a T4P was essential for this function, most likely by stabilizing the interaction between the donor and recipient bacteria. With the exception of TP114, only six other plasmids (R64, pIP72, RIP71a, R1, pOX38 and R6K) were able to transfer in our *in situ* conjugation assays. Both IncI₁ plasmid R64 and IncB plasmid pIP72 encode T4P albeit without any significant sequence homology with TP114. Conjugative plasmids RIP71a, R1 as well as pOX38 all belong to the IncF group and rely on TraN, an adhesin expressed at the surface of the donor host cell, to stabilize the mating pair by binding to OmpA or LPS³⁰. While no gene involved in MPS has been identified yet in R6K, this plasmid was reported to transfer with almost the same efficiency on agar or in broth³¹ (Fig. 2.1a and 2.1b). Since MPS is required for efficient conjugation in broth, it is likely that R6K has the capacity to stabilize the interaction between its host and a recipient bacterium although the exact genes responsible for this function remain to be identified. In fact, all plasmids that transferred *in situ* were also able to yield transconjugants in broth with relatively high efficiencies relative to agar solid support (3.7±4.0-fold lower transfer rates in broth relative to solid support). Conversely, conjugative plasmids that did not generate transconjugants *in situ* transferred at lower rates in broth relative to solid support (156.7±10.6-fold lower transfer rates in broth relative to solid support) suggesting that they could not perform MPS

in the context of our assay. A clear example is IncP1 α plasmid RK24, which had the highest transfer rate *in vitro* on agar support. However, RK24 lacks MPS genes³², transfers ~500-fold less efficiently in broth relative to agar support, and had virtually no conjugation activity *in situ*. These observations further support the importance of MPS for plasmid conjugation in the intestinal microbiota.

Plasmids that failed to produce transconjugants in our *in situ* assays could possibly encode genes that bear no homology to known MPS systems, and use them only in other contexts. For example, their adhesins may recognize structures found on other bacterial species that are not displayed on the EcN recipients. MPS would hence play a very important role in determining plasmid transfer range towards specific species or strains in certain unstable environments. *In silico* analyses were performed to identify putative MPS genes in the six conjugative plasmids that had no detectable transfer in the mouse intestinal tract, and no putative adhesin-related protein domains could be found in RK24, R388 or pIP69. However, possible orthologs of TraN were identified in pVCR94, pRts1 and TP123. The protein sequence of these orthologs were found to diverge from the TraN present in IncF plasmids, suggesting that their role or specificity could differ (Supplementary Fig. 2.6b). It is also possible that the TraN variants of pVCR94, pRts1 and TP123 interact more strongly with other bacterial species that are potentially more closely related to their original hosts (Supplementary Table 2.1) or simply lack the capacity to perform MPS. Another factor that might influence conjugation rates is the availability of commensal recipient bacteria that could sequester broad-host-range plasmids and limit their ability to be transferred to our introduced EcN recipient strain. However, the transfer range of most conjugative plasmids is still largely unknown, making this hypothesis difficult to assert at this point.

The ability of conjugative plasmids to transfer *in situ* involves MPS but also requires other specialized machineries such as the T4SS and the relaxosome. However, using computational analyses to associate each plasmid to a specific incompatibility (Inc; replication initiation protein (RepA)), mobilization machinery (MOB; relaxase), or mating pore formation (MPF; T4SS VirB4 subunit) group showed that diverse plasmid phylogenetic families can support conjugative transfer

in the mouse intestinal tract (Supplementary Table 2.1). This observation implies that *in situ* conjugation is not restricted to a specific type of T4SS or plasmid family.

Conjugation in the gut microbiota was initially thought to occur in a relatively stable matrix akin to an agar plate¹⁶. This conclusion was based on experiments with IncF plasmid R1 in which transfer frequencies were compared *in situ* and *in vitro* in broth or on agar. Based on these results, it was suggested that bacterial conjugation occurs most frequently in the mucus layer of the intestinal lumen after biofilm formation¹⁶. However, several factors can influence conjugation rates *in situ*. We identified MPS as an important molecular mechanism for conjugation in the gut at early stages of colonization. This suggests that DNA transfer occurs in a rather unstable environment. The mouse model used in this study most likely represents a situation in which an incoming bacterium is introduced in the gut microbiota and can disseminate plasmids in the microbiota. Plasmids that were unable to transfer *in situ* in our assays could potentially transfer at later stages of colonization, for instance when embedded in dense biofilms. RK24 is an interesting example as this plasmid was previously used as a mobilization machinery for microbiome editing, and was shown to produce interspecies transconjugants^{7,33}. However, conjugation rates were not reported⁷ or shown to be very low³³, which is consistent with our results.

IncI₂ plasmids likely share the ability to transfer DNA at high frequencies in the gut. Surveillance of plasmids from this family could provide insights on the likeliness of antibiotic resistance gene dissemination in animals and humans. As several plasmids are often found in a single host³⁴, investigating the interactions between different conjugative plasmid families *in situ* could provide further insights on the mobility of resistance genes. For example, it is possible that mobile genetic elements unable to perform MPS *in situ* could take advantage of the MPS mechanism of other plasmids to transfer more efficiently in the gut. If so, interfering with MPS could become an interesting strategy to prevent dissemination of antibiotic resistance in infected patients and animal husbandry.

Given that the microbiota plays major roles in human health, different strategies to correct imbalanced microbiome compositions have been recently proposed. For example, microbiota-

editing strategies could rely on bacteriophages or conjugative plasmids to disseminate CRISPR or toxin genes to kill specific bacteria^{33,35-37}. However, these studies often highlight the lack of efficient conjugative plasmid for DNA delivery^{33,35}. Highly effective conjugative systems such as TP114 could thus become important DNA delivery tools in the development of microbiome editing technologies.

2.7 Methods:

2.7.1. Strains, plasmids and growth conditions.

Strains and plasmids are described in Supplementary Table 2.3 and are available upon request. Cells were typically grown in Luria broth Miller (LB) or on LB agar medium supplemented, when needed, with antibiotics at the following working concentrations: ampicillin 100 µg/mL, chloramphenicol 34 µg/mL, kanamycin 50 µg/mL, nalidixic acid 4 µg/mL, rifampicin 66.7 µg/mL spectinomycin 100 µg/mL, Sm 50 µg/mL, sulfamethoxazole 160 µg/mL, tetracycline 15 µg/mL, and trimethoprim 32 µg/mL. Diaminopimelic acid (DAP) auxotrophy was complemented by adding DAP at a final concentration of 57 µg/mL in the medium. All cultures were routinely grown at 37°C. Cells with thermosensitive plasmids (pSIM6, pE-FLP, pGRG36, pFG036) were grown at 30°C. Bacterial cultures were grown for no more than 18 hours.

2.7.2. DNA manipulations.

A detailed list of oligonucleotide sequences is found in Supplementary Table 2.3. Cloning vectors were prepared using EZ10-Spin Column Plasmid Miniprep kit (Bio Basic) whereas genomic DNA (gDNA) and conjugative plasmids were extracted using the Quick-gDNA miniprep (Zymo Research) according to the manufacturer's instructions. PCR amplifications were performed using Veraseq DNA polymerase (Enzymatics) or TaqB (Enzymatics) for DNA parts amplification and screening, respectively. Digestion with restriction enzymes (NEB) were incubated for 1 hour at 37°C following manufacturer's recommendations. Plasmids were constructed by Gibson assembly³⁸ using the NEBuilder Gibson Assembly mix (NEB) following manufacturer's protocol.

These procedures were applied to the construction of pPiS that are further detailed in the Supplementary methods section.

2.7.3. DNA purification.

Purification of DNA was performed between each step of plasmid assembly to prevent buffer incompatibility or interference with enzymatic reactions. For general purposes, DNA was purified by Solid Phase Reversible Immobilization (SPRI) using HiPrep PCR DNA binding beads (MagBio). For sequencing, DNA was purified using Agencourt Ampure DNA binding beads (Beckman Coulter). Protocols were identical for both types of beads and followed manufacturer's guidelines. For restriction enzyme digestions, DNA was purified using DNA Clean and Concentrator (Zymo Research) following manufacturer's recommendation to prevent buffer incompatibility between the digestion's CutSmart buffer (NEB) and the polyethylene glycol-based buffer of the SPRI beads, which greatly decreased DNA yield during purification. After purification, DNA concentration and purity were routinely assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific).

2.7.4. DNA transformation into *E. coli* by electroporation.

Routine plasmid transformations were performed by electroporation. Electrocompetent *E. coli* strains were prepared by sub-culturing in 20 mL of LB broth. Cultures reaching an optical density of 0.6 at 600 nanometers (OD_{600nm}) were washed three times in sterile distilled water. Cells were then resuspended in 200 μ L of water and distributed in 40 μ L aliquots. The DNA was then added to the electrocompetent cells and the mixture was transferred in a 1 mm electroporation cuvette. Cells were electroporated with a pulse of 1.8 kV, 25 μ F, and 200 Ω for 50 ms. Cells were then resuspended in 1 mL of non-selective LB medium to recover for 1 hour at 37°C or 30°C (for thermosensitive plasmids) before plating on selective media.

2.7.5. DNA transformation into *E. coli* by heat-shock.

Heat-shock transformation was mostly used for Gibson assembly products transformation. Chemically competent cells were prepared according to the rubidium chloride protocol as described previously³⁹. Chemically competent cells were flash-frozen and conserved at -80°C before use. Routinely, up to 10 µL of DNA was added to 100 µL of EC100D*pir*⁺ chemically competent cells before transformation with a 45 seconds heat-shock at 42°C. Cells were then resuspended in 1 mL of non-selective LB medium to recover for 1 hour at 30°C or 37°C before plating on selective media.

2.7.6. Recombineering.

All recombineering experiments were performed using pSIM6 as described previously^{40,41}. Briefly, the recombineering cassette was electroporated in an induced *E. coli* + pSIM6 strain. Then, cells were left to recover overnight at room temperature before plating on selective medium. Single isolated colonies were stabbed in a new plate before homogenization in 100 µL of 5% w/v Chelex solution (Bio-Rad). Afterwards, the Chelex mixture was heated to 56°C for 25 minutes and 100°C for 10 minutes in a PCR machine for DNA extraction. For PCR screening, 1 µL of freshly extracted DNA was added to the PCR mix. Positive clones were sub-cultured in 5 mL of selective LB broth overnight and frozen in 25% glycerol for storage. This protocol was applied to the deletion of the *pilS* in TP114 and the deletion of *dapA* in the chromosome of EcN, which are further described in the Supplementary methods section.

2.7.7. Generation of modified EcN strains.

The modified EcN strains were generated by Tn7 insertion of the antibiotic resistance cassettes as described previously²⁴ with minor adaptations. More specifically, the pGRG36 vector was purified from *E. coli* EC100D*pir*⁺ and digested using SmaI and XhoI. The inserts were amplified by PCR using their corresponding primers (Supplementary Table 2.3) and cloned by Gibson assembly (NEBuilder Gibson Assembly kit) between *attL*_{Tn7} and *attR*_{Tn7} sites of the digested pGRG36

plasmid (Supplementary Fig. 2.1abc) following manufacturer's recommendations. The Gibson assembly products were then transformed in a chemically competent *E. coli* EC100Dpir+. Integrity of the constructions was confirmed by restriction enzyme digestions and positive clones were transformed into *E. coli* MFDpir+⁴². Plasmid pGRG36 derivatives were then mobilized towards EcN by the RK24 conjugative machinery of the MFDpir+ strain. To mediate cassette insertion into the terminator of *glmS*, EcN was first grown at 30°C in LB with 1% arabinose until 0.6 OD_{600nm}. Then, cells were heat-shocked at 42°C for 1 hour and incubated at 37°C overnight to allow for plasmid clearance. Cells were then streaked onto an LB agar plate selecting only the insert. Plasmid elimination was tested by streaking ≥ 20 colonies on both plates with or without ampicillin, which selected for the pGRG36 backbone. Colonies that only grew in absence of ampicillin but contained the insert selection markers were then confirmed for insertion using the corresponding primers from Supplementary Table 2.3.

2.7.8. In vitro conjugation assays.

All *in vitro* conjugation assays used derivatives of KN01 Δ dapA as the donor strain and KN02 or KN03 as the recipient strain depending on the antibiotic markers carried by the plasmid tested. Both strains were sub-cultured from frozen stocks up to 18 hours prior to conjugation experiments and were then mixed at a 1:1 volume ratio (100 μ L each), centrifuged at 20,000 x g for one minute and washed in 200 μ L of LB without antibiotics. The cell mix was then centrifuged again and either resuspended in 5 μ L of LB broth and deposited on a LB agar plate with DAP or resuspended to 1.0 OD_{600nm} in LB broth with DAP. The cell mixture was then incubated at 37°C for 2 hours unless specified otherwise. Conjugation experiments in shaking broths were performed on a tube tumbler rotating mixer (CAT# SBS550-2) at 20 rpm while conjugation in static broth were performed in microtubes on a tube-holder. After incubation, cells were resuspended in sterile PBS, diluted 1/10 serially and 5 μ L of each dilution were spotted in triplicates on LB plate with antibiotics selecting donors, recipients, or transconjugants. All conjugation frequencies were reported as a factor of the recipient CFUs. The conjugation frequencies calculated on a "per donor" basis were equivalent since cells were mixed 1:1. For experiments involving variation in temperature during conjugation, the mix of donor and recipient strains were incubated either at 23°C, 30°C or 37°C for 2 hours before CFU assessment. For experiments with feces in the

conjugative mix, feces were obtained from healthy C57 BL/6 mice, which received no treatments. Feces were then homogenized following procedures described in the Feces sampling section of the Supplementary Method section. Then, feces were transferred directly in the conjugation mix and incubated for 2 hours at 37°C before plating for CFU. No colonies that could have been erroneously interpreted as transconjugants were detected when feces were plated on selective medium containing appropriate antibiotics. Conjugation under hypoxic condition were carried using Gaspack anaerobe (BD 260001) for anaerobic condition and Gaspack CO₂ (BD 260679) for microaerobic condition. Conjugation were incubated in an airtight bag for 24 hours to allow for the creation of hypoxic conditions. All conjugation experiments were repeated in at least three independent biological replicates.

2.7.9. Mouse Model.

All mice-related protocols followed the *Université de Sherbrooke Animal Care Comity Guidelines* and were strictly evaluated to avoid animal suffering. Animals were provided with water and standard chow (Charles River) *ad libitum* during the experiments. Animals were housed in individually ventilated cages and no more than 5 individuals shared the same cage. All animals used were C57 BL/6 females of 16-20 g (Charles River) and were given ≥ 3 -days rest upon arrival. Animal weight and health was evaluated daily, and no notable health or weight loss was noted for all mice throughout the experiments. For Sm-treated mice, 1000 mg/L of Sm was added to drinking water 2 days prior to gavages. From that point, water bottles were changed every 3 days to maintain Sm efficiency. The bacterial strains were prepared as described in the mice inoculum preparation section of the Supplementary Method section. Briefly, for each time a bacterial strain was introduced in mice, approximately 1×10^8 viable cells (quantified by CFU) in PBS were administered by gavage to each mouse using a curved 1.5-inch 20g ball-tip needle. Bacterial presence was monitored in feces using CFU counts at specified time points during the experiments. Further details on the preparation of the inoculums needed for mice gavages as well as feces sampling and mice dissection procedures are available in the Supplementary Methods section.

2.7.10. In situ conjugation assays.

For *in situ* conjugation experiments, mice received the recipient strain 2 or 12 hours prior to the introduction of donor strain to avoid possible plasmid transfer in the PBS solution prior to gavage. Conjugation was then monitored in feces samples at specified time points. Mice were sacrificed at the end of the experiment and the caecum was extracted to measure conjugation levels in the murine gut. Feces were homogenized and CFUs were counted on MacConkey plates as described in the *Feces sampling* section.

2.7.11. DNA sequencing.

Illumina sequencing of conjugative plasmids and of HDTM mutants was performed at the *Plateforme RNomics* (<https://rnomics.med.usherbrooke.ca/>) of the *Université de Sherbrooke*. Oxford Nanopore sequencing was also performed on TP114 to help resolve regions difficult to assemble such as the shufflon. Additionally, Sanger sequencing was performed by the *Plateforme de séquençage et de phénotypage* (<http://www.sequences.crchul.ulaval.ca/>) of the *Université Laval* on TP114 to confirm its sequence and fill gaps. Specific details on the sequencing and assembly of conjugative plasmids and of the HDTM experiments can be found in the Supplementary Methods section.

2.7.12. Analysis of plasmid gene function and generation of phylogenetic trees.

In silico analysis of plasmid gene function was performed using both CDsearch⁴³ and BLASTp v2.8.1⁴⁴. A protein multi-fasta file was first generated for all Open Reading Frames (ORF) predicted by RAST⁴⁵. Using CDSearch, TP114 genes with known protein domains or superfamilies were attributed functions. For genes in which CDsearch failed to identify any protein domain with high confidence (e-value $<1 \times 10^{-15}$), their corresponding protein sequence file was also submitted to BLAST v2.8.1 to identify putative protein homologs. Analyses were performed using default parameters. BLAST hits with high identity levels (>98% identity on >95% query coverage) were used to attribute putative functions only when more than five hits showed

the same result. Proteins that failed to match these criteria were considered of unknown function. The protein homologs of VirB4, the relaxase and the replication initiation protein were identified following the same procedures described above for all 13 conjugative plasmids and were next aligned with each other using ClustalW and inputted in MEGA-X v10.0.5⁴⁶ for the generation of neighbour-joining phylogenetic trees with 1,000 bootstrapping replicates.

2.7.13. High-density transposon mutagenesis.

A conjugation-assisted random transposon mutagenesis experiment was performed. The transposition system was composed of pFG036 (a plasmid coding for a cI transcription repressor) and pFG051 (a Pi-dependent suicide plasmid coding for the Tn5 transposon machinery under the repression of cI, a RK24-based origin of transfer and a spectinomycin resistance transposon) cloned into *E. coli* MFD pir^+ ⁴². This DAP-auxotrophic strain contains an RK24 conjugative machinery integrated in its genome and expresses the Pi protein required for pFG051 replication. The HDTM experiment was performed in several successive steps to clearly identify the function of genes involved at each one of these steps. First, pFG051 was transferred by conjugation from MFD pir^+ to EcN containing TP114 for 2 hours at 30°C on LB + DAP plates in triplicates. Once in EcN, the Tn5 machinery was constitutively expressed from pFG051 to mediate random transposon insertions in TP114. Then, transconjugants were entirely spread onto 6 plates per replicate and incubated overnight at 37°C to form a cell lawn. After the incubation, cells were collected using a cell scraper and subsequently resuspended in LB broth with selective antibiotics. A 100 µL aliquot of the transconjugants was used in two subsequent conjugative transfer experiments performed *in vitro* and *in situ* using KN02 and next KN03 as recipients. For each of these experiments, transconjugants forming the mutant library were washed, resuspended in 4.5 mL of LB + 25% glycerol and cryopreserved prior to DNA extraction and sequencing.

2.7.14. Mouse model for *in situ* HDTM library conjugation.

A cryopreserved 500 µL aliquot of the HDTM mutant library was inoculated in 20 mL selective LB broth and incubated at 37°C for 4 hours before gavage. The recipient strains were prepared as

described in the *in situ* conjugation section. Briefly, frozen stocks were streaked on MacConkey for overnight growth before sub-culturing in 5 mL LB. Then, the recipient strains were sub-cultured again in 20 mL LB 2 hours prior to gavage. When ready, cells were washed once in PBS and concentrated in a volume equivalent to 6.0 OD_{600nm}. Mice were administered the recipient strain 3 hours prior to the introduction of the donor strain. Conjugation was then monitored after feces sampling at 24 and 48 hours. At 48 hours, mice were sacrificed, and the caecum was extracted. At each time-point, 4 x 100 µL of sample per mice were also plated to maximize the number of transconjugants for sequencing.

2.7.15. HDTM libraries analysis.

Reads were first trimmed based on their quality and the presence of the Nextera Illumina adapter using Trimmomatic v0.32⁴⁷ with the parameters SLIDINGWINDOW:4:20 and MINLEN:30. The quality of the reads, before and after trimming, was assessed with FastQC v0.11.4 using the default parameters⁴⁸. Reads mapping on the EcN chromosome were filtered out and the remaining reads were mapped onto TP114. These alignments were done with BWA MEM v0.7.2 using the default parameters⁴⁹. Alignments with a mapping quality score lower than 30 were discarded. The position of the middle base pair of the 9-bp Tn5 insertion site duplication was then used to represent every corresponding alignment⁵⁰. Insertion sites represented by only one read were discarded to filter out sequencing noise. The insertion maps with normalized read counts (based on samples sequencing depths) were then visualized using UCSC Genome Browser in a Box v1.12⁵¹. The essentiality of the genes in condition 1 (initial library) was verified manually, searching for low coverage regions that were mappable and reproducible in all three replicates. A read count table was then generated by calculating the normalized read counts per gene for each condition. Insertion sites in the first 5% and last 15% of the gene were discarded from the read count as they may lead to functional gene fragments. The genes important for *in vitro* and *in situ* conjugation were determined based on the normalized gene read count ratio between condition 1 and the test condition in the following manner: (Normalized Read count x – Normalized Read count 1)/Normalized Read count 1. A core set of genes visually identified in UCSC Genome Browser as essential for conjugation *in vitro* (*traABCDEFGHIJK*, *trbJ*, *nikAB*) and *in situ* (*pilNOPQRSUV*)

was then used to set the maximal ratio value for each condition. All genes with gene count ratios below the maximal value were considered essential in the given condition.

2.7.16. Comparative genomics using BRIGG.

Gene content comparison was performed on TP114 against a database of 7 randomly selected plasmids of the IncI₁ and IncI₂ subfamilies. The BRIGG stand-alone software⁵² was used to perform BLAST-based homology analysis between TP114 and each plasmid group. Homology was analysed using both the nucleotide sequence of the whole plasmids and amino-acid sequences of the coding genes. Conservation of genes was evaluated using the sequence identity cut-offs of 100%, 70%, and 50%. The identity percentage was calculated by attributing scores of -2 for mismatches, +1 for matches and a linear cost for insertion/deletion. Genes were then categorized as core genes when present in 100% of the plasmids, soft core genes when present in more than 50% of the plasmids, or accessory genes when present in less than 50% of the plasmids.

2.7.17. Statistics and reproducibility.

Statistical significance was performed on the logarithmic value of the data using One-way ANOVA unless specified otherwise. P-values are directly indicated on the graphs and represent statistical significance of the difference between the two data groups. Differences in the data were considered significant when the P-value was below 0.05. When P-value are not noted directly on graph, statistical significance was noted as: ns $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Number of replicates is specified in “n=” statements in figure legends. All *in vitro* experiments were performed in biological triplicates using three independently grown cultures. For experiments involving mice, a minimum of 4 C57 BL/6 female mice were used for each sampling.

2.7.18. Data availability.

The sequence of conjugative plasmids is available in GenBank (see Supplementary Table 2.1 for accession numbers). Transposon insertion sites for the HDTM experiments (Fig. 2.2) are available

on UCSC Genome Browser under http://genome.ucsc.edu/cgi-bin/hgTracks?genome=TP114&hubUrl=https://datahub-103-cu2.p.genap.ca/Neil_2020/hub.txt.

Reads for the sequencing of TP114, other conjugative plasmids and HDTM experiments are available at NCBI under BioProjects: PRJNA510561, PRJNA576482 and PRJNA510811 respectively.

2.8 References:

1. Soucy, S. M., Huang, J. & Gogarten, J. P. Horizontal Gene Transfer: Building the Web of Life. *Nat. Rev. Genet.* **16**, 472–482 (2015).
2. Thomas, C. M. & Nielsen, K. M. Mechanisms of and Barriers to, Horizontal Gene Transfer between Bacteria. *Nat. Rev. Microbiol.* **3**, 711–721 (2005).
3. Hospenthal, M. K., Costa, T. R. D. & Waksman, G. A Comprehensive Guide to Pilus Biogenesis in Gram-negative Bacteria. *Nat. Rev. Microbiol.* **15**, 365–279 (2017).
4. Bradley, D., Taylor, D. E. & Cohen, D. R. Specification of Surface Mating Systems Among Conjugative Drug Resistance Plasmids in *Escherichia coli* K-12. *J. Bacteriol.* **143**, 1466–1470 (1980).
5. Arutyunov, D. & Frost, L. S. F conjugation: Back to the beginning. *Plasmid* **70**, 18–32 (2013).
6. Smillie, C., Garcillan-Barcia, M. P., Francia, M. V., Rocha, E. P. C. & de la Cruz, F. Mobility of Plasmids. *Microbiol. Mol. Biol. Rev.* **74**, 434–452 (2010).
7. Ronda, C., Chen, S. P., Cabral, V., Yaung, S. J. & Wang, H. H. Metagenomic Engineering of the Mammalian Gut Microbiome *in situ*. *Nat. Methods* **16**, 167–170 (2019).
8. Aviv, G., Rahav, G. & Gal-mor, O. Horizontal Transfer of the *Salmonella enterica* serovar infantis Resistance and Virulence Plasmid pESI to the Gut Microbiota of. *Am. Soc. Microbiol.* **7**, 1–12 (2016).
9. Doucet-Populaire, F., Trieuquot, P., Andremont, A. & Courvalin, P. Conjugal Transfer of Plasmid DNA from *Enterococcus faecalis* to *Escherichia coli* in Digestive Tracts of Gnotobiotic Mice. *Antimicrob. Agents Chemother.* **36**, 502–504 (1992).
10. Igimi, S. *et al.* Transfer of conjugative plasmid pAM β 1 from *Lactococcus lactis* to mouse intestinal bacteria. *Lett. Appl. Microbiol.* **23**, 31–35 (1996).

11. Gumpert, H. *et al.* Transfer and Persistence of a Multi-Drug Resistance Plasmid *in situ* of the Infant Gut Microbiota in the Absence of Antibiotic Treatment. *Front. Microbiol.* **8**, 1–10 (2017).
12. Stecher, B. *et al.* Gut Inflammation can Boost Horizontal Gene Transfer between Pathogenic and Commensal *Enterobacteriaceae*. *Proc. Natl. Acad. Sci.* **109**, 1269–1274 (2012).
13. Machado, A. M. D. & Sommer, M. O. A. Human Intestinal Cells Modulate Conjugational Transfer of Multidrug Resistance Plasmids between Clinical *Escherichia coli* Isolates. *PLoS One* **9**, e100739 (2014).
14. García-Quintanilla, M., Ramos-Morales, F. & Casadesús, J. Conjugal Transfer of the *Salmonella enterica* Virulence Plasmid in the Mouse Intestine. *J. Bacteriol.* **190**, 1922–1927 (2008).
15. Bakkeren, E. *et al.* *Salmonella* Persisters Promote the Spread of Antibiotic Resistance Plasmids in the Gut. *Nature* **573**, 276–280 (2019).
16. Licht, T. R., Christensen, B. B. & Krogfelt, K. A. Plasmid Transfer in the Animal Intestine and Other Dynamic Bacterial Populations : the Role of Community Structure and Environment. *Microbiology* **145**, 2615–2622 (1999).
17. Hillman, E. T., Lu, H., Yao, T. & Nakatsu, C. H. Microbial Ecology along the Gastrointestinal Tract. *Microbes Environ.* **32**, 300–313 (2017).
18. Huttenhower, C. *et al.* Structure, Function and Diversity of the Healthy Human Microbiome. *Nature* **486**, 207–214 (2012).
19. Penders, J., Stobberingh, E. E., Savelkoul, P. H. M. & Wolffs, P. F. G. The Human Microbiome as a Reservoir of Antimicrobial Resistance. *Front. Microbiol.* **4**, 1–7 (2013).
20. Doucet-Populaire, F., Trieu-Cuot, P., Dosbaa, I., Andremont, A. & Courvalin, P. Inducible Transfer of Conjugative Transposon Tn1545 from *Enterococcus faecalis* to *Listeria monocytogenes* in the Digestive Tracts of Gnotobiotic Mice. *Antimicrob. Agents Chemother.* **35**, 185–187 (1991).
21. Mataseje, L. F. *et al.* Results from the Canadian Nosocomial Infection Surveillance Program on Carbapenemase-Producing *Enterobacteriaceae*, 2010 to 2014 on behalf of the Canadian Nosocomial Infection Surveillance Program. *Antimicrob. Agents Chemother.* **60**, 1359–16 (2016).
22. Nordmann, P., Naas, T. & Poirel, L. Global Spread of Carbapenemase Producing *Enterobacteriaceae*. *Emerg. Infect. Dis.* **17**, 1791–1798 (2011).
23. Rozwandowicz, M. *et al.* Plasmids Carrying Antimicrobial Resistance Genes in *Enterobacteriaceae*. *J. Antimicrob. Chemother.* **73**, 1121–1137 (2018).

24. McKenzie, G. J. *et al.* Fast, Easy and Efficient: Site-Specific Insertion of Transgenes into Enterobacterial Chromosomes using Tn7 Without need for Selection of the Insertion Event. *BMC Microbiol.* **6**, 1–7 (2006).
25. Krøl, J. E. *et al.* Invasion of *E. coli* Biofilms by Antibiotic Resistance Plasmids. *Plasmid* **70**, 110–119 (2011).
26. Carraro, N., Matteau, D., Luo, P. & Burrus, V. The Master Activator of IncA / C Conjugative Plasmids Stimulates Genomic Islands and Multidrug Resistance Dissemination. *PLoS Genet.* **10**, e1004714 (2014).
27. Guynet, C., Cuevas, A., Moncalián, G. & de la Cruz, F. The *stb* Operon Balances the Requirements for Vegetative Stability and Conjugative Transfer of Plasmid R388. *PLoS Genet.* **7**, (2011).
28. Schinner, S. A. C. *et al.* *Escherichia coli* EDL933 Requires Gluconeogenic Nutrients to Successfully Colonize the Intestines of Streptomycin-treated Mice Precolonized with *E. coli* Nissle 1917. *Infect. Immun.* **83**, 1983–1991 (2015).
29. Ishiwa, A. & Komano, T. The Lipopolysaccharide of Recipient Cells is a Specific Receptor for PilV Proteins, Selected by Shufflon DNA Rearrangement, in Liquid Matings with Donors Bearing the R64 Plasmid. *Mol. Gen. Genet.* **263**, 159–164 (2000).
30. Klimke, W. A. & Frost, L. S. Genetic Analysis of the Role of the Transfer Gene, *traN*, of the F and R100-1 Plasmids in Mating Pair Stabilization during Conjugation. *J. Bacteriol.* **180**, 4036–4043 (1998).
31. Burmølle, M., Norman, A., Sørensen, S. J. & Hansen, L. H. Sequencing of IncX Plasmids Suggests Ubiquity of Mobile forms of a Biofilm-promoting Gene Cassette Recruited from *Klebsiella pneumoniae*. *PLoS One* **7**, (2012).
32. Samuels, A. L., Lanka, E. & Davies, J. E. Conjugative Junctions in RP4-mediated Mating of *Escherichia coli*. *J. Bacteriol.* **182**, 2709–2715 (2000).
33. López-Igual, R., Bernal-Bayard, J., Rodríguez-Patón, A., Ghigo, J. M. & Mazel, D. Engineered Toxin–intein Antimicrobials can Selectively Target and Kill Antibiotic-Resistant Bacteria in Mixed Populations. *Nat. Biotechnol.* **37**, 755–760 (2019).
34. McMillan, E. A. *et al.* Antimicrobial Resistance Genes, Cassettes, and Plasmids present in *Salmonella enterica* Associated with United States food animals. *Front. Microbiol.* **10**, 1–18 (2019).
35. Bikard, D. *et al.* Development of Sequence Specific Antimicrobials based on Programmable CRISPR-Cas Nucleases. *Nat. Biotechnol.* **32**, 1146–1150 (2014).
36. Citorik, R. J., Mimee, M. & Lu, T. K. Sequence-specific Antimicrobials using Efficiently Delivered RNA-guided Nucleases. *Nat. Biotechnol.* **32**, 1141–1145 (2014).

37. Neil, K., Allard, N., Jordan, D. & Rodrigue, S. Assembly of Large Mobilizable Genetic Cargo by Double Recombinase Operated Insertion of DNA (DROID). *Plasmid* **104**, 102419 (2019).
38. Gibson, D. G. *et al.* Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **6**, 343–345 (2009).
39. Green, R. & Rogers, E. J. Chemical Transformation of *E. coli*. *Methods Enzymol.* **529**, 329–336 (2013).
40. Datsenko, K. A. & Wanner, B. L. One-step Inactivation of Chromosomal Genes in *Escherichia coli* K-12 using PCR Products. *PNAS* **97**, 6640–6645 (2000).
41. Datta, S., Costantino, N. & Court, D. L. A set of Recombineering Plasmids for Gram-negative Bacteria. *Gene* **379**, 109–115 (2006).
42. Ferrières, L. *et al.* Silent Mischief: Bacteriophage Mu Insertions Contaminate Products of *Escherichia coli* Random Mutagenesis Performed using Suicidal Transposon Delivery Plasmids Mobilized by Broad-host-range RP4 Conjugative Machinery. *J. Bacteriol.* **192**, 6418–6427 (2010).
43. Marchler-Bauer, A. *et al.* CDD/SPARCLE: Functional Classification of Proteins via Subfamily Domain Architectures. *Nucleic Acids Res.* **45**, D200–D203 (2017).
44. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **215**, 403–10 (1990).
45. Aziz, R. K. *et al.* The RAST Server : Rapid Annotations using Subsystems Technology. *BMC Genomics* **15**, 75–90 (2008).
46. Kumar, S., Tamura, K. & Nei, M. MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. *Brief. Bioinform.* **5**, 150–163 (2004).
47. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **30**, 2114–2120 (2014).
48. Andrew, S. FastQC: a Quality Control Tool for High Throughput Sequence Data. Available online at: www.bioinformatics.babraham.ac.uk/projects/fastqc (2010).
49. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN] (2013).
50. Goryshin, I. Y., Miller, J. A., Kil, Y. V., Lanzov, V. A. & Reznikoff, W. S. Tn5/IS50 Target Recognition. *Proc. Natl. Acad. Sci.* **95**, 10716–10721 (1998).
51. Haeussler, M. *et al.* Navigating Protected Genomics Data with UCSC Genome Browser in a Box. *Bioinformatics* **31**, 764–766 (2015).

52. Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. BLAST Ring Image Generator (BRIG): Simple Prokaryote Genome Comparisons. *BMC Genomics* **12**, 1–10 (2011).

2.9 End notes:

2.9.1. Acknowledgments:

We thank Dominick Matteau for assistance in the sequencing of TP114, the *Plateforme Rnomics* of *Université de Sherbrooke* for assistance with Illumina sequencing, as well as Compute Canada and Compute Quebec for access to bioinformatics resources and support. We are grateful to Pr. Laura Frost for the kind gift of pOX38. We also acknowledge all lab members for thoughtful discussions. This work was supported by the Canadian Institutes of Health Research (CIHR #159817). S.R. holds a Chercheur boursier junior 2 fellowship from the Fonds de recherche du Québec – Santé (FRQS), and K.N. is the recipient of a graduate research scholarship from the Fond de Recherche du Québec – Nature et Technologies (FRQNT) and from the Natural Science and Engineering Research Council of Canada (NSERC). N.A. is supported by a doctoral scholarship from the *Université de Sherbrooke*.

2.9.2. Author contributions:

K.N. and S.R. designed the experiments; K.N. and N.A. performed the experiments; K.N. analyzed the data; F.G. performed bioinformatic analyses for the HDTM and conjugative plasmids sequence assembly except for TP114, which was performed by K.N.; S.R. supervised the project. K.N. and S.R. wrote the manuscript; N.A., V.B. and F.G. revised the manuscript.

2.9.3. Competing interests:

The work presented in the manuscript is part of US provisional patent application 62/696,367. All of the authors of the present manuscript except for F.G. are also co-authors of this provisional patent application. K.N. and S.R. have a financial interest in TATUM bioscience. F.G. declares no competing financial interests and all authors declare no competing non-financial interests.

2.10 Supporting information

2.10.1. Page titre

Supplementary Materials

Highly efficient gene transfer in the mouse gut microbiota is enabled by the Incl_2 conjugative plasmid TP114

Kevin Neil, Nancy Allard, Frédéric Grenier, Vincent Burrus and Sébastien Rodrigue*.

*Correspondence to: Sebastien.Rodrigue@USherbrooke.ca

This file includes:

Supplementary Methods

Supplementary Figs. 2.1 to 2.6

Supplementary Tables 2.1 to 2.3

Supplementary References 1 to 8

2.10.2. Supplementary Methods

2.10.2.1 *Construction of E. coli KN01ΔdapA.*

A DAP auxotrophic variant was obtained by deleting the *dapA* gene in EcN¹ by recombineering using pSIM6. To generate a DAP auxotrophic strain, the *aph-IIIa* resistance cassette of pKD4 was amplified by PCR with added homology for the regions flanking *dapA*. A second PCR round on the purified PCR product then allowed to increase the length of homology. EcN containing an activated pSIM6 was electroporated with the purified PCR product, followed by kanamycin resistance selection. Insertion of the cassette and deletion of *dapA* were then verified by PCR with corresponding primers (Supplementary Table 2.3). After confirmation of *dapA* gene locus replacement by the pKD4 resistance cassette, the strain was cured from pSIM6 by heat-shock at 42°C for 1 hour followed by overnight incubation at 37°C. The culture was then streaked on selective plates to screen for ampicillin sensitive clones. Then, the strain was transformed with pE-FLP to knock-out the resistance cassette inserted in the genome as described previously². pE-FLP plasmid was cured by heat-shock following the same procedure as for pSIM6. Next, the insert from pGRG36-SmSp was added in the genome of EcNΔ*dapA* strain to obtain KN01Δ*dapA*.

2.10.2.2 *Generation of rifampicin and nalidixic resistant mutants of EcN.*

Plasmid TP123, RIP71a and R1 cannot be tested using our standard set of 3 strains (KN01, KN02, KN03) because they harbor the resistance genes used to select donor and recipient pairs (Supplementary Tables 2.1 and 2.4). As such, an additional resistance was needed to test the transfer of these plasmids. We opted for rifampicin and nalidixic acid resistances, as both are chromosomally encoded and easy to obtain. To generate spontaneous resistant variants of KN01, KN01Δ*dapA* and KN02, each strain was first started from frozen stock and allowed to grow overnight. Then, 200 μL of each culture was transferred in LB broth containing 0.5X of the working concentration of either rifampicin or nalidixic acid and incubated overnight at 37°C. The resulting cultures were then streaked on plates containing the working concentration of rifampicin or nalidixic acid and plates were incubated at 37°C overnight. A single colony was isolated,

creating KN01Rf^R, KN01 Δ dapARf^R, and KN02Nx^R. For all three plasmids, KN01 Δ dapARf^R was next used as the donor bacterium for *in vitro* conjugation experiment and KN01Rf^R as the donor for *in situ* conjugation experiment while the recipient strain was always KN02Nx^R.

2.10.2.3 Mice inoculum preparation.

Two days prior to gavage, a frozen stock of the strain(s) used for colonization was streaked onto MacConkey selective plates and incubated overnight at 37°C. The next day, several colonies were inoculated in selective LB broth at 37°C. Three to four hours prior to mice oral challenge, strain(s) were sub-cultured again with a large inoculum (200 μ L or 500 μ L) in 20 mL selective LB broth and incubated at 37°C until an OD_{600nm} of 0.6 ± 0.1 was reached. The cells were then washed once in PBS and concentrated in a volume equivalent to 6.0 OD_{600nm} according to initial OD_{600nm} in LB. A volume of 100 μ L of the cell suspension was force fed to each mouse (corresponding to $\sim 1 \times 10^8$ CFU). Part of the inoculum was plated and analyzed for CFU counts. Cell concentration ranged from 10^8 to 10^9 CFU/mL for all experiments.

2.10.2.4 Feces sampling.

Collection tubes were prepared prior to the experiment by aseptically adding 500 μ L of PBS and a single 0.2 mm glass bead to a 1.5 mL tube. Tubes were weighted before and after sampling to normalize CFU by sample weight. Then, samples were homogenized in a FastPrep-24 (MP) beads beater for 1 minute at maximum speed. The homogenates were centrifuged at 500 x g for 30 seconds to prevent possible pipetting of larger debris. The centrifugation step did not show a significant impact on retrieved CFU. The samples were then serially diluted 1/10 in sterile PBS from 10^0 to 10^{-7} of the initial concentration and 2.5 μ L of each dilution was spotted on selective MacConkey plates in technical duplicates. For each experiment, total *Enterobacteriaceae* clearance was also followed on MacConkey plates without antibiotics as a control for streptomycin treatment.

2.10.2.5 Mice dissection and *EcN* colonization pattern assessment.

On the last day of the experiment, the mice were anesthetized with isoflurane, sacrificed by cervical dislocation and dissected to extract the duodenum, jejunum, ileum, caecum, ascending colon and/or descending colon. To distinguish the parts of the intestine, the first 3 cm of small intestine attached to the stomach were considered to be duodenum, the 6 central cm the jejunum and the 6 last cm (closest to the caecum) the ileum. The ascending and descending colons were the exact halves of the colon. Two spaced quarters of each section and the longitudinal half of the caecum were sampled for CFU analysis. Since the caecum is a large and distinct structure of the mouse intestine, and *EcN* was shown to strongly colonize the caecum (Figure 2.2a), this region was chosen as a representative part of the intestine to study colonization.

2.10.2.6 TP114 sequencing.

TP114 was acquired from DSMZ (DSM-4246) and transferred from *E. coli* K-12 J53-2 by conjugation into *E. coli* MG1655Nx^R. The resulting strain was then grown at 37°C in selective LB broth to obtain enough DNA for sequencing. An Illumina library was prepared using the QIAseq FX Library kit (Qiagen) from size-selected genomic DNA fragments of approximately 400 to 600 base pair (bp). The Illumina library was sequenced on a MiSeq instrument using paired-end reads of 300 bp to assemble longer composite reads covering the entire insert³. A MinION (Oxford Nanopore Technologies, UK) sequencing library was also prepared using 1.5 µg of high-molecular weight genomic DNA and the R9 Nanopore sequencing kit (SQK-NSK007, Oxford Nanopore Technologies, UK). Illumina sequencing reads were assembled with the Roche gsAssembler v2.6 either *de novo* or using reference sequences from other conjugative plasmids from the IncI₂ family (R721, AP002527.1; pChi7122, FR851304; pRM12761, CP007134.1; pSLy21, NZ_CP016405.1). Large *de novo* and reference contigs were then manually assembled and scaffolded with high-quality MinION reads using BLASTn v2.8.1. Finally, 10 regions of 1.5 kb were selected based on lower read coverage and were re-sequenced by Sanger sequencing to confirm the assembly with corresponding primers (Supplementary Table 2.3). The resulting circular sequence was submitted to the RAST v2.0 annotation server⁴, and a total of 92 open reading frames were predicted. The

annotation was then adjusted to name homologous genes consistently between TP114 and the reference IncI₂ plasmid R721 (GenBank: AP002527.1).

2.10.2.7 Sequencing and annotation of other conjugative plasmids.

Plasmids were acquired from various sources (Supplementary Table 1). Upon arrival, strains containing the plasmids were grown overnight in LB broth at 37°C to obtain enough material for DNA extraction. Illumina libraries were prepared using NEBNext Ultra II DNA Library Prep Kit selecting DNA fragment of 150-300 bp. The Illumina libraries were sequenced on a NextSeq 500 instrument using paired-end reads of 75 bp to assemble longer composite reads covering the entire insert³. Illumina sequencing reads were assembled with the Roche gsAssembler v2.6 either *de novo* or using reference sequences when available. Contigs mapping on *E. coli* chromosome were screened out of the assembly to purify reads covering the plasmid DNA. The resulting circular sequences or large contigs were submitted to the RAST v2.0 annotation server⁴. The resulting DNA assembly was submitted to Genbank when significant differences were found with the reference sequence or when the sequence for the plasmid was missing from Genbank.

2.10.2.8 HDTM libraries sequencing.

For each sample, a 1.5 mL frozen stock aliquot of mutant library was thawed on ice for 15 minutes. The aliquot was centrifuged, and cells were resuspended in 300 µL of cell lysis buffer from the Quick gDNA Miniprep kit (ZymoResearch). DNA was fragmented using a Bioruptor UCD-200 sonication system (Diagenode) for 12 cycles of 30 seconds ON, 30 seconds OFF at 4°C. After fragmentation, the Quick gDNA Miniprep kit protocol for cell suspension was performed and DNA was eluted in 50 µL of molecular grade water. 10 µg of the eluted DNA was end-repaired using End-repair Mix HC (Enzymatics) followed by DNA purification with AMPure DNA XP magnetic beads (Agencourt). Purified DNA was then adenylated by TaqB (Enzymatics) supplemented with dATP for 30 minutes at 68°C and purified again with AMPure DNA XP beads (Agencourt). Nextera adaptor B was generated by annealing 2 oligonucleotides: 5'-PO₄-CTGTCTCTTATACACATCTCCGAGCCCACGAGAC-InvdT-3' and 5'-

CAAGCAGAAGACGGCATAACGAGATTTCGCCTTAGTCTCGTGGGCTCGGAGATGTGTA
TAAGAGACAGT-3' together. Annealing was performed by heating 40 μ M of each oligonucleotide in annealing buffer (10 mM Tris NaCl pH 7.5, 50 mM NaCl) to 98°C and then slowly decreasing the temperature by 0.1°C each 10 seconds until 4°C was reached. Nextera adapter B was ligated using T4 DNA ligase (Enzymatics) overnight at 16°C. DNA was purified again using DNA Ampure XP beads (Agencourt) and barcoding was performed in a qPCR machine using Veraseq DNA polymerase (Enzymatics). Amplification reaction was stopped at the end of the exponential phase. DNA was purified again and quantified using Quant-it PicoGreen dsDNA assay (Thermo Fisher). Quality and size distribution of the amplified mutant library was assessed on an Agilent 2100 Bioanalyzer instrument using a High Sensitivity DNA Chip. Mutant libraries were then pooled and sequenced by Illumina using the Nextera technology.

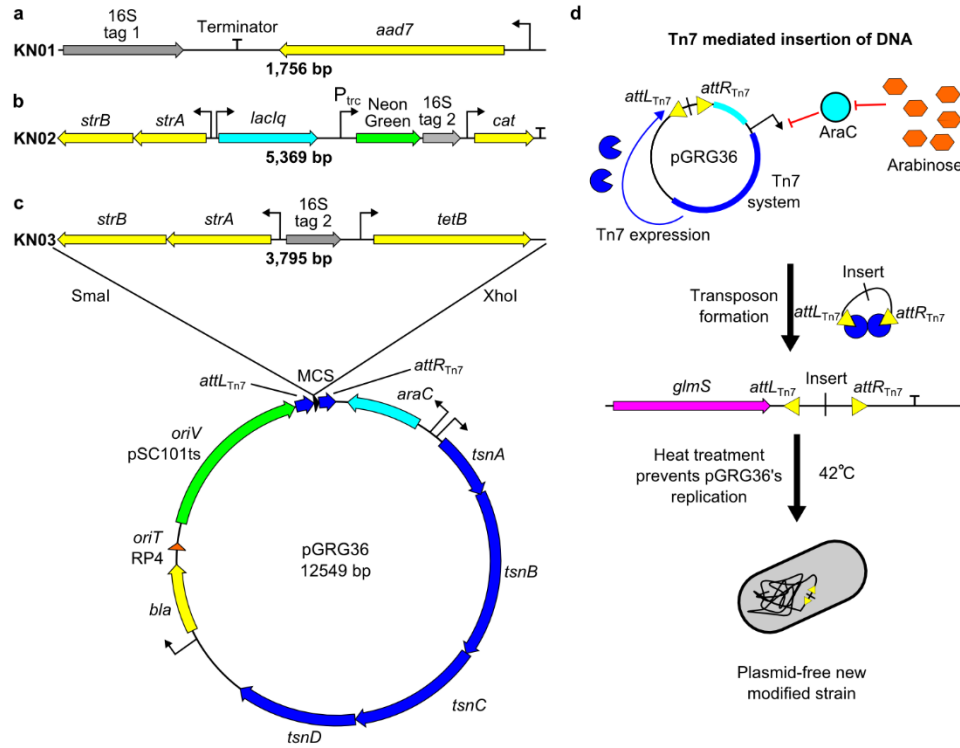
2.10.2.9 Deletion of *pilS* in TP114.

An FRT flanked *cat* gene was amplified from pKD3 (Supplementary Table 2.3). The PCR product was used to delete *pilS* in TP114 by recombineering. The recombinant clones were then screened using appropriate primers (Supplementary Table 3). Cells were then heat-shocked as described in the recombineering section to get rid of the pSIM6 plasmid. Then, to eliminate the *cat* gene from TP114 Δ *pilS*::*cat*, cells were transformed with pE-FLP. Chloramphenicol sensitive clones of TP114 Δ *pilS* were confirmed for *cat* loss by PCR and transferred to *E. coli* strain KN01. The ability of the wild-type and *pilS* mutant versions of TP114 to transfer from *E. coli* KN01 to *E. coli* KN03 was assayed under solid, liquid (static), liquid (shaking) and *in situ* conditions.

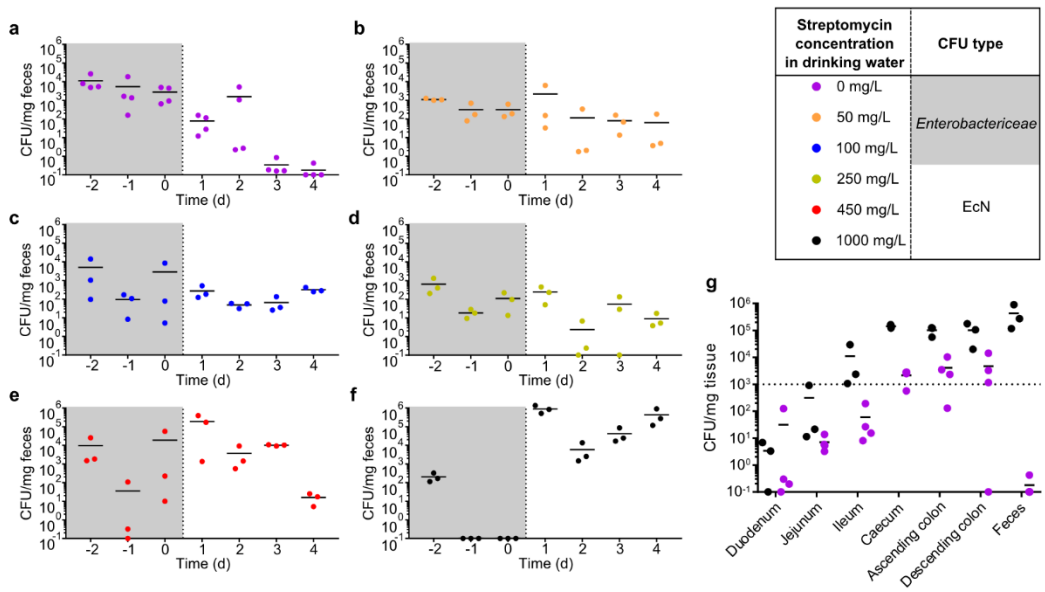
2.10.2.10 Induction of *pilS* expression.

Plasmid p*PilS* was constructed by amplifying the *pilS* gene from TP114, *oriV*_{p15A-araC}-P_{BAD} from pBAD30 and *cat* from pSB1C3 using primers listed in Supplementary Table 3 and joining the amplification products by Gibson assembly. Plasmid p*PilS* was then transformed into KN01 + TP114 Δ *pilS* for complementation studies. The *pilS* gene is under the regulation of AraC⁵, its expression is therefore inducible by arabinose. For *pilS* complementation experiments, donor and

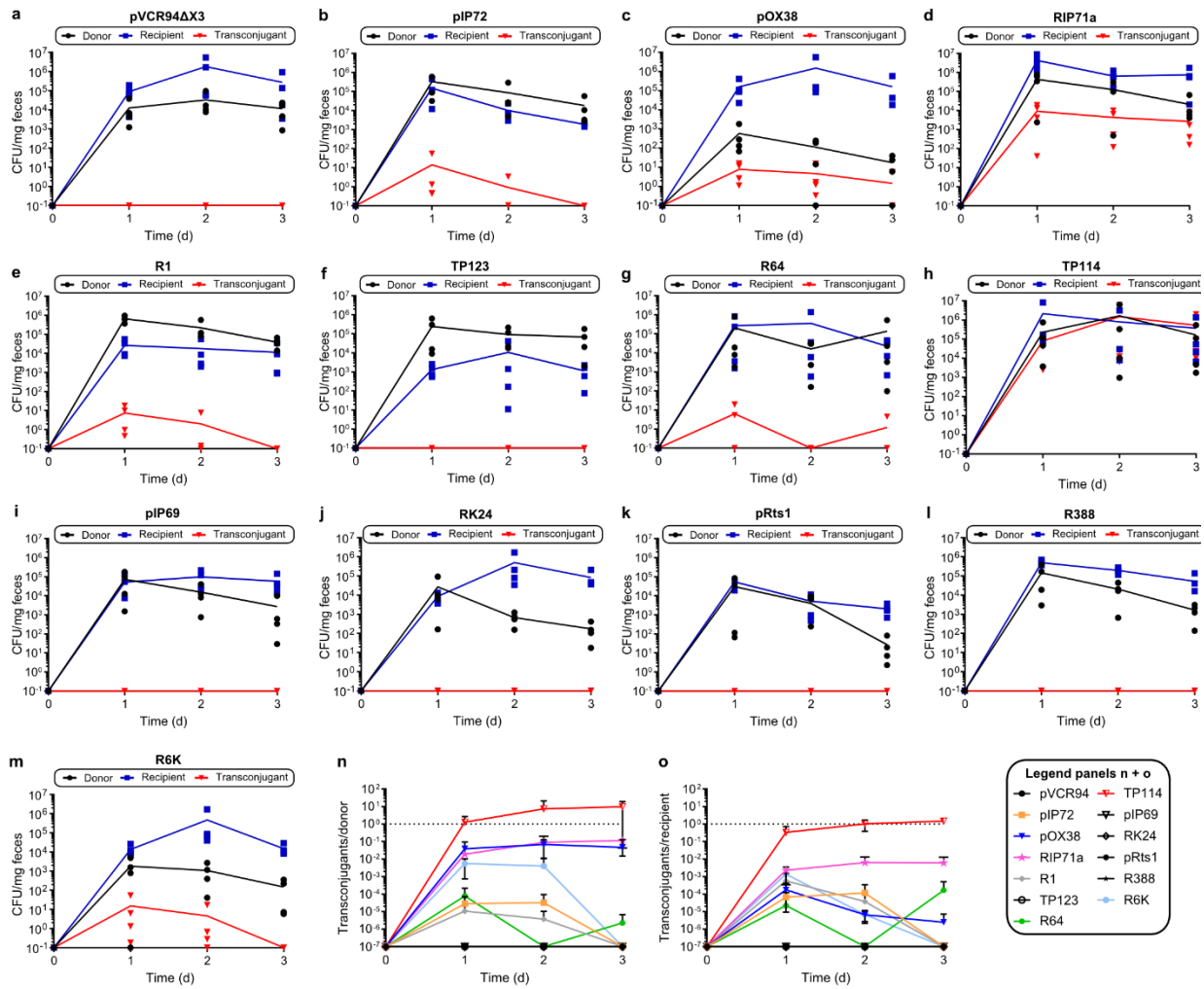
recipient strains were grown overnight at 37°C. Two hours before conjugation, arabinose was added to the donor strain cultures at a final concentration of 1% w/v. Then, OD_{600nm} of each culture was measured and cells were washed in LB + 1% arabinose then resuspended in a volume equivalent to 40 OD_{600nm} in LB + 1% arabinose. A volume of 2.5 µL of the donor and recipient strain were then mixed together and deposited on an LB + 1% arabinose plate for solid conjugation or mixed with 195 µL of pre-warmed LB + 1 % arabinose for both liquid static and liquid shaking conjugations. Conjugation experiments in all conditions were then incubated at 37°C for 2 hours. Additionally, liquid shaking conjugation were placed on a rotary agitator. After incubation, the conjugation experiments were serially diluted 1/10 and plated on selective media for CFU analysis of the donor, recipient, and transconjugant strains.



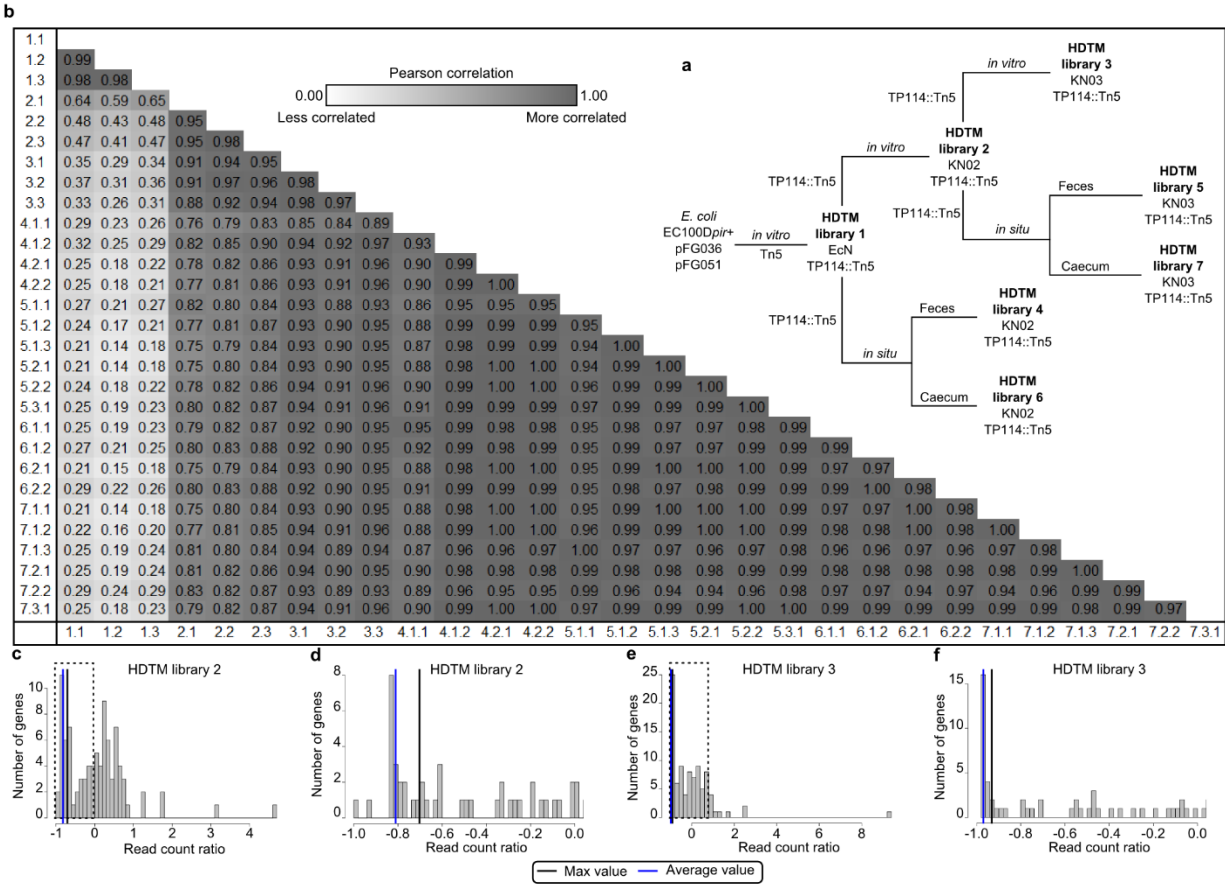
Supplementary Figure 2.1: Specific features of strains used in the conjugation mouse model. Antibiotic resistance inserts were designed to allow quantification of conjugation. Each insert was added into EcN genome using site-specific recombination, creating KN01 (a), KN02 (b), and KN03 (c). Schematic representation of the insertion strategy for the antibiotic resistance markers in EcN genome using an arabinose inducible Tn7 plasmid system (d).



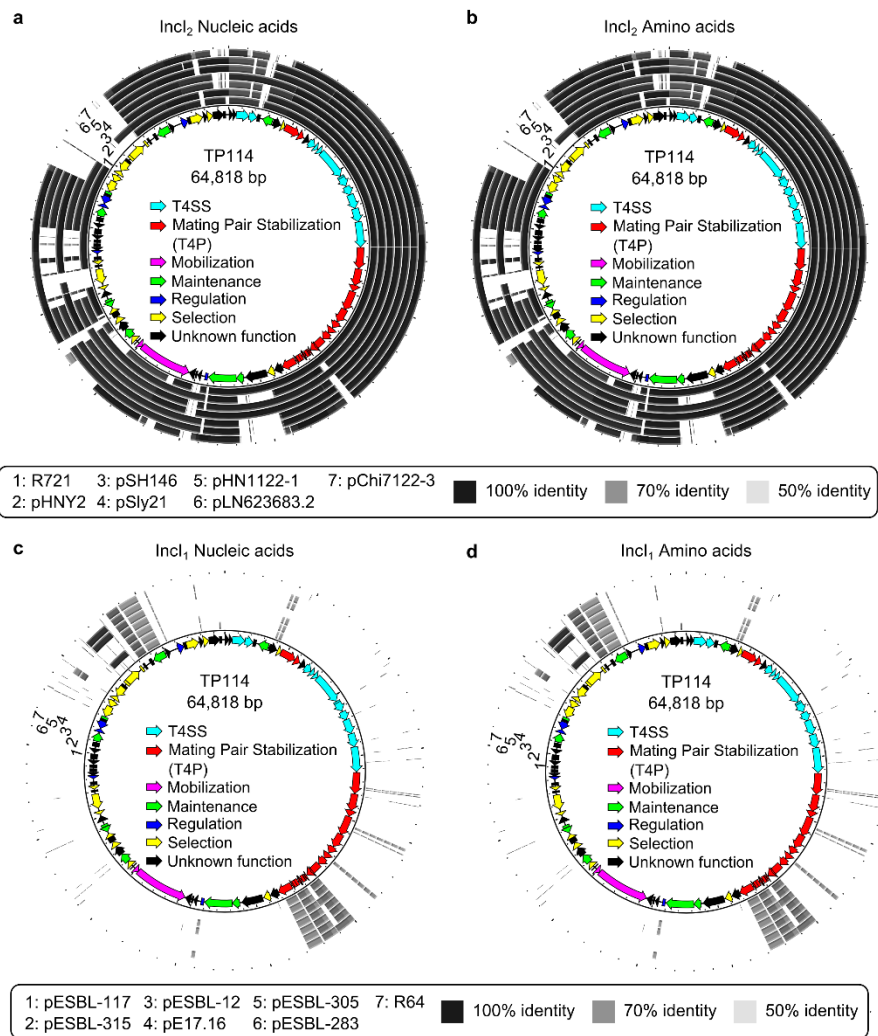
Supplementary Figure 2.2. EcN colonization optimization. *Enterobacteriaceae* depletion (gray area) and EcN colonization (white area) followed by CFU counts from feces of mice treated with 0 mg/L (n=4, biologically independent mice) (a), 50 mg/L (n=3, biologically independent mice) (b), 100 mg/L (n=3, biologically independent mice) (c), 250 mg/L (n=3, biologically independent mice) (d), 450 mg/L (n=3, biologically independent mice) (e), and 1000 mg/L (n=3, biologically independent mice) (f) of streptomycin in drinking water starting at day -2. g, Colonization levels of EcN KN01 in various sections of the intestinal tract for Sm-treated (black) or untreated (purple) mice 4 days after introduction in mice (n=4, biologically independent mice).



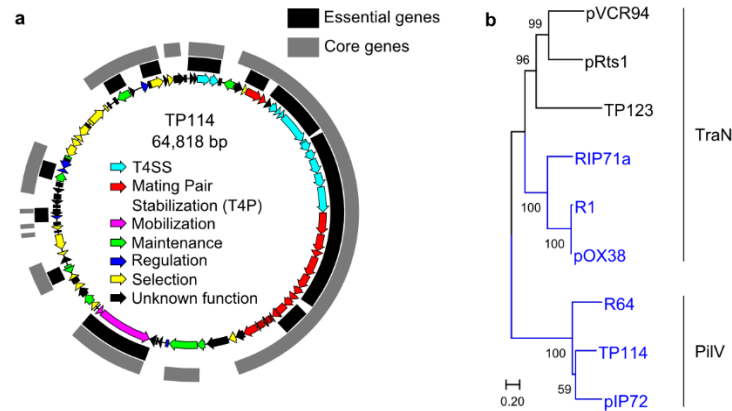
Supplementary Figure 2.3: Transfer rates of 13 plasmids in the mouse intestinal tract. Raw CFU counts of the donor strain (KN01), the recipient strain (KN02 or KN03) and transconjugants in feces collected on day 0, 1, 2 and 3 after the introduction of KN01 for plasmid pVCR94ΔX3 (a), pIP72 (b), pOX38 (c), RIP71a (d), R1 (e), TP123 (f), R64 (g), TP114 (h), pIP69 (i), RK24 (j), pRts1 (k), R388 (l), and R6K (m) (n=4). Calculated transfer rates per donor (n) and per recipient (o) are shown (n=4, biologically independent mice).



Supplementary Figure 2.4: Analysis of TP114 gene essentiality by high-density transposon mutagenesis (HDTM). **a**, Initial mutant library 1 was selected for mutants able to perform *in vitro* or *in situ* conjugation, generating HDTM library 2 to 7. **b**, Pearson correlations between HDTM libraries annotated in a X.Y.Z format where X=HDTM library, Y=library replicate and Z=Mouse when applicable. Gene essentiality was determined by calculating the difference in read counts between HDTM library 1 and test conditions. Essentiality thresholds (max value, black line) were set using a core set of predicted essential genes, as exemplified with HDTM library 2 (**c**) and 3 (**e**). Panel **d** and **f** are a zoom of the framed section of panel **c** and **e**, respectively.



Supplementary Figure 2.5: Sequence homology between TP114 and other plasmids of the IncI family. Sequence homology was evaluated using BRIGG, a BLAST-like program that shows sequence identity in circular pattern. Sequence identity threshold were set at 100%, 70% and 50% for all analysis. The sequence of TP114 was first compared to seven members of the IncI₂ subfamily based on the nucleic acid sequences (**a**), and the amino acid sequences of its coding genes (**b**). Gene conservation among IncI₂ plasmids is also compiled in Supplementary Data 1. TP114 was then compared to seven members of the IncI₁ subfamily based on the nucleic acid sequences (**c**), and the amino acid sequences of its coding genes (**d**). Numbers on homology rings correspond to the plasmids in the legend with 1 being the innermost ring and 7 being the outermost ring.



Supplementary Figure 2.6: TP114 gene essentiality and conservation among IncI₂ plasmid family members. **a**, the TP114 plasmid map displayed in the innermost circle starts at TP114-01 at the top of the circle and continues clockwise with TP114-02 up until TP114-92. Gene essentiality in any HDTM condition as experimentally tested in Figure 3 is highlighted in black while core IncI₂ genes as determined by comparative genomics analysis shown in Supplementary Figure 5 are highlighted in gray. **b**, phylogenetic analysis of known adhesin genes involved in mating pair stabilization. Conjugative plasmids that transferred *in situ* are shown in blue.

Supplementary Table 2.1: Conjugative Plasmids used in this study.

Name	Inc group	MOB group	MPF group	<i>In situ</i> conjugation	Resistance	Isolated in	Size (bp)	CDS count	Genbank	Source
pVCR94ΔX3	C	H12	F	-	Km, Su	<i>Vibrio cholerae</i>	121,195	152	KF551948.1	V. Burrus lab
pIP72	B/O	P7	I	+	Km	<i>Escherichia coli</i>	Unknown	140	MN612051	CRBIP19.42
pOX38	FI	F12	F	+	Sp, Tc, Su	<i>Escherichia coli</i>	59,705	76	MF370216.1	L. Frost lab
RIP71a	FII	F12	F	+	Ap, Tc, Cm, Sm, Sp	<i>Escherichia coli</i>	Unknown	112	MN626601	CRBIP19.47
R1	FII	F12	F	+	Km, Cm, Su, Sp	<i>Salmonella enterica</i>	97,566	135	KY749247.1	CRBIP19.58
TP123	HI	H11	F	-	Sm, Cm, Su, Sp	<i>Salmonella typhi</i>	Unknown	222	MN6266042	CIP TP123
R64	I1	P12	I	+	Sm, Tc	<i>Salmonella typhimurium</i>	120,826	150	NC_005014.1	CRBIP19.30
TP114	I2	P6	T	+	Km	<i>Escherichia coli</i>	64,818	92	MF521836.1	DSM-4246
pIP69	L/M	P13	I	-	Ap, Km, Tc	<i>Salmonella enterica</i>	70,500	111	MN626603	CIP pIP69
RK24	P1α	P11	T	-	Ap, Tc	<i>Escherichia coli</i>	60,096	70	BN000925.1	DSM-3876
pRts1	T	H12	F	-	Km, Sp	<i>Proteus vulgaris</i>	Unknown	162	MN626604	CRBIP19.43
R388	W	F11	T	-	Su, Tm	<i>Escherichia coli</i>	33,913	43	NC_028464	DSM-5189
R6K	X2	P3	T	+	Ap, Sm	<i>Escherichia coli</i>	39,872	52	NZ_LT827129.1	DSM-4245

Supplementary Table 2.2: High Density Transposon Mutagenesis Tnseq results.

Library. Replicate. Mouse	Expected number of reads	Raw reads	% pFG051	Trimmed reads	% trimmed aligned on EcN	Aligned on TP114	% trimmed aligned on TP114	Insertion sites	Insertion sites per Kilobase	Average distance between sites (bp)
1.1	5.20E+07	5.19E+07	4.43	4.81E+07	79.02%	5.55E+06	11.53%	11385	175.65	5.69
1.2	5.20E+07	6.37E+07	8.00	6.02E+07	75.05%	6.59E+06	10.95%	10873	167.75	5.96
1.3	5.20E+07	4.96E+07	8.59	4.68E+07	76.05%	5.25E+06	11.21%	12710	196.09	5.10
2.1	6.20E+05	2.38E+06	6.68	2.25E+06	65.79%	5.35E+05	23.77%	11470	176.96	5.65
2.2	6.20E+05	1.48E+06	0.33	1.36E+06	18.60%	1.07E+06	79.06%	9745	150.34	6.65
2.3	6.20E+05	6.93E+05	0.30	6.53E+05	13.60%	5.50E+05	84.21%	8929	137.75	7.26
3.1	6.20E+05	9.29E+05	0.21	8.49E+05	3.52%	7.78E+05	91.71%	5050	77.91	12.84
3.2	6.20E+05	7.89E+05	1.09	7.23E+05	11.61%	6.03E+05	83.39%	6335	97.74	10.23
3.3	6.20E+05	1.45E+06	2.80	1.36E+06	28.97%	8.85E+05	65.22%	11045	170.40	5.87
4.1.1	6.20E+05	4.72E+06	4.75	4.42E+06	48.56%	1.89E+06	42.79%	14331	221.10	4.52
4.1.2	6.20E+05	2.29E+06	5.62	2.15E+06	58.09%	6.88E+05	32.06%	12582	194.11	5.15
4.2.1	6.20E+05	1.81E+06	3.31	1.67E+06	34.44%	9.93E+05	59.45%	10901	168.18	5.95
4.2.2	6.20E+05	1.86E+06	1.88	1.68E+06	39.15%	8.96E+05	53.26%	8503	131.18	7.62
4.3.1	6.20E+05	1.85E+06	4.16	1.75E+06	40.66%	8.94E+05	51.09%	7924	122.25	8.18
4.3.2	6.20E+05	1.59E+06	4.08	1.51E+06	37.14%	8.51E+05	56.49%	7667	118.29	8.45
5.1.1	6.20E+05	3.15E+05	0.76	2.99E+05	8.01%	2.70E+05	90.21%	4212	64.98	15.39
5.1.2	6.20E+05	4.79E+05	<0.2	4.51E+05	2.41%	4.35E+05	96.60%	4927	76.01	13.16
5.1.3	6.20E+05	6.14E+05	<0.2	5.79E+05	0.96%	5.70E+05	98.45%	4164	64.24	15.57
5.2.1	6.20E+05	6.32E+05	<0.2	6.02E+05	0.52%	5.94E+05	98.76%	4517	69.69	14.35
5.2.2	6.20E+05	5.78E+05	<0.2	5.47E+05	1.45%	5.33E+05	97.44%	4356	67.20	14.88
5.3.1	6.20E+05	7.88E+05	<0.2	7.48E+05	2.33%	7.21E+05	96.44%	7285	112.39	8.90

Supplementary Table 2.2: (continued)

6.1.1	6.20E+05	1.55E+06	0.22	1.45E+06	3.63%	1.38E+06	94.90%	8777	135.41	7.38
6.1.2	6.20E+05	7.44E+05	0.27	6.97E+05	4.21%	6.57E+05	94.33%	7604	117.31	8.52
6.2.1	6.20E+05	7.47E+05	0.12	6.94E+05	2.93%	6.65E+05	95.81%	7771	119.89	8.34
6.2.2	6.20E+05	2.41E+06	1.45	2.14E+06	37.94%	1.10E+06	51.20%	8858	136.66	7.32
6.3.1	6.20E+05	7.81E+05	<0.2	7.43E+05	3.90%	7.04E+05	94.77%	3834	59.15	16.91
6.3.2	6.20E+05	1.17E+06	<0.2	1.12E+06	1.72%	1.09E+06	97.56%	3600	55.54	18.01
7.1.1	6.20E+05	3.12E+05	<0.2	2.97E+05	0.65%	2.92E+05	98.48%	4738	73.10	13.68
7.1.2	6.20E+05	5.97E+05	<0.2	5.63E+05	1.42%	5.49E+05	97.52%	5397	83.26	12.01
7.1.3	6.20E+05	1.52E+06	<0.2	1.44E+06	1.55%	1.41E+06	97.53%	8136	125.52	7.97
7.2.1	6.20E+05	2.70E+06	0.13	2.53E+06	2.38%	2.44E+06	96.48%	9081	140.10	7.14
7.2.2	6.20E+05	7.15E+05	0.27	6.73E+05	4.30%	6.34E+05	94.23%	7003	108.04	9.26
7.3.1	6.20E+05	8.22E+05	0.12	7.75E+05	1.96%	7.51E+05	96.92%	6943	107.12	9.34

Supplementary Table 2.3: List of strains, plasmids and primers used in this study

<i>E. coli</i>	Relevant genotype/phenotype	Source/Reference
EC100Dpir+	F– <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) φ80 <i>dlacZ</i> ΔM15 Δ <i>lacX74 recA1 endA1 araD</i> 139 Δ(<i>ara, leu</i>)7697 <i>galU galK</i> λ– <i>rpsL nupG pir+</i> (DHFR)	#ECP09500 (Lucigen)
KN01	Sm ^R Sp ^R Nissle 1917	This study
KN01Rf ^R	Spontaneous Rifampicin resistant KN01	This study
KN01Δ <i>dapA</i>	Δ <i>dapA</i> KN01	This study
KN01Δ <i>dapA</i> Rf ^R	Spontaneous Rifampicin resistant KN01Δ <i>dapA</i>	This study
KN02	Sm ^R Cm ^R Nissle 1917	This study
KN02Nx ^R	Spontaneous Nalidixic acid resistant KN02	This study
KN03	Sm ^R Tc ^R Nissle 1917	This study
MFD <i>pir+</i>	MG1655 RP4-2-Tc::[Δ <i>Mu1::aac</i> (3)IV-Δ <i>aphA-Δnic35-ΔMu2::zeo</i>] Δ <i>dapA::(erm-pir) ΔrecA</i>	Ferrière <i>et al.</i> 2010 ⁶
MG1655Nx ^R	K-12 F– λ– <i>ilvG– rfb-50 rph-1</i> Nx ^R	Carraro <i>et al.</i> 2014 ⁷
Nissle 1917	Wildtype probiotic strain	DSM-6601 (DSMZ)
Plasmid	Relevant genotype/phenotype	Source/Reference
pBAD30	<i>oriV</i> _{p15A} , <i>bla</i> , <i>araC</i> , P _{BAD}	Guzman <i>et al.</i> 1995 ⁵
pE-FLP	<i>oriV</i> _{pSC101ts} , <i>flp</i> , Ap ^R	Addgene #45978
pFG036	<i>cl</i> , <i>tetM</i> , <i>oriV</i> ColE1	Addgene #137996
pFG051	<i>aad7</i> , <i>cat</i> , Tn5 <i>tnp</i> , <i>oriV</i> R6K, <i>oriT</i> RP4	Addgene #137997
pGRG36	<i>oriV</i> _{pSC101ts} , Tn7 insertion machinery, <i>araC</i> , Ap ^R	Addgene #16666
pGRG36-SmCm	pGRG36 with Sm ^R Cm ^R insert	This study
pGRG36-SmSp	pGRG36 with Sm ^R Sp ^R insert	This study
pGRG36-SmTc	pGRG36 with Sm ^R Tc ^R insert	This study
pKD3	R6K <i>oriV</i> , FRT flanked Cm ^R , Ap ^R	Addgene #45604
pKD4	R6K <i>oriV</i> , FRT flanked Km ^R , Ap ^R	Addgene #45605
pPiS	<i>oriV</i> _{p15A} , <i>cat</i> , <i>araC</i> , P _{BAD} , <i>pilS</i>	This study

Supplementary table 2.3: (continued)

pSB1C3	<i>oriV_{pMB1}</i> , <i>cat</i> , Biobricks	IGEM
pSIM6	<i>oriV_{pSC101ts}</i> , Lambda Red recombinase, Ap ^R	Datsenko and Wanner 2000 ⁸
TP114Δ <i>pilS</i> :: <i>cat</i>	TP114Δ <i>pilS</i> :: <i>cat</i>	This study

Primer	Sequence*	Usage
oGST1-F	<u>GATCCTAGTAAGCCACGTTTTAATTAATCAGATCCCGGGCTAGTATGACGTCTGTGCGCAC</u>	<i>strA-strB</i> for pGRG36-SmTc
oGST1-R	<u>GTAGTTATTTATATAATTCATCCATCCATAACATCTGTTTACAGCTAGCTCAGTCCT</u>	
oGST2-F	<u>AAGCTAGCATAATACCTAGGACTGAGCTAGCTGTAACAGATGTTATGGGAATGGATGAA</u>	16S tag for pGRG36-SmTc
oGST2-R	<u>CCCCAAAACCTTCCCAAAAACCTTCCCAAAAACCTGGCTATACTCGAGGCATGCCTGCGAG</u>	
oGST3-F	<u>GGATTAGATACCCTGGTAGTCTGCGAGGCATGCCTCGAGTATAGCCAGTTTTGGGAAGG</u>	<i>tetB</i> for pGRG36-SmTc
oGST3-R	<u>GGGGTCGACGCGGCCGTGGCGCGCCTCCTAGGTGCTCGAGCTAGGTCGACGCTTGGATTG</u>	
oGSS1-F	<u>GGATCCTAGTAAGCCACGTTTTAATTAATCAGATCCCGGGCTACGGGAGGCAGCAGTGG</u>	16S tag for pGRG36-SmSp
oGSS1-R	<u>CGGGGAAGTACGAGGGTATGGTGC GCGCATGAAAAGACTACCAGGGTATCTAATCCTGTT</u>	
oGSS2-F	<u>CGCACCATACCCTCCTAGTTCCTCGGTTATCTCTCCTGTCTCTTATACACATCTGACGCT</u>	<i>aad7</i> for pGRG36-SmSp
oGSS2-R	<u>GGGGTCGACGCGGCCGTGGCGCGCCTCCTAGGTGCTCGAGGCAAGCGAACCGGAATTGCC</u>	
oGSC1-F	<u>GATCCTAGTAAGCCACGTTTTAATTAATCAGATCCCGGGCTAGTATGACGTCTGTGCGCAC</u>	<i>strA-strB</i> for pGRG36-SmCm
oGSC1-R1	<u>TGCTAGCTTTGAAAATTAAGAGGTATATATTGAATCGAACTAATATTTTTTTGGTG</u>	
oGSC1-R2	<u>GTGTCAACGTTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCTTTGAAAATTAAGAGG</u>	Add P1-U8 for pGRG36-SmCm
oGSC2-F1	<u>TACCTAGGACTGAGCTAGCTGTAACGTTGACACCATCGAATGGTGCAAAACCTTTCGCG</u>	<i>lacI-P_{trc}</i> for pGRG36-SmCm
oGSC2-F2	<u>ATATCCCGAATGTGCAGTTAACGACGTTGACACCATCGAATGGTGCAAAACCTTTCGCGG</u>	
oGSC2-R	<u>TAATATATACCTCTTTAATTTTTAATAATAAAGTTAATCG</u>	
oGSC3-F	<u>TATTATTAATAAATAAGAGGTATATATTAATGGTTTCTAAAGGAGAAGAAAAAATATG</u>	NeonGreen for pGRG36-SmCm
oGSC3-R	<u>CCTCTTCCACTGCTGCCTCCCGTAGGTTATTTATATAATTCATCCATCCATAACATC</u>	
oGSC4-F	<u>AGCATTTACAGATGTTATGGGAATGGATGAATTATATAAATAACCTACGGGAGGCAGCAG</u>	16S tag for pGRG36-SmCm
oGSC4-R	<u>ACCTCTTACGTGCCCCGATCAACTCGAGGCATGCCTGCAGGACTACCAGGGTATCTAATCC</u>	
oGSC5-F	<u>GAACAGGATTAGATACCCTGGTAGTCTGCGAGGCATGCCTCGAGTTGATCGGGCAGGTAA</u>	<i>cat</i> for pGRG36-SmCm
oGSC5-R	<u>GGTCGACGCGGCCGTGGCGCGCCTCCTAGGTGCTCGAGCTCGAGGCTTGGATTCTACCA</u>	

Supplementary table 2.3: (continued)

oGSC6-F	<u>GAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGACGCTTAATGCGCCGCTACAG</u>	Screen pGRG36 plasmid presence
oGSC6-R	<u>CTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACC<u>CGCGCAATTAACCTCACTA</u></u>	
oDTD1-F	<u>CGCGTGCCTCGGCAAAATGCCCTTCTGCTGCCAGTTTGCAGTGTAGGCTGGAGCTGCTTC</u>	<i>aph-IIIa</i> for <i>dapA</i> deletion
oDTD1-R	<u>AAACGTACCATTGAGACACTTGTGGCACAGAGGATGGCCCTGGGAATTAGCCATGGTCC</u>	
oDTD2-F	<u>CA*A*A*TAGTTTGTGTGTAATGGCATCAGACGCTGATTAATAACGCGTGCCTCGGCAAAAT</u>	Add homology
oDTD2-R	<u>AC*G*G*TTCTGTCTGCTTGTCTTTAATGCCATACCA<u>ACGTACCATTGAGACACTTGT</u>TTGC</u>	
oDTD3-int	<u>CCGTTTCTGCGGACTGGCTT</u>	Screen <i>dapA</i> deletion
oDTD23-F	<u>TAGCGGCTATCACCAACATC</u>	
oDTD3-R	<u>GTGAAGCGCCTTATGAACAATG</u>	
oTPS1-F	<u>TGCTGAACCAGTAACAACCACC</u>	TP114 1 bp sanger
oTPS1-R	<u>GTCGCCGCTGTGGATTCAAC</u>	
oTPS2-F	<u>GTTCAATACACATTACAGCCCACC</u>	TP114 10 kb sanger
oTPS2-R	<u>CTGCGCTCAAAGTCACGTATGG</u>	
oTPS3-F	<u>TTACGCAACAGAATCTGAAAGCAC</u>	TP114 20 kb sanger
oTPS3-R	<u>GAAGGTGGCCTGTCATCGAG</u>	
oTPS4-F	<u>TGTCCGATTCGTCCTGGTTG</u>	TP114 25 kb sanger
oTPS4-R	<u>GTATTTGTCCAGCGCCCGG</u>	
oTPS5-F	<u>TTCAGATGCGTCGTGCAATG</u>	TP114 37 kb sanger
oTPS5-R	<u>CACACTTGAGCGTCTTTCTGA</u>	
oTPS6-F	<u>AGAAGCTCTTGAGTCCGACC</u>	TP114 42,5 kb sanger
oTPS6-R	<u>GACTTATTCGCCAACCCAAATI</u>	
oTPS7-F	<u>GGCCCGCTCAAGGCTTTTC</u>	TP114 47 kb sanger
oTPS7-R	<u>GCTGGAGAACCCTGATTATGT</u>	
oTPS8-F	<u>AAAGTTCTTTGCGCCTGTCATAGC</u>	TP114 50 kb sanger
oTPS8-R	<u>GAAGCCAGGTTTGTGCTGTG</u>	
oTPS9-F	<u>TTTCTCTGCTACAGCATCTTCTTC</u>	TP114 52,5 kb sanger

Supplementary table 2.3: (continued)

oTPS9-R	<u>GGAAGTGCCTCGGTGAAT</u>	
oTPS10-F	<u>GGCATAAGGCGTGACAAATGG</u>	TP114 57,5 kb sanger
oTPS10-R	<u>CAAACGTGCTAATCGCCTGGC</u>	
opilS1-F	ATGTCTTCTATTAATATTTTAAATATGCGTTCTGTTTTTT <u>GTGTAGGCTGGAGCTGCTTC</u>	Generate TP114Δ <i>pilS</i> :: <i>cat</i>
opilS1-R	TCAGGAATCAGTGCTGAAGGTCAGCGTATTGCTGTCAGATAT <u>TGGGAATTAGCCATGGTCC</u>	
opilS2-F	<u>TACGTCCTGCAACACTAAT</u>	Screen <i>pilS</i> deletion clones
opilS2-R	<u>GCTTATCCGATGCACATGAA</u>	
opPiIS-F1	<u>CTGTCAGACCAAGTTTACTC</u>	<i>oriV_{p15A}</i> , <i>araC P_{BAD}</i> for pPiIS
opPiIS-R1	TTTTGCCTCCTAC <u>GCTAGCCCAAAAAACGGG</u>	
opPiIS-F2	<u>AACCGTATTACCGCCTTTGAG</u>	<i>cat</i> for pPiIS
opPiIS-R2	TCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGG <u>TAACTTGGTCTGACAGCTCG</u>	
opPiIS-F3	TCTCCATACCCGTTTTTTTTGGGCTAGCGTAGGAGGCAAAA <u>ATGCTTCTATTAATATTTT</u>	<i>pilS</i> for pPiIS
opPiIS-R3	GGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTT <u>CAGGAATCAGTGCTGAAGG</u>	

Supplementary Data 2.1: TP114 *in silico* analysis of gene function, essentiality and conservation

Locus tag	Gene function			Essentiality confidence level ^a				Conservation
	name in TP114	Protein family/Identity	Known homolog/Function	Maintenance	<i>in vitro</i> conjugation	<i>in situ</i> conjugation only	<i>in vitro</i> + <i>in situ</i> conjugation	
TP114-001	<i>yhaB</i>	None	Unknown	-	++	++	++	Core
TP114-002	<i>yhaA</i>	None	Unknown	-	+	++	++	Core
TP114-003	<i>traA</i>	pfam04610	<i>virB6</i>	-	++	++	++	Core
TP114-004	<i>trbJ</i>	pfam07996	<i>virB5</i>	-	++	++	++	Core
TP114-005	<i>TP114-005</i>	None	Unknown	-	-	-	-	Accessory
TP114-006	<i>ygiA</i>	BLASTx 98% identity	DNA primase	-	-	-	-	Soft-core
TP114-007	<i>yggB</i>	None	Unknown	-	-	-	-	Core
TP114-008	<i>kikA</i>	pfam07424	p-toluenesulfonate degradation (<i>trbM</i>)	-	-	++	+	Core
TP114-009	<i>pilL</i>	pfam10671	<i>tcpQ</i>	-	-	++	++	Core
TP114-010	<i>pilM</i>	pfam07419	<i>pilM</i>	-	-	-	-	Core
TP114-011	<i>ygeA</i>	None	Unknown	-	+	++	++	Core
TP114-012	<i>traB</i>	pfam01464	<i>virB1</i>	-	++	+	++	Core
TP114-013	<i>traC</i>	BLAST 99% identity	<i>virB2</i>	-	++	++	++	Core
TP114-014	<i>traD</i>	pfam05101	<i>virB3</i>	-	++	++	++	Core
TP114-015	<i>traE</i>	pfam03135	<i>virB4</i>	-	++	++	++	Core
TP114-016	<i>traF</i>	BLAST 100% identity	R721 <i>traF</i> (<i>virB7</i>)	-	-	-	-	Core
TP114-017	<i>traG</i>	pfam04335	<i>virB8</i>	-	++	++	++	Core
TP114-018	<i>traH</i>	pfam03524	<i>virB9</i>	-	++	++	++	Core
TP114-019	<i>traI</i>	pfam03743	<i>virB10</i>	-	++	++	++	Core
TP114-020	<i>traJ</i>	pfam00437	<i>virB11</i>	-	++	++	++	Core
TP114-021	<i>traK</i>	pfam02534	<i>virD4</i>	-	++	++	++	Core
TP114-022	<i>pilN</i>	TIGR02520	<i>pilN</i>	-	-	++	++	Core

Supplementary Data 2.1: (continued)

TP114-023	<i>pilO</i>	pfam06864	<i>pilO</i>	-	-	++	++	Core
TP114-024	<i>pilP</i>	TIGR03021	<i>pilP</i>	-	-	++	++	Core
TP114-025	<i>pilQ</i>	pfam00437	<i>virB11</i>	-	-	++	++	Core
TP114-026	<i>pilR</i>	pfam00482	T2SSF	-	-	++	++	Core
TP114-027	<i>pilS</i>	pfam08805	<i>pilS</i>	-	-	++	++	Core
TP114-028	<i>pilT</i>	cl00222	<i>virB1</i>	-	-	-	-	Core
TP114-029	<i>pilU</i>	pfam01478	Peptidase_A24	-	-	++	++	Core
TP114-030	<i>pilV</i>	pfam04917	Shufflon_N	-	-	++	++	Core
TP114-031	<i>rci</i>	pfam00589	Phage integrase	-	-	-	-	Core
TP114-032	<i>yebA</i>	None	Unknown	-	-	-	-	Core
TP114-033	<i>yeaA</i>	pfam00565	Endonuclease	-	-	-	-	Core
TP114-034	<i>YdjA</i>	None	Unknown	-	-	-	-	Accessory
TP114-035	<i>ydiA</i>	BLASTx 99% identity	DNA topoisomerase III	-	-	-	-	Accessory
TP114-036	<i>ydgA</i>	cl27597	DNA topoisomerase IA	-	-	-	-	Core
TP114-037	<i>hha</i>	pfam05321	Haemolysin expression modulating protein	-	-	-	-	Core
TP114-038	<i>TP114-038</i>	None	Unknown	-	-	-	-	Accessory
TP114-039	<i>TP114-039</i>	None	Unknown	-	-	-	-	Accessory
TP114-040	<i>yddA</i>	None	Unknown	-	-	-	-	Core
TP114-041	<i>nikB</i>	pfam03432	Relaxase	-	++	++	++	Core
TP114-042	<i>nikA</i>	pfam05713	<i>mobC</i>	-	++	++	++	Core
TP114-043	<i>TP114-043</i>	BLASTx 98% identity	Putative zinc transporter	-	+	-	++	Soft-core
TP114-044	<i>traL</i>	cl27521	<i>parA</i>	-	-	+	++	Soft-core
TP114-045	<i>TP114-045</i>	None	Unknown	-	-	-	-	Soft-core

Supplementary Data 2.1: (continued)

TP114-046	<i>ycgB</i>	BLASTx 98% identity	valyl-tRNA synthase	-	-	-	-	Accessory
TP114-047	<i>TP114-047</i>	None	Unknown	-	-	-	-	Accessory
TP114-048	<i>TP114-048</i>	None	Unknown	-	-	-	-	Accessory
TP114-049	<i>TP114-049</i>	BLASTx 88% identity	valyl-tRNA synthase	-	-	-	-	Core
TP114-050	<i>ycfB</i>	pfam15919	<i>hicB</i> Anti-toxin	++	++	++	++	Core
TP114-051	<i>ycfA</i>	pfam07927	<i>hicA</i> toxin	++	++	++	++	Core
TP114-052	<i>yceB</i>	None	Unknown	-	-	-	-	Core
TP114-053	<i>hicB</i>	pfam05534	Rnase H	-	-	-	-	Accessory
TP114-054	<i>TP114-054</i>	pfam01385	Transposase	-	-	-	-	Accessory
TP114-055	<i>TP114-055</i>	None	Unknown	-	-	-	-	Core
TP114-056	<i>yajC</i>	BLASTx 99% identity	cobalamin biosynthesis (<i>cibX</i>)	-	-	-	-	Accessory
TP114-057	<i>TP114-057</i>	None	Unknown	-	-	-	-	Core
TP114-058	<i>yajA</i>	pfam15731	Transcription regulator	-	++	++	++	Accessory
TP114-059	<i>yaiB</i>	None	Unknown	-	-	++	++	Accessory
TP114-060	<i>TP114-060</i>	None	Unknown	-	++	++	++	Core
TP114-061	<i>yaiA</i>	None	Unknown	-	-	-	-	Accessory
TP114-062	<i>yahA</i>	None	Unknown	-	-	-	-	Accessory
TP114-063	<i>TP114-063</i>	None	Unknown	-	-	-	-	Soft-core
TP114-064	<i>TP114-064</i>	None	Unknown	-	-	-	-	Core
TP114-065	<i>yagA</i>	None	Unknown	-	-	-	-	Core
TP114-066	<i>TP114-066</i>	None	Unknown	-	-	-	-	Core
TP114-067	<i>TP114-067</i>	None	Unknown	-	-	-	-	Core
TP114-068	<i>parA</i>	cd02042	<i>parA</i>	++	++	++	++	Core

Supplementary Data 2.1: (continued)

TP114-069	<i>yafA</i>	BLASTx 99% identity	<i>cogG</i> Transcription regulator	-	++	++	++	Core
TP114-070	<i>yaeC</i>	cd00236	<i>finO</i> Transcription repressor	-	-	-	-	Core
TP114-071	<i>TP114-071</i>	None	Unknown	-	-	-	-	Core
TP114-072	<i>TP114-072</i>	cl10143	Post-Segregation killing protein	-	-	-	-	Core
TP114-073	<i>yadA</i>	cl26233	Transposase	-	-	-	-	Accessory
TP114-074	<i>yadB</i>	cl26146	Transposase	-	-	-	-	Accessory
TP114-075	<i>tnpA</i>	BLASTx 100% identity	Tn3 family transposase	-	-	-	-	Accessory
TP114-076	<i>aph-III</i>	cd05150	Aminoglycoside 3'-phosphotransferase	-	-	-	-	Accessory
TP114-077	<i>TP114-077</i>	None	Unknown	-	-	-	-	Accessory
TP114-078	<i>TP114-078</i>	BLASTx 91% identity	Transposase	-	-	-	-	Accessory
TP114-079	<i>TP114-079</i>	pfam04986	Transposase	-	-	-	-	Accessory
TP114-080	<i>TP114-080</i>	BLASTx 100% identity	IS91 transposase	-	-	-	-	Core
TP114-081	<i>TP114-081</i>	None	Unknown	-	-	-	-	Core
TP114-082	<i>TP114-082</i>	None	Unknown	++	++	++	++	Core
TP114-083	<i>repA</i>	pfam02387	<i>repA</i>	++	++	++	++	Core
TP114-084	<i>TP114-084</i>	None	Unknown	-	-	-	-	Core
TP114-085	<i>yheC</i>	pfam07180	CaiF/GriA transcriptional regulator	-	-	+	-	Core
TP114-086	<i>yheB</i>	pfam06688	Unknown	-	-	++	++	Core
TP114-087	<i>ydhA</i>	cd08826	Stomatin-like protein	-	-	++	++	Core
TP114-088	<i>TP114-088</i>	None	Unknown	-	-	+	-	Accessory
TP114-089	<i>yhcA</i>	cl00686	Regulation of membrane protease activity	-	-	+	-	Core

Supplementary Data 2.1: (continued)

TP114-090	<i>TP114-090</i>	None	Unknown	-	-	-	+	Core
TP114-091	<i>yhbB</i>	None	Unknown	-	-	-	+	Soft-core
TP114-092	<i>TP114-092</i>	None	Unknown	-	-	+	-	Soft-core

a. Essentiality confidence score: ++ essential in all replicate, + essential in at least one replicate, - essential in no replicate

2.10.3. Supplementary References:

1. Born, T. L. & Blanchard, J. S. Structure/function studies on enzymes in the diaminopimelate pathway of bacterial cell wall biosynthesis. *Curr. Opin. Chem. Biol.* **3**, 607–613 (1999).
2. St-Pierre, F. *et al.* One-step cloning and chromosomal integration of DNA. *ACS Synth. Biol.* **2**, 537–541 (2013).
3. Rodrigue, S. *et al.* Unlocking short read sequencing for metagenomics. *PLoS One* **5**, e11840 (2010).
4. Aziz, R. K. *et al.* The RAST Server : Rapid Annotations using Subsystems Technology. *BMC Genomics* **15**, 75–90 (2008).
5. Guzman, L., Belin, D. & Carson, M. J. Tight regulation , modulation , and high-level expression by vectors containing the arabinose P BAD promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
6. Ferrières, L. *et al.* Silent mischief: Bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *J. Bacteriol.* **192**, 6418–6427 (2010).
7. Carraro, N., Matteau, D., Luo, P. & Burrus, V. The master activator of IncA / C conjugative plasmids stimulates genomic islands and multidrug resistance dissemination. *PLoS Genet.* **10**, e1004714 (2014).
8. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* **97**, 6640–6645 (2000).

CHAPITRE 3

ASSEMBLY OF LARGE MOBILIZABLE GENETIC CARGO BY DOUBLE RECOMBINASE OPERATED INSERTION OF DNA (DROID)

3.1 Présentation de l'article et contributions

Les outils de manipulation de l'ADN ont grandement évolué dans les dernières décennies. Une panoplie de techniques sont maintenant disponibles chez *E. coli* afin d'assembler des constructions sous forme de plasmides ou de faire des modifications génétiques. Par contre, très peu d'outils sont efficaces pour la fusion de deux grandes molécules d'ADN. En effet, les méthodes d'assemblage comme le Gibson permettent la création de plasmides en collant plusieurs morceaux d'ADN ensemble, mais plus ces constructions sont grandes, plus il est difficile de les transformer chez *E. coli* (Gibson *et al*, 2009). De la même façon, le Recombineering permet la délétion de grandes sections de génome, mais son efficacité est inversement proportionnelle à la taille du fragment à insérer (Kuhlman & Cox, 2010). La fusion de TP114 (~65 kpb) avec le module CRISPR (~6 kbp) demeure donc un défi en utilisant ces techniques. De plus, les méthodes de manipulation de l'ADN directement dans *E. coli* provoquent régulièrement des événements de recombinaison indésirés (Datsenko & Wanner, 2000). Pour des modules CRISPR à plusieurs ARNg, la recombinaison pourrait engendrer la perte de certains ARNg par recombinaison entre les séquences identiques entre les ARNg. Ce chapitre décrit les efforts fournis afin de développer la méthode DROID et permettre la fusion de TP114 à un module CRISPR. Cette nouvelle méthode permet la manipulation de deux grandes molécules d'ADN chez *E. coli* et peut être appliquée à l'insertion de modules thérapeutiques chez des plasmides conjugatifs pour livraison dans le microbiote, ou à l'insertion de gros fragments d'ADN dans le génome de *E. coli*.

Cet article est le résultat des efforts de quatre personnes. J'ai moi-même fait la plupart des expériences et analyses nécessaires à la publication de ces travaux. J'ai également participé à

l'écriture de l'article en collaboration avec Sébastien Rodrigue. Nancy Allard a participé aux expériences de mise au point de la méthode DROID et aux tests d'efficacité du système. David Jordan a conçu et construit la version thermosensible des plasmides du DROID. Sébastien Rodrigue a supervisé le projet, aidé à sa conception et à la rédaction du manuscrit.

Référence bibliographique : Neil K., Allard N., Jordan D. and Rodrigue S. (2019) Assembly of Large Mobilizable Genetic Cargo by Double Recombinase Operated Insertion of DNA (DROID). *Plasmid*, 104:102419.

3.2 Page titre

Assembly of Large Mobilizable Genetic Cargo by Double Recombinase Operated Insertion of DNA (DROID)

Authors: Kevin NEIL^a, Nancy ALLARD^a, David JORDAN^a, Sébastien RODRIGUE^{a,b}.

^a Department of Biology, University of Sherbrooke, Sherbrooke, QC, J1K 2R1, Canada

^b Centre de recherche du CHUS, Sherbrooke, QC, J1H 5N4, Canada

Declaration of interest:

The work presented in this manuscript is part of US provisional patent application 62/696,367.

Corresponding Author: Sébastien RODRIGUE

E-mail address: Sebastien.Rodrigue@USherbrooke.ca).

Full postal address : Département de biologie, Université de Sherbrooke, 2500 Boulevard de l'Université, Sherbrooke (QC), Canada, J1K 2R1

3.3 Abstract:

There is an important need to develop new therapeutic tools to modulate the gene content of microbiomes. A promising strategy for microbiome engineering relies on the delivery of CRISPR-Cas9 systems by conjugative plasmids, allowing DNA sequence-specific elimination of genes in target bacteria. Yet, fusion of large DNA molecules such as conjugative plasmids and CRISPR-Cas9 systems can be challenging. Here, we describe the Double Recombinase Operated Insertion of DNA (DROID), an efficient method to assemble large DNA constructs. We show that the Bxb1 integrase and FLP recombinase can be used successively to stably insert CRISPR-Cas9 systems in a conjugative plasmid without adding any antibiotic marker. The resulting CRISPR-Cas9 mobilization system was able to cure a multi-copy antibiotic resistance plasmid from a target bacterium. The DROID method facilitates the assembly of delivery vehicles for microbiome editing and is readily adaptable to a multitude of applications that require the manipulation of large DNA molecules.

Keywords: DNA Assembly, Plasmid, conjugation, CRISPR-Cas9, Recombinase.

Highlights:

- Double Recombinase Operated Insertion of DNA (DROID) is a new high-efficiency method to introduce large DNA cassettes into another DNA molecule.
- Plasmids were created to enable delivery of CRISPR-Cas9 systems for microbiome engineering.

Abbreviations:

bp: base pair

Cm: Chloramphenicol

DROID: Double Recombinase Operated Insertion of DNA

EcN: *Escherichia coli* Nissle 1917

Incl2: Incompatibility Group I2

Km: Kanamycin

LB: Luria Broth Miller

LBA: Luria Broth Miller Agar

OD_{600nm}: Optical Density at 600 nanometers

PAM: Protospacer Adjacent Motif

pT: Target plasmid

pNT: Non-Target plasmid

3.4 Introduction

Many estimates suggest that the human body hosts a greater number of micro-organisms than human cells (Sender et al., 2016). These micro-organisms form the microbiota and play key roles in several processes such as nutrient absorption (Ley et al., 2005), host metabolism (Bäckhed et al., 2007), defense against pathogens (Cash et al., 2006), drug processing (Guthrie et al., 2017), etc. However, repeated infections and antibiotic treatments can lead to the accumulation of antibiotic resistance and virulence associated genes (Penders et al., 2013). Furthermore, the gut microbiota is a suspected hot-spot for horizontal gene transfer, hence promoting the dissemination of potentially harmful genes between bacteria (Schaik, 2015). New tools to modulate the composition of the microbiome could offer a powerful avenue to prevent or treat antibiotic resistant infections and dysbiosis.

One promising way to engineer the microbiome is through the delivery of DNA payloads to the microbiota residents using either bacteriophages or conjugative plasmids. For example, phage-based delivery was used to transfer CRISPR-Cas9 systems and eliminate antibiotic resistant pathogens (Bikard et al., 2014; Citorik et al., 2014) or to resensitize them to antibiotics (Yosef et al., 2015). Both bacteriophages and conjugative plasmids present their own pros and cons for DNA delivery. Phage-based transport systems can reach high DNA transfer efficiencies and are very precise tools due to their narrow host range (Villarroel et al., 2016), but present limited DNA payload size (Nurmemmedov et al., 2012) and are susceptible to conditions found in the gut environment (pH, bile salts, proteases) (Nobrega et al., 2016). In contrast, conjugative plasmids are often large DNA molecules that can accommodate DNA payloads of considerable sizes and transfer to a large variety of hosts, provided a system efficient enough for DNA delivery *in situ* is found. Conjugative plasmids therefore have the potential to mobilize larger constructs, such as biosynthetic pathways and complex DNA circuits to supplement deficient functions in an imbalanced microbiota. Yet, efficiently loading large DNA payloads (e.g. CRISPR-Cas9) into phages and conjugative plasmids remains challenging.

Several methods have been described to mediate the association of two or more DNA molecules *in vitro*. For example, Gibson assembly (Gibson et al., 2009), digestion-ligation (Cohen et al., 1973), Golden Gate (Engler et al., 2008), USER-cloning (Geu-flores et al., 2007) are commonly used but the assembly of large DNA molecules often remains challenging. Indeed, the efficiency of these techniques quickly drops with increasing numbers of DNA fragments to assemble, usually limiting them to about 6 DNA fragments. The total size of the plasmid also can decrease cloning efficiency, generally limiting these techniques to constructs of <20 kbp. Furthermore, these protocols often rely on the PCR amplification of several DNA fragments to assemble, each of which are limited by the length of DNA amplifiable by PCR. *In vivo* DNA recombination-based techniques such as Recombineering (Datsenko and Wanner, 2000; Datta et al., 2006), Gene doctoring (Lee et al., 2009), or NO-SCAR (Reisch and Prather, 2015) can delete large DNA fragments. However, the size of the DNA molecule that can be introduced is limited as insertion of DNA fragment larger than 2.5 kbp using these techniques greatly decreases their efficiency (Kuhlman and Cox, 2010). Alternative recombination based techniques for the fusion of large DNA fragments (>4 kbp) such as ALFIRE (Rivero-Muller et al., 2007) exists, but are time-consuming (Zhao et al., 2011) and often rely on counter selections with toxic genes that can quickly mutate, leading to variable recombination efficiencies. Consequently, there is a need for more efficient tools facilitating the manipulation of large DNA molecules such as conjugative plasmids that are often too large (>20 kbp) to conveniently modify using *in vitro* DNA assembly techniques, particularly for the insertion of large (>4 kbp) genetic cargos that often depend on recombineering.

Here, we present the Double Recombinase Operated Insertion of DNA (DROID) method, a novel DNA assembly strategy that allows the size-independent fusion of DNA molecules with high efficiency. DROID takes advantage of two sets of recombination sites to join DNA molecules without accumulating antibiotic markers. We exemplify the DROID assembly method by loading CRISPR-Cas9 systems into TP114, a conjugative plasmid of incompatibility group I2 (IncI2) (Grindley et al., 1972). We also show that the delivery of a CRISPR-Cas9

system by TP114 can cure a specific multicopy resistance plasmid in a target bacterium thus confirming the functionality of genetic payload introduced by DROID.

3.5 Material and Methods

3.5.1. Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used are described in Table S3.1. Cells were typically grown in Luria broth Miller (LB) or on Luria broth agar Miller (LBA) medium supplemented, when needed, with antibiotics or medium additives at the following concentrations: ampicillin 100 µg/mL, chloramphenicol (Cm) 34 µg/mL, diaminopimelic acid (DAP) 57 µg/mL, kanamycin (Km) 50 µg/mL, nalidixic acid 4 µg/mL, spectinomycin 100 µg/mL, streptomycin 50 µg/mL, and tetracycline 15 µg/mL. Conjugation experiments were performed between two modified strains of *Escherichia coli* Nissle 1917 (EcN) on LBA plates without antibiotics at 37°C for 2 hours unless specified otherwise in the figure legend. Cells were then resuspended, diluted and plated by spotting 5 µL of each dilution in triplicate on selective LBA plates. All cultures were grown at 37°C, except for bacteria bearing thermosensitive plasmids (pE-FLP, pKill1, pKill3) which were grown at 30°C. Bacterial cultures used in the experiments were grown for no more than 18 hours.

3.5.2. DNA manipulations

All primers used are described in Table S3.2. Plasmid DNA sequences were deposited in Genbank, accession number can be found in Table S3.1. Plasmid DNA was purified using the EZ10 Spin Column Plasmid DNA Miniprep kit (BIOBASIC) while genomic DNA was purified using Quick-DNA Miniprep kit (ZYMO RESEARCH). DNA fragments for plasmid construction were amplified by PCR using the TransStart PFU Fly DNA polymerase (CIVIC BIOSCIENCE). PCR products were purified using Ampure DNA XP (BECKMAN COULTER) following manufacturer's recommendations. DNA fragments were then pooled and assembled

using NEBuilder Hifi DNA Assembly (NEB). Gibson assemblies were then treated with 0.5 μ L DpnI for 15 minutes at 37°C to digest the DNA template. The mixture was then transformed into chemically competent *E. coli* EC100Dpir+ as described previously (Green and Rogers, 2013).

For clone screening, isolated colonies were picked and inoculated in up to 100 μ L of 5% p/v chelex beads. The mixture was then heated at 55°C for 25 minutes and 100°C for 10 minutes. The supernatant was used as a DNA template for PCR screening using the TaqB (ENZYMATICS) or OneTaq (NEB) DNA polymerases.

3.5.3. Construction of the loading docks pREC1 and pBXB1

The DNA fragments for pREC1 were amplified from pOSIP-TT for the *tetB* gene and pKD4 for the *oriV*_{R6K}. *FRT* and *attB*_{Bxb1} sites were provided within the PCR priming oligonucleotides. Plasmid pBXB1 contains *oriV*_{pMB1}-*bla* (ampicillin resistance gene) from pSB1A3, and the *Bxb1* integrase gene from pRAD-module-G8-C1 (Bonnet et al., 2012). The P_{tetO} promoter controlling the expression of the Bxb1 integrase is constitutive since its repressor (*tetR*) is absent on the plasmid construct. The resulting plasmids maps of pREC1 are shown in Figure 3.1.A and pBXB1 in Figure 3.1.B.

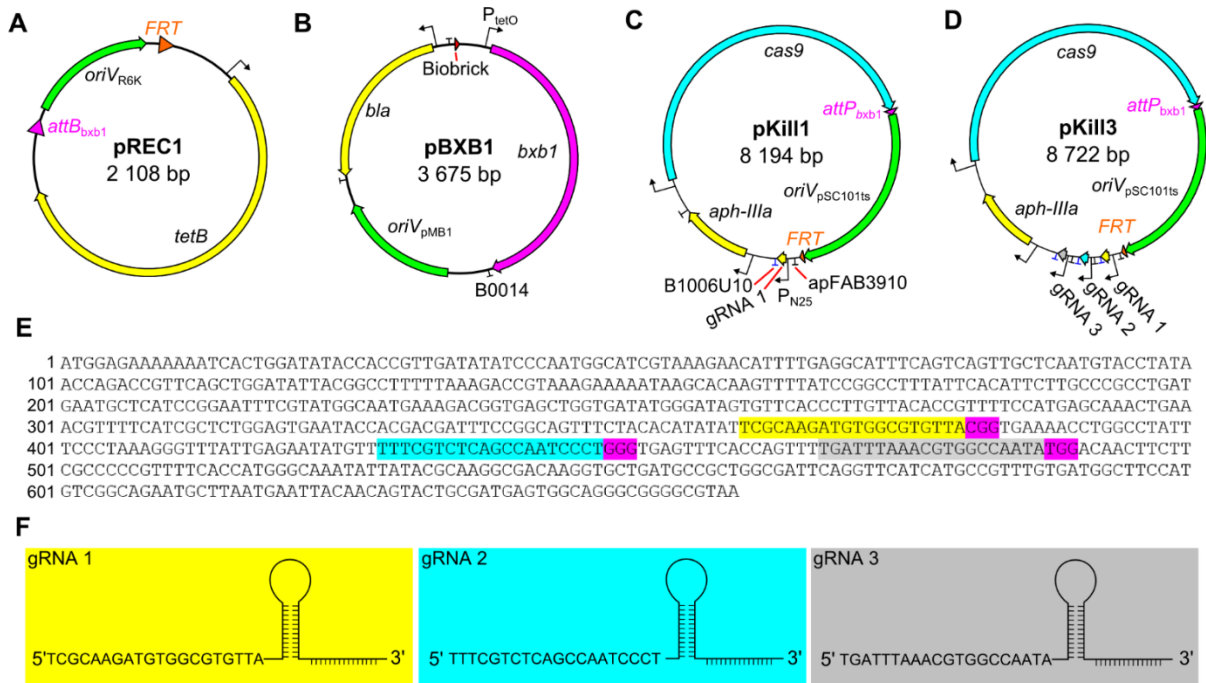


Figure 3.1: Plasmid maps and design. Maps of pREC1 (A), pBxB1 (B), pKill1 (C) and pKill3 (D) are shown to scale. Total size of the plasmid in base pair (bp) is displayed below the plasmid name. Promoters and terminators of the gRNAs from pKill3 are identical to pKill1's gRNA. E, complete coding sequence of the target chloramphenicol resistance gene *cat* with the protospacers highlighted in yellow (gRNA 1), blue (gRNA 2) and gray (gRNA 3) and their corresponding PAM highlighted in purple. F, schematic representation of each gRNA with the sequence of their spacers. The gRNA number and color are conserved between panels.

3.5.4. Construction of the CRISPR-Cas9 modules pKill1 and pKill3

Two CRISPR-Cas9 modules, pKill1 (Figure 3.1.C) and pKill3 (Figure 3.1.D), were created. First, gRNAs were designed to target the Cm resistance gene *cat* (Figure 3.1.E) using DNA 2.0 (ATUM) web-based software with *E. coli* K-12's genome as an off-target sequence. The most potent gRNA spacers (highest score based on the ΔG , off-target frequency) were then analyzed by BLASTn(Altschul et al., 1990) against the EcN donor strain and *Enterobacteriaceae*

genomes to eliminate any candidate gRNAs with high off-targeting (less than 2 mismatches). Three gRNAs (Figure 3.1.F) identified by this bioinformatics strategy were cloned into two different CRISPR-Cas9 system bearing plasmids named pKill1 and pKill3. Plasmids pKill1 and pKill3 were obtained by Gibson assembly, with pKill1 carrying gRNA 1 and pKill3 carrying gRNA 1, 2 and 3. Both plasmids were assembled using a similar strategy combining the *cas9* gene from pCas9, the kanamycin resistance gene *aph-IIIa* from pKD4, and the *oriV_{pSC101ts}* from pGRG36 (Table S3.2). The gRNA 1 was amplified entirely from a previous construct (pKN02), while the spacers of gRNA 2 and 3 were added in PCR primer tails. For pKill3, assembly tags were placed between each gRNAs to prevent miss-assembly due to the repetitive nature of this gene locus. Junctions between DNA segments were created by introducing short sequences in the PCR primers. The integrity of all gRNA sequences was confirmed by Sanger sequencing using the [Plateforme de séquençage et de génotypage des génomes](#) (Université Laval, QC, Canada). The activity of pKill plasmids was assessed by electroporation in EcN modified strain KN03 + pT (Target plasmid) and spread on LBA plates selecting pKill only or pKill + pT supplemented with 1% arabinose. Plasmid pT encodes the targeted Cm resistance gene *cat* and an arabinose inducible *gfp* gene. Plasmid curation can therefore be easily followed in the recipient and transconjugant by fluorescence and Cm resistance loss.

3.5.5. Loading CRISPR-*cas9* into TP114 by DROID

The loading dock consisting of FRT-*tetB*-*attB_{bxb1}* was amplified from pREC1 and inserted in TP114 by recombineering using pSIM7 as previously described (Datta et al., 2006). The modified TP114::*tetB* was first transferred by conjugation at 37°C in *E. coli* MG1655N_x^R to purify it from potential contaminating wildtype TP114. Then, TP114::*tetB* was transferred at 30°C into *E. coli* EC100Dpir⁺ bearing pBXB1 and one of the pKill plasmid. Cells were incubated overnight to allow the Bxb1-mediated insertion of the pKill plasmid into TP114::*tetB*. The resulting plasmids TP114::*tetB*-Kill1 and TP114::*tetB*-Kill3 were then transferred by conjugation at 30°C to *E. coli* MG1655N_x^R. From this point on, the strains were grown at 37°C to avoid any activity of the thermosensitive replication origin (*oriV_{pSC101ts}*) from the integrated

payload that might destabilize TP114::*tetB*-Kill. *E. coli* MG1655Nx^R + TP114::*tetB*-Kill1 or TP114::*tetB*-Kill3 were then transformed by electroporation with pE-FLP(St-Pierre et al., 2013). The final TP114::Kill1 and TP114::Kill3 constructs were finally transferred by conjugation to EcN KN01 Δ *dapA* for assessment of CRISPR-Cas9 activity.

3.5.6. Plasmid pT copy number determination

An overnight culture of KN03 + pT was diluted 1/50 in fresh LB broth and incubated at 37°C. The culture was sampled at an OD_{600nm} value of 0.20, 0.80 and >2.00 to obtain plasmid copy number in early, mid-exponential as well as stationary growth phases. The gDNA of each sample was extracted using the Quick DNA Miniprep Kit (ZYMO RESEARCH) and quantified using a NanoDrop spectrophotometer (THERMO FISHER SCIENTIFIC). All qPCR reactions were performed following the iQ SYBR Green Supermix guidelines (BIO-RAD). The primers used to amplify a portion of the *cat* gene from plasmid pT, and of the housekeeping gene *rpoB*, are listed in Table S3.2. After optimization, plasmid pT copy number was then determined for each test conditions (OD_{600nm} of 0.20, 0.80 and 2.00) as previously described (Anindyajati et al., 2016).

3.5.7. CRISPR-Cas9 module activity test

To test the activity of the CRISPR-Cas9 systems, TP114, TP114::Kill1 and TP114::Kill3 were transferred by conjugation in KN03 + pT. The cells were then plated on selective LB agar medium + 1% arabinose to induce the expression of the *gfp* gene from the targeted plasmid pT. Colony GFP fluorescence was visualized with a Typhoon FLA 9500 (GE HEALTHCARE LIFE SCIENCE). GFP fluorescence was detected using a Low Pass Blue filter and a 473 nm laser while the brightfield image was taken with a Low Pass Red filter and a 635 nm laser. The fluorescent and brightfield images were merged using ImageFiji software (Schindelin et al., 2012). Then, the green fluorescent and non-fluorescent colonies were counted. Plasmid curing was assessed by plating transconjugants on normal or selective media for pT. The same

procedure was performed for the highly similar but untargeted plasmid pNT, which bears an ampicillin resistance gene (*bla*) for selection instead of the *cat* gene.

3.5.8. Statistical analysis

Statistical significance was tested on the logarithmic value of the data by One-way ANOVA. Differences between two test groups were considered significant when the P-value was below 0.05. P-values are noted directly in the figures: ns ($P \geq 0.05$), * ($P < 0.05$), ** ($P < 0.01$) and *** ($P < 0.001$).

3.6 **Results and discussion**

3.6.1. DROID method rationale

DROID uses two recombinases to join two DNA molecules, independently of their respective sizes and without the need of cumulating antibiotic resistance markers. The two DNA molecules to combine during DROID are referred to as the “acceptor DNA molecule” and the “payload”. The acceptor DNA molecule is first equipped with a loading dock, composed of an *attB*_{bxb1} site, a selection marker and an *FRT* site (Figure 3.2.A). The acceptor DNA molecule is then able to recombine with virtually any payload containing a corresponding *attP*_{bxb1} site. To promote recombination, the Bxb1 integrase (Bonnet et al., 2012) is expressed from a helper plasmid (pBXB1) and recognizes the *attB*_{bxb1}/*attP*_{bxb1} pair (Figure 3.2.B). After the integration of the genetic payload, two appropriately located *FRT* sites are recombined with the assistance of the Flp flipase expressed from pE-FLP. This second step removes any unnecessary genes from the final construction (Figure 3.2.C) and deletes the *attL*_{bxb1} site, preventing any excision of the genetic payload.

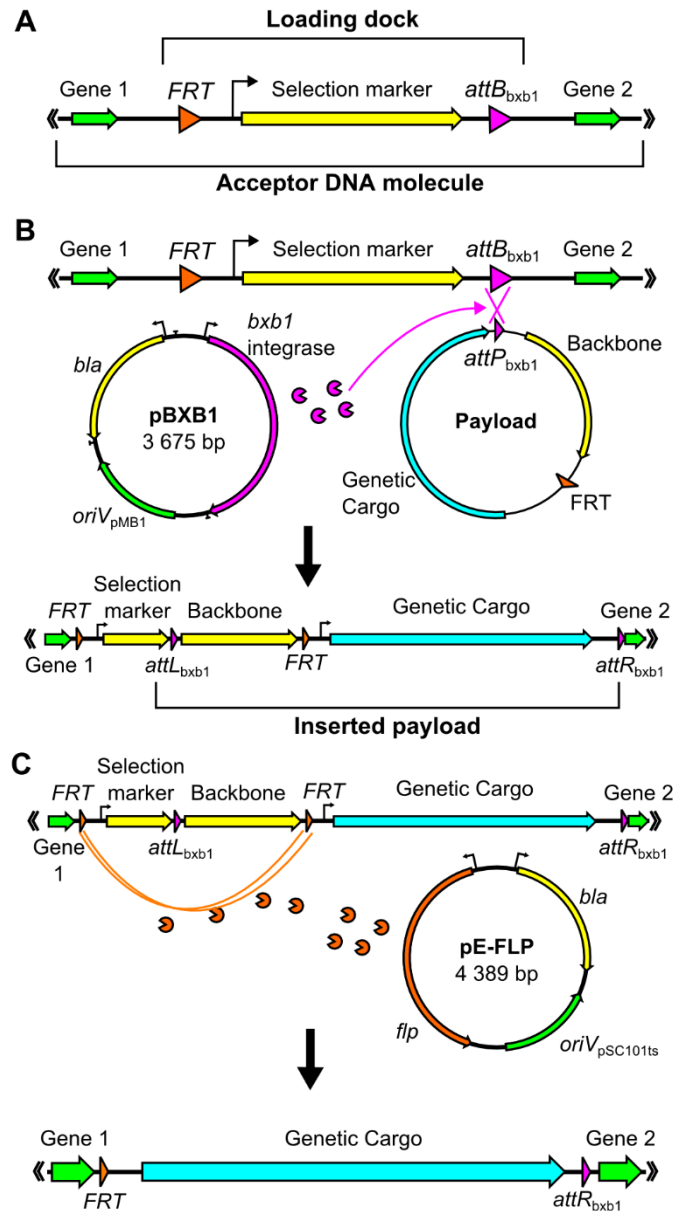


Figure 3.2: Key steps of the DROID method. The DROID method mediates the fusion between an acceptor DNA molecule and a payload. **A**, the acceptor DNA molecule is a large construct that possess a loading dock, composed of an *FRT* site, a selection marker and an $attB_{bxb1}$ site. **B**, a helper plasmid expresses the Bxb1 integrase that recognizes the $attB_{bxb1}$ site on the loading dock and manages the recombination with the $attP_{bxb1}$ site from the payload. After recombination, the complete payload is inserted in the acceptor DNA molecule and is

flanked by an *attL* and *attR* sites formed during the recombination process mediated by Bxb1. C, a helper plasmid then expresses the *flp* recombinase to promote the recombination between two FRT sites and remove all DNA sequences between them, in this case, the acceptor DNA molecule selection marker and the payload backbone. A-C, features are represented as arrows to facilitate the interpretation of insertion directionality.

The DROID method is particularly useful when a genetic payload cannot be efficiently introduced using recombineering based techniques, and the acceptor DNA molecule is too large for easy *in vitro* manipulation. We chose to exemplify this situation by the insertion of a CRISPR-Cas9 system into the IncI2 TP114 conjugative plasmid. CRISPR-Cas9 systems (≥ 6.0 kbp) exceed the length recommended for effective recombineering while the size of TP114 (64,818 bp) is impractical for *in vitro* manipulation. Recombineering is a proficient method for the deletion of large fragments of DNA but its efficiency was shown to drop considerably for the insertion of DNA molecules larger than 2.5 kbp (Kuhlman and Cox, 2010). Considering that each recombineering step usually include a ~ 1 kbp antibiotic resistance gene, only ~ 1.5 kbp would remain available for the genetic payload to preserve good recombineering efficiency. This means that insertion of CRISPR-Cas9 in TP114 by recombineering would likely have to be divided in 5-6 steps, each requiring at least 2-3 days. In contrast, DROID allows the integration of a whole CRISPR-Cas9 system at once and is easily repeatable to allow fast gRNA switching and evaluation.

3.6.2. Insertion of the loading dock in TP114

Since TP114 doesn't include recombination sites essential for the DROID method, we first inserted a loading dock in this acceptor plasmid. The loading dock was designed to contain an *FRT* site, the *tetB* gene and the *attB*_{bxb1} sequence. Plasmid pREC1 was created to contain this loading dock along with the *oriV*_{R6K} replication origin, which maintenance depends on a chromosomally encoded replication initiation protein (Pi) (Figure 3.1.A). This strategy was reported to lower the false positive recombinant rate in *pir*- hosts during the recombineering

procedures (Datsenko and Wanner, 2000). The order and orientation of the recombination sites in the payload and acceptor DNA molecules must be designed carefully. Indeed, once inserted in the *attB_{Bxb1}* site, the payload will provide a *FRT* element that must be correctly aligned with the *FRT* located in the loading dock to allow the deletion of the *tetB* gene by Flp recombination.

The kanamycin resistance gene (*aph-IIIa*) of TP114 was chosen as the insertion point for the loading dock. The loading dock was amplified from pREC1 and inserted by recombineering into TP114, generating TP114::*tetB* (Figure 3.3.A). The replacement of *aph-IIIa* by the loading dock *tetB* was confirmed by PCR (Figure 3.3.B), and had no impact on TP114 conjugation efficiency (Figure 3.3.C).

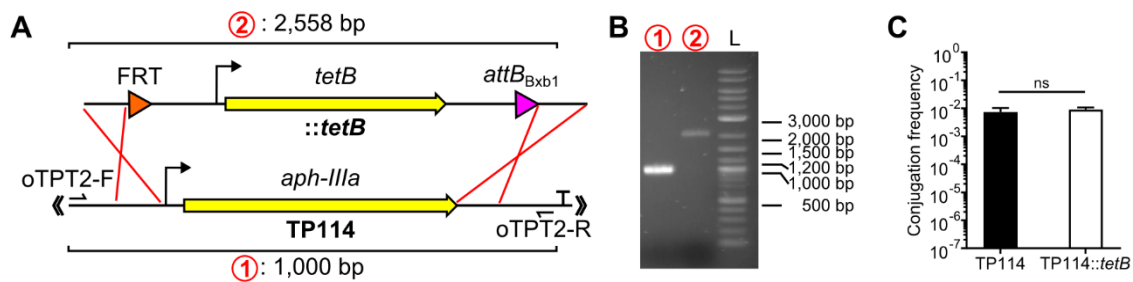


Figure 3.3: Loading dock insertion have no negative effect on TP114 conjugation. **A**, schematic representation of the insertion of the loading dock in the *aph-IIIa* kanamycin resistance gene of TP114. Screening primers oTPT2-F and oTPT2-R are shown with the predicted size of amplified fragments for both TP114 and TP114::*tetB*. **B**, PCR screening for the insertion of the loading dock. Lane 1 is a TP114 negative control while lane 2 is a positive clone of TP114::*tetB*, lane L is a 1-Kb Plus DNA ladder (NEB). Lane numbers can be associated with the predicted size fragments depicted in panel A. **C**, conjugation rate after 2 hours at 37°C on LBA, as a function of the recipient bacteria CFU for both TP114 and TP114::*tetB* (n=3). Statistical significance is indicated as follows: ns (P>0.05), * (P<0.05), ** (P<0.01) or *** (P<0.001). Bars represent the average value of the data while error bars show standard deviation.

3.6.3. Engineering of two functional CRISPR-Cas9 systems

Two CRISPR-Cas9 systems were assembled to demonstrate the DROID method's ability to quickly introduce DNA payloads in TP114::*tetB*. Plasmid pKill1 (Figure 3.1.C) contains the *cas9* gene along with one gRNA (gRNA #1) whereas plasmid pKill3 (Figure 3.1.D) includes three gRNAs (gRNA #1, 2 and 3) instead. The use of pKill1 and pKill3 was interesting to determine if DNA payloads are prone to recombination during the DROID procedure since pKill3 contains three gRNAs composed of identical promoter and terminator sequences. All gRNAs were designed to target the Cm resistance gene *cat* (Figure 3.1.E and F). Both pKill1 and pKill3 are otherwise identical, share the same selection marker, and were designed to facilitate the insertion of the CRISPR-Cas9 systems in TP114 by DROID. Importantly, pKill1 and pKill3 display impaired replication at $\geq 37^{\circ}\text{C}$ (*oriV_{pSC101ts}*), preventing possible instability at this non-permissive temperature after their insertion in TP114::*tetB*. Also, these CRISPR-Cas9 encoding plasmids carry the *attP_{bxb1}* and the *FRT* sites that are located to allow the precise excision of the *oriV_{pSC101ts}* after insertion in TP114.

Plasmid pT was next used to measure the activity of the CRISPR-Cas9 system by transformation of pKill1 and pKill3 into *E. coli* EC100D*pir*⁺ bearing pT (Figure 3.4). Plasmid pT includes the targeted *cat* gene in addition to an inducible *gfp* gene, and is present at high copy numbers (7-140 copies/cell), thus requiring robust expression of the CRISPR-Cas9 system for its curing. Transformation efficiency dropped drastically when selecting for the presence of both pT and the pKill plasmid in the same cells. All pKill transformants without selection of pT showed no fluorescence, suggesting that these clones had all lost pT. Furthermore, subculturing of these non-fluorescent colonies (n=24) in medium containing Cm showed no signs of growth, indicating Cm sensitivity. Taken together, these results indicate that the CRISPR-Cas9 system of pKill1 and pKill3 plasmids cleaves the targeted protospacers located on pT, leading to loss of the Cm resistance through plasmid curation.

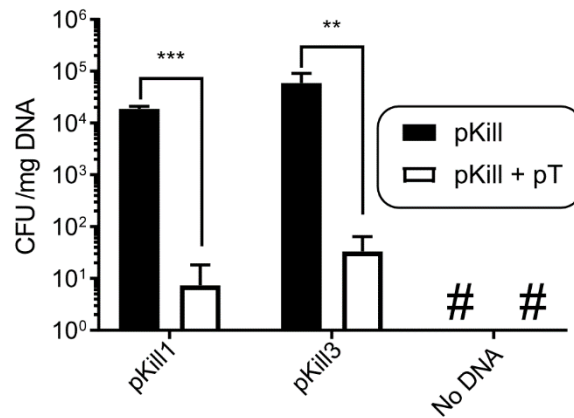


Figure 3.4: The CRISPR-Cas9 systems plasmids pKill1 and pKill3 specifically cleave their target sequence. The number of pKill3 and pKill1 transformants was quantified in EcN KN03 bearing pT or not (n=3). Selection of pKill transformants with and without co-selection for pT, a targeted chloramphenicol resistance plasmid. # denotes the absence of transformant colonies. Statistical significance is indicated as follows: ns (P>0.05), * (P<0.05), ** (P<0.01) or *** (P<0.001). Bars represent the average of the values with the error bar showing the standard deviation.

3.6.4. Loading the CRISPR-cas9 systems on TP114::tetB

We next sought to insert the CRISPR-*cas9* systems pKill1 and pKill3 in the DROID-ready TP114::*tetB*. This step of the DROID method is mediated by the Bxb1 integrase, which is expressed from the pBXB1 helper plasmid. TP114::*tetB* and one of the pKill plasmids were introduced in *E. coli* EC100Dpir⁺ containing pBXB1 to catalyze their fusion (Figure 3.5.A). To quantify the efficiency of the insertion step managed by Bxb1, the insertion of pKill in TP114::*tetB* was repeated in biological triplicates for both pKill1 and pKill3. The insertion of the pKill plasmids in TP114::*tetB* was confirmed by PCR screening (Figure 3.5.B). A total of 46 clones of EC100Dpir⁺ bearing TP114::*tetB*, pBXB1 and a pKill plasmid were screened by PCR for insertion of the pKill plasmid in TP114 (Figure 3.5.C). Although both TP114::*tetB* + pKill and TP114::*tetB*-Kill have the same antibiotic resistances, thereby preventing the selection

for the insertion, most of the clones obtained showed a positive pattern for pKill insertion (27/46) and very few clones lacked a pKill insertion (4/46). These results demonstrate that Bxb1 can very efficiently recognize the *attB_{bxb1}*/*attP_{bxb1}* pair and operate DNA fusion even in the absence of antibiotic selection pressure. Surprisingly, a high proportion of clones (15/46) contained a mixed population of TP114::*tetB* and TP114::*tetB*-Kill. The presence of these mixed population of TP114::*tetB* lacking or not the Kill insert in the same clonal colonies might indicate that in those cells, TP114::*tetB* most probably was present in higher copy number than the pKill plasmid when the Bxb1 mediated fusion reaction took place. As the presence of uninserted pKill plasmid was not followed, it is also unclear if the *bxb1* integrase might have suffered from inactivating mutations throughout the insertion process, which could also explain the mixed population observed. Also, the insertion of pKill1 seemed to be less efficient than the insertion of pKill3. This difference might be due to a higher frequency of inactivating mutations in either the *bxb1* integrase gene or most probably in the *attP_{bxb1}* site on pKill1 specifically. Nevertheless, 42/46 clones contained at least one copy of TP114::*tetB*-Kill and could therefore be used in further steps of the DROID method.

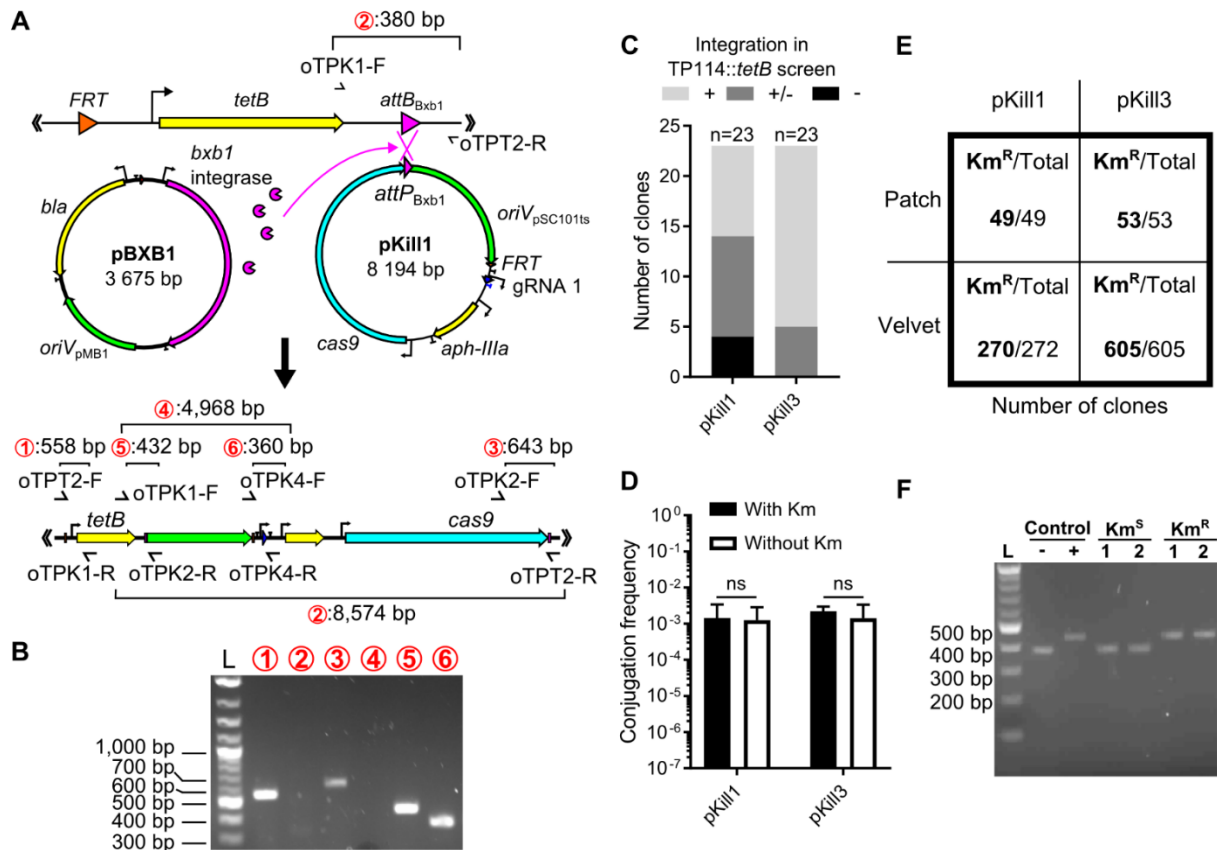


Figure 3.5: Bxb1 allows efficient manipulation of large DNA molecules. **A**, schematic representation of pKill1 insertion. Screening primers are shown with the predicted length of their amplicon. **B**, PCR screening of pKill1 integrity after insertion in TP114. Amplicon numbering is consistent with panel A, lane L is a 1-Kb Plus DNA ladder (NEB). **C**, EC100D*pir*+ pKill + pBXB1 + TP114::tetB colony PCR screening of pKill insertion in TP114::tetB using primers oTPK1-F, oTPK2-R and oTPT2-R. Clones presenting an amplicon corresponding to the integrated form of pKill were considered positive for pKill integration (+) while clones presenting the PCR pattern corresponding to TP114::tetB without the pKill insert were considered negative (-). Some clones displayed both the positive and negative PCR amplicon and were thus assigned to the +/- group. **D**, frequency of TP114::tetB (without Km selection) and TP114::tetB-Kill (with Km selection) transconjugants after a transfer from EC100D*pir*+ to MG1655Nx^R (n=3). Statistical significance is indicated as follows: ns (P>0.05), * (P<0.05), ** (P<0.01) or *** (P<0.001). Bar represent the means of the data with error bars being the standard

deviation **E**, TP114::*tetB* transconjugant colonies from plates without Km were screened for presence of the pKill insert by replica plating (velvet) and colony patching (patch) on medium with and without Km. **F**, PCR screening of transconjugant colonies found to be Km^S or Km^R. Negative control was TP114::*tetB* purified DNA while the positive control was TP114::*tetB*-Kill1 purified DNA, lane L is a 1-Kb Plus DNA ladder (NEB).

One clone per replicate of the pKill1 and pKill3 insertion in TP114::*tetB* was randomly selected and transferred by conjugation into *E. coli* MG1655Nx^R. Since the pKill plasmids encode a kanamycin resistance gene, their integration in TP114::*tetB* should lead to the mobilization of *aph-III*. Therefore, transconjugants were plated on LBA with or without kanamycin as a first assessment of pKill integration efficiency (Figure 3.5.D). Surprisingly, no significant difference was observed in the presence or absence of Km selection for pKill, thus revealing a high proportion of TP114::*tetB*-Kill in the transconjugant population although some of the donor clones contained a mixed population of TP114::*tetB* with or without the Kill insert. We further investigated this result by subculturing transconjugant colonies recovered from LBA plates without Km onto solid medium supplemented with Km by replica plating and colony patching. Out of the 879 colonies replicated and 102 colonies patched, only 2 were unable to grow on Km plates selecting for pKill (Figure 3.5.E). To confirm that Km resistance could be used to select for TP114::*tetB*-Kill, 41 Km resistant clones and the 2 Km sensitive clones were screened by PCR. All Km resistant clones showed the amplification product associated with the integrated form of pKill while the 2 Km sensitive clones were negatives (Figure 3.5.F). The discrepancy between the results in figure 3.5C and figures 3.5DEF might be due to the cultivations steps preceding the transfer of TP114 to *E. coli* MG1655Nx^R. The overnight incubation at 37°C of clones in LB selecting for both Tc and Km resistance might have favored the growth of daughter cells containing the TP114::*tetB*-Kill construct. Also, as PCR is not a quantitative technique, the presence of a mixed population of amplicon after PCR doesn't imply that these clones carried equal proportions of both constructs. Taken together, these results support the observation that the Bxb1 integration step in the DROID method mediates high-efficiency merging of two DNA molecules, without the need for selection nor extensive screening.

3.6.5. FLP mediated deletion of antibiotics markers

The last step of the DROID method is to use properly located *FRT* sites to eliminate any unwanted DNA from the final plasmid construct. In the present case, the TP114::*tetB*-Kill construction contains two antibiotic selection markers and two origins of replication. As only one of each is required for full plasmid functionality, *FRT* sites were designed to eliminate the *tetB* resistance gene and *oriV_{pSC101ts}*. *FRT* recombination is performed by the Flp recombinase expressed from helper plasmid pE-FLP (Figure 3.6.A). Transformation of pE-FLP led to loss of tetracycline resistance in all screened colonies (n=12). The loss of both *tetB* and *oriV_{pSC101ts}* was also verified by PCR (Figure 3.6.B). The high efficiency of pE-FLP mediated *FRT* recombination in TP114::*tetB*-Kill is in good agreement with previous reports (St-Pierre et al., 2013). The integrity of gRNAs in TP114::Kill1 and TP114::Kill3 final constructs was next investigated by Sanger sequencing and revealed no mutation nor recombination, hereby showing that the DROID method doesn't promote recombination event within the DNA payload. Both TP114::Kill1 and TP114::Kill3 were next transferred in KN01Δ*dapA* for assessment of CRISPR-Cas9 functionality.

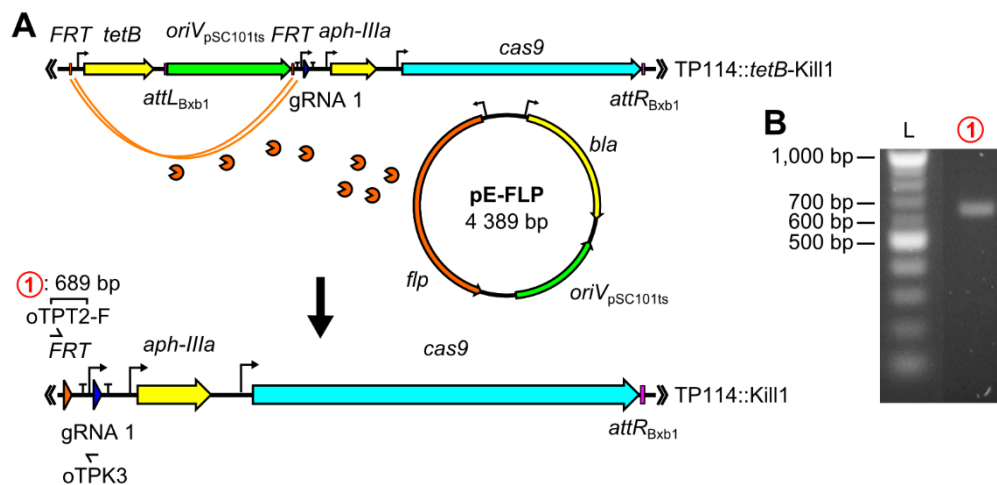


Figure 3.6: The FRT recombination by the FLP flipase removes any unnecessary genes from the final construction. A, schematic representation of FLP mediated recombination of *FRT* sites. Primers used for PCR screening are shown with their predicted amplicon size **B**, PCR

screening of FRT recombination using primers identified in panel A. Size of the amplified fragment is consistent with the amplicon of panel A, lane L is a 1-Kb Plus DNA ladder (NEB).

3.6.6. TP114 can perform in-cis mobilization of CRISPR-Cas9 systems

Both TP114::Kill1 and TP114::Kill3 can mobilize a CRISPR-Cas9 system to cleave the Cm resistance gene *cat*. To assess the functionality of the DNA payload after insertion in the acceptor DNA molecule by DROID, TP114, TP114::Kill1 and TP114::Kill3 were transferred by conjugation in a recipient strain (KN03) bearing plasmid pT. Curation was followed both by counting CFU on selective medium (Figure 3.7.A) and loss of GFP fluorescence (Figure 3.7.B). Transfer of TP114::Kill1 and TP114::Kill3 led to plasmid pT curation in most transconjugants. Moreover, transconjugants obtained using wild-type TP114 retained plasmid pT, indicating that the CRISPR-Cas9 systems were responsible for plasmid loss. Importantly, to verify the specificity of the CRISPR-Cas9 systems to the Cm resistance gene, the same experiment was repeated using plasmid pNT. This plasmid is highly similar to pT but encodes for ampicillin resistance instead of Cm resistance. TP114::Kill3, which is expected to show the highest off-targeting activity, was tested for its ability to cure pNT (Figure 3.7.C). No difference was found in the number of transconjugants when selecting or not for pNT. This suggests that TP114::Kill3 had no off-targeting activity on pNT. Taken together, these results show that the CRISPR-Cas9 systems introduced in TP114 can be mobilized *in cis* while maintaining their activity and specificity.

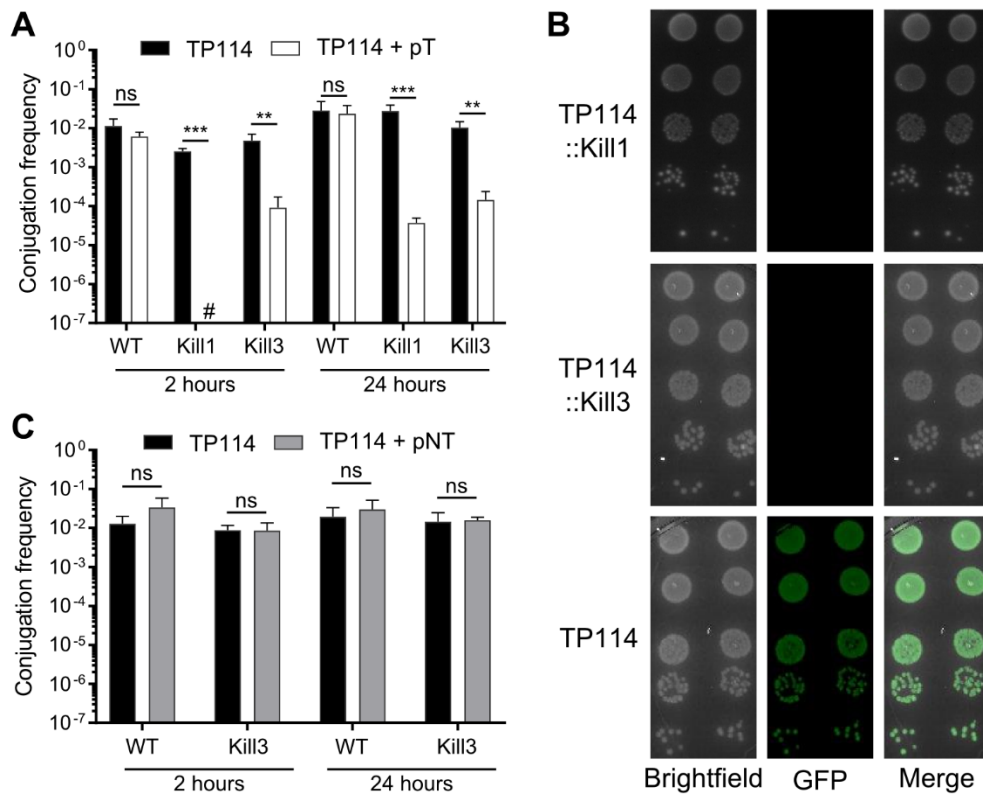


Figure 3.7: TP114 can deliver functional CRISPR-Cas9 systems *in cis* to specifically cure an antibiotic resistance plasmid. **A**, conjugation of TP114 (WT), TP114::Kill1 (Kill1) and TP114::Kill3 (Kill3) from KN01 Δ *dapA* to KN03 + pT. Transconjugants were selected for TP114 acquisition only (black) or co-selected for TP114 and pT (white). # denotes the absence of transformant colonies. **B**, representative colony picture of TP114::Kill1, TP114::Kill3 and TP114 transconjugants after 24 hours of conjugation towards KN03 + pT without selection for pT. Plasmid pT is producing green fluorescence. **C**, conjugation of TP114 (WT) and TP114::Kill3 towards KN03 + pNT. Plasmid pNT is very similar to plasmid pT but does not include the targeted *cat* gene. Transconjugants were selected for TP114 acquisition only (black) or co-selected for TP114 and pNT (gray). **A-C**, Conjugation was performed over 2 and 24 hours for all experiments with at least three independent biological replicates. Bars indicate the average value with error bars representing the standard deviation. Statistical significance is indicated as follows: ns ($P > 0.05$), * ($P < 0.05$), ** ($P < 0.01$) or *** ($P < 0.001$).

3.6.7. DROID is a potent DNA assembly method that is readily adaptable to multiple applications

The DROID method efficiently joined two DNA molecules of 64 kbp and 8 kbp without the need for selection and extensive screenings. In our hands, the overall procedure took 14 days to complete (Table 3.1) and required minimal screening as each step produced a majority of clones carrying the expected product. Importantly, once the acceptor DNA molecule contains a loading dock, the procedure can be completed in as little as 7 days. Also, each step requires very little manipulation (~7 hours of labor over 7 days). As such, the insertion of several different payloads in the acceptor DNA molecule could be performed in parallel.

While the DROID offers several advantages such as high efficiency, no DNA size limitation, and no undesired mutations resulting from the stimulation of recombination, the method also has certain limitations. First, DROID might not be advantageous to insert DNA payload < 5 kbp as most recombineering techniques could be completed faster with sufficient efficiency. Also, the current protocol of DROID implies that the acceptor DNA molecule must be mobilizable. While it is possible to mobilize very large DNA molecules (e. g. BACs), some applications such as the insertion of DNA in the chromosome of a bacterium cannot be performed with the current set of plasmids. To address this limitation, we generated a thermosensitive version of pBXB1 (pBXB1ts) to allow the loss of pBXB1ts by heat shock as opposed to conjugative transfer after the first step of the DROID. Furthermore, the procedure employs three antibiotic resistance genes to select for each component. In the present example, we used ampicillin resistance for pBXB1 and pE-FLP, tetracycline resistance for the loading dock and kanamycin resistance for the payload. We designed an alternative loading dock (pREC2) encoding chloramphenicol resistance instead of tetracycline to facilitate manipulation of systems bearing multiple resistance genes. Loading dock pREC1 was also designed to allow applications such as marker-free chromosomal insertions, which required minor changes to the recombination site placement (Figure 3.8). One last limitation implies the use of homologous recombination for the insertion of the loading dock. Consequently, the insertion of the loading dock must be accomplished in a cell that can efficiently perform homologous recombination.

The DROID method could be further improved through many minor adjustments. First, as both recombination events present >99% efficiency, performing both recombination steps at the same time could further accelerate the procedure. This modification could allow completion of DROID in less than 5 days. Importantly, the usage of two or more orthologous *attB_{bxb1}/attP_{bxb1}* sites could enable recursive insertions of payloads in a same acceptor DNA molecule. Such iterative DNA assembly could be used to assemble more complex DNA constructs like large metabolic pathways and even complete chromosomes. The assembly of complete chromosomes would however require careful design and the repartition of the whole chromosome into clonable DNA fragments flanked with the recombination sites. This assembly would be completed by several iterations until the whole chromosome would be assembled.

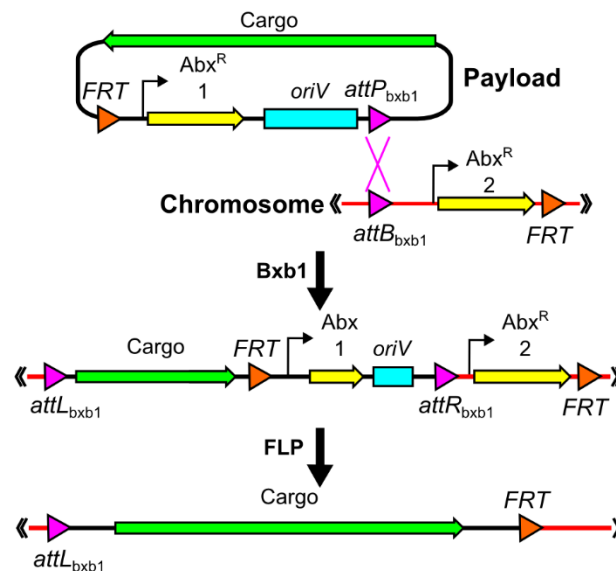


Figure 3.8: Marker-free insertion of large payloads in the chromosome. A loading dock can be inserted in the chromosome of a bacteria. The chromosome of the bacteria is then considered as the acceptor DNA molecule and can be used to insert different payloads. The *FRT* and *attB_{bxb1}/attP_{bxb1}* sites are placed so once the payload is inserted in the chromosome, both *FRT* sites can recombine and eliminate all antibiotics selection markers as well as the *oriV* from the payload. Number 1 and 2 are used to distinguish between two different genes sharing the same

function. Abbreviations and designations: Abx^{R} , antibiotics resistance gene; *oriV*, origin of replication.

Table 3.1: DROID protocol summary.

Main Steps	Insertion of the loading dock by recombineering	Bxb1 mediated insertion of the payload	FLP construction cleaning
Days	Day 1 to 5	Day 5 to 11	Day 11 to 14
Protocol	<p>Transform the pSIM plasmid in an <i>E. coli</i> strain bearing the acceptor DNA molecule. Incubate cells overnight at 30°C. (Day 1)</p> <p>Parallely, amplify the loading dock from pREC1 by PCR (x8), pool and purify the products by SPRI. (Day 1)</p> <p>Subculture a colony of transformants and incubate overnight at 30°C. (Day 2)</p> <p>Perform recombineering as previously described (Datsenko and Wanner, 2000; Datta et al., 2006) (Day 3)</p> <p>Let cells recuperate overnight at 30°C before plating on LBA selecting the loading dock insertion. (Day 3)</p> <p>Confirm insertion by colony PCR and subculture positive clones for frozen stock. Proceed with the insertion of the payload. (Day 4-5)</p>	<p>Transfer by conjugation the acceptor DNA molecule to the <i>E. coli</i> strain bearing pBxB1. Incubate the transconjugants selecting plate overnight at 30°C. (Day 5)</p> <p>Pick a colony and subculture it in broth overnight at 30°C. (Day 6)</p> <p>Transform the payload in an <i>E. coli</i> strain with pBxB1 and the acceptor DNA molecule. Incubate the plate overnight at 30°C. (Day 7)</p> <p>Pick a colony to inoculate in broth and incubate the culture overnight at 30°C. (Day 8)</p> <p>Transfer the acceptor DNA molecule to a plasmid free <i>E. coli</i> strain, selecting for transconjugants bearing the selection marker from the payload. Incubate transconjugant selecting plates overnight at 37°C. (Day 9)</p> <p>Pick a colony for frozen stock generation. Proceed with the construction cleaning. (Day 10-11)</p>	<p>Transform the <i>E. coli</i> strain bearing the assembly between the acceptor DNA molecule and the payload with pE-FLP. Incubate selective plates overnight at 30°C. (Day 11)</p> <p>Inoculate a colony in broth and verify by patching that it has lost the unwanted selective markers. Incubate the cultures overnight at 30°C. (Day 12)</p> <p>Transfer the final construction to the strain of interest for payload mobilization. Incubate the transconjugants plate overnight at 37°C. (Day 13)</p> <p>Subculture a single colony in broth for frozen stock. Screening for insert integrity is recommended at this step. (Day 14)</p>
Approximate hands-on time	10 hours	4 hours	3 hours

3.7 Concluding remarks

The DROID method enables rapid, stable and size independent fusion of two DNA molecules. Here, we used DROID to rapidly insert CRISPR-Cas9 systems into a conjugative plasmid, and showed that the resulting molecule can serve as a delivery vehicle to efficiently suppress antibiotic resistance in targeted cells. As such, DROID facilitates to the development of microbiome engineering tools. Other types of genetic payloads such as genes encoding metabolic functions or anti-microbial peptides could be introduced in conjugative plasmids using DROID. This could help supplement the functions of the microbiota and potentially contribute to the treatment of dysbiosis. Finally, DROID could be readily adapted for other applications in synthetic biology that would require the assembly of two or more large DNA molecules.

3.8 Acknowledgments

This work was supported by the Canadian Institutes of Health Research (grant no. 159817) and by internal funds from the Université de Sherbrooke. S.R. holds a Chercheur boursier Junior 2 award from the *Fonds de recherche du Québec – Santé* (FRQS). K.N. acknowledges graduate research fellowships from the *Fonds de recherche du Québec – Nature et technologies* (FRQNT) and from the Natural Science and Engineering Research Council of Canada (NSERC). N.A. is supported by a doctoral scholarship from the Université de Sherbrooke. We thank Rodrigue lab members for thoughtful discussions and Alain Lavigueur for comments on this manuscript.

3.9 References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)

Anindyajati, Anita Artarini, A., Riani, C., Retnoningrum, D.S., 2016. Plasmid copy number determination by quantitative polymerase chain reaction. *Sci. Pharm.* 84, 89–101. <https://doi.org/10.3797/scipharm.ISP.2015.02>

Bäckhed, F., Manchester, J.K., Semenkovich, C.F., Gordon, J.I., 2007. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc. Natl. Acad. Sci.* 104, 979–984. <https://doi.org/10.1073/pnas.0605374104>

Bikard, D., Euler, C., Jiang, W., Nussenzweig, P.M., Gregory, W., Duportet, X., Fischetti, V.A., Marraffini, L.A., 2014. Development of sequence specific antimicrobials based on programmable CRISPR-Cas nucleases. *Nat. Biotechnol.* 32, 1146–1150. <https://doi.org/10.1038/nbt.3043.Development>

Bonnet, J., Subsoontorn, P., Endy, D., 2012. Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proc. Natl. Acad. Sci.* 109, 8884–8889. <https://doi.org/10.1073/pnas.1202344109>

Cash, H.L., Whitham, C. V, Behrendt, C.L., Hooper, L. V, 2006. Symbiotic bacteria direct expression of an intestinal bacterial Lectin. *Science.* 313, 1126–1130. <https://doi.org/10.1126/science.1127119.Symbiotic>

Citorik, R.J., Mimee, M., Lu, T.K., 2014. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat. Biotechnol.* 32, 1141–1145. <https://doi.org/10.1086/498510.Parasitic>

Cohen, S.N., Chang, A.C.Y., Boyert, H.W., Helling, R.B., 1973. Construction of Biologically Functional Bacterial Plasmids *In Vitro*. *PNAS* 70, 3240–3244.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* 97, 6640–6645.

Datta, S., Costantino, N., Court, D.L., 2006. A set of recombineering plasmids for gram-negative bacteria. *Gene* 379, 109–115. <https://doi.org/10.1016/j.gene.2006.04.018>

Engler, C., Kandzia, R., Marillonnet, S., 2008. A One Pot , One Step , Precision Cloning Method with High Throughput Capability. *PLoS One* 3, e3647. <https://doi.org/10.1371/journal.pone.0003647>

Geu-flores, F., Nour-eldin, H.H., Nielsen, M.T., Halkier, B.A., 2007. USER fusion : a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.* 35, e55. <https://doi.org/10.1093/nar/gkm106>

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. <https://doi.org/10.1038/nmeth.1318>

- Green, R., Rogers, E.J., 2013. Chemical Transformation of *E. coli*. *Methods Enzymol.* 529, 329–336. <https://doi.org/10.1016/B978-0-12-418687-3.00028-8>.Chemical
- Grindley, N.D.F., Grindley, J.N., Anderson, E.S., 1972. R Factor Compatibility Groups. *Mol. Gen. Genet.* 119, 287–297.
- Guthrie, L., Gupta, S., Daily, J., Kelly, L., 2017. Human microbiome signatures of differential colorectal cancer drug metabolism. *Biofilms and Microbiomes* 3, doi:10.1038/s41522-017-0034-1. <https://doi.org/10.1038/s41522-017-0034-1>
- Kuhlman, T.E., Cox, E.C., 2010. Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Res.* 38, e92. <https://doi.org/10.1093/nar/gkp1193>
- Lee, D.J., Bingle, L.E.H., Heurlier, K., Pallen, M.J., Penn, C.W., Busby, S.J.W., Hobman, J.L., 2009. Gene doctoring : a method for recombineering in laboratory and pathogenic *Escherichia coli* strains. *BMC Microbiol.* 9, 1–14. <https://doi.org/10.1186/1471-2180-9-252>
- Ley, R.E., Backhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D., Gordon, J.I., 2005. Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci.* 102, 11070–11075. <https://doi.org/10.1073/pnas.0504978102>
- Nobrega, F.L., Costa, A.R., Santos, J.F., Siliakus, M.F., Lent, J.W.M. Van, Kengen, S.W.M., Azeredo, J., Kluskens, L.D., 2016. Genetically manipulated phages with improved pH resistance for oral administration in veterinary medicine. *Sci. Rep.* 6, 1–12. <https://doi.org/10.1038/srep39235>
- Nurmemmedov, E., Castelnovo, M., Medina, E., Catalano, C.E., Evilevitch, A., 2012. Challenging Packaging Limits and Infectivity of Phage λ . *J. Mol. Biol.* 415, 263–273. <https://doi.org/10.1016/j.jmb.2011.11.015>
- Penders, J., Stobberingh, E.E., Savelkoul, P.H.M., Wolffs, P.F.G., 2013. The human microbiome as a reservoir of antimicrobial resistance. *Front. Microbiol.* 4, 1–7. <https://doi.org/10.3389/fmicb.2013.00087>
- Reisch, C.R., Prather, K.L.J., 2015. The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Sci. Rep.* 5, 15096. <https://doi.org/10.1038/srep15096>
- Rivero-Muller, A., Lajic, S., Huhtaniemi, I., 2007. Assisted large fragment insertion by Red / ET-recombination (ALFIRE)— an alternative and enhanced method for large fragment recombineering. *Nucleic Acids Res.* 35, e78 doi:10.1093/nar/gkm250. <https://doi.org/10.1093/nar/gkm250>
- Schaik, W. Van, 2015. The human gut resistome. *Phil. Trans. R. Soc. B* 370, 1–9.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. <https://doi.org/10.1038/nmeth.2019>

Sender, R., Fuchs, S., Milo, R., 2016. Are We Really Vastly Outnumbered ? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* 164, 337–340. <https://doi.org/10.1016/j.cell.2016.01.013>

St-Pierre, F., Cui, L., Priest, D.G., Endy, D., Dodd, I.B., Shearwin, K.E., 2013. One-step cloning and chromosomal integration of DNA. *ACS Synth. Biol.* 2, 537–541. <https://doi.org/10.1021/sb400021j>

Villarroel, J., Kleinheinz, K.A., Jurtz, V.I., Zschach, H., Lund, O., Nielsen, M., Larsen, M.V., 2016. HostPhinder: A Phage Host Prediction Tool. *Viruses* 6, 1–22. <https://doi.org/10.3390/v8050116>

Yosef, I., Manor, M., Kiro, R., Qimron, U., 2015. Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. *Proc. Natl. Acad. Sci.* 112, 7267–7272. <https://doi.org/10.1073/pnas.1500107112>

Zhao, Y., Wang, S., Zhu, J., 2011. Original Article A multi-step strategy for BAC recombineering of large DNA fragments. *Int J Biochem Mol Biol* 2, 199–206.

3.10 Supplementary Data

Table S3.1: List of strains and plasmid used in this study

Strain or plasmid	Relevant phenotype or genotype	Source/Reference
<i>E. coli</i>		
EC100Dpir+	F ⁻ mcrA Δ(mrr-hsdRMS-mcrBC) φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ ⁻ rpsL nupG pir ⁺ (DHFR)	#ECP09500 (Lucigen)
KN01ΔdapA	ΔdapA, Sp ^R , Sm ^R , Nissle1917	Neil <i>et al.</i> 2020
KN03	Sm ^R Tc ^R Nissle 1917	Neil <i>et al.</i> 2020
MG1655Nx ^R	Nx ^R K-12 F ⁻ λ ⁻ ilvG ⁻ rfb-50 rph-1	(Carraro <i>et al.</i> 2014)
Nissle 1917	Wildtype probiotics strain	DSM-6601 (DSMZ)
Plasmid		
pBXB1	pMB1 oriV, <i>bxb1</i> integrase, <i>bla</i> (Ap ^R)	This study (Genbank: MK756311)
pBXB1ts	pSC101ts oriV, <i>bxb1</i> integrase, <i>bla</i> (Ap ^R)	This study (Genbank: MK756316)
pCas9	<i>cas9</i> , <i>cat</i> , <i>oriV</i> _{p15A} , tracrRNA	Addgene #42876
pE-FLP	<i>oriV</i> _{pSC101ts} , <i>flp</i> , <i>bla</i> (Ap ^R)	Addgene #45978
pGRG36	<i>oriV</i> _{pSC101ts} , Tn7 insertion machinery, <i>araC</i> , Ap ^R	Addgene #16666
pKD4	<i>oriV</i> _{R6K} , <i>FRT</i> flanked Kn ^R , Ap ^R	Addgene #45605
pKill1	<i>oriV</i> _{pSC101ts} , <i>attP</i> _{bxb1} , <i>FRT</i> , 1 gRNA vs <i>cat</i> , <i>aph-IIIa</i> (Km ^R), <i>cas9</i>	This study (Genbank: MK756312)
pKill3	<i>oriV</i> _{pSC101ts} , <i>attP</i> _{bxb1} , <i>FRT</i> , 3 gRNA vs <i>cat</i> , <i>aph-IIIa</i> (Km ^R), <i>cas9</i>	This study (Genbank: MK756313)
pKN02	<i>oriT</i> _{RP4} , <i>oriV</i> _{R6K} , <i>cas9</i> , <i>bla</i> , gRNA against <i>cat</i>	This study (Genbank: MK756314)
pNT	<i>oriV</i> _{pMB1} , <i>bla</i> (Ap ^R)	pSB1A3 (IGEM)
pOSIP-TT	<i>tetB</i>	Addgene #45989
pREC1	<i>oriV</i> _{R6K} , <i>attB</i> _{bxb1} , <i>tetB</i> (Tc ^R), <i>FRT</i>	This study (Genbank: MK756315)
pREC2	<i>oriV</i> _{R6K} , <i>attP</i> _{bxb1} , <i>cat</i> (Cm ^R), <i>FRT</i>	This study (Genbank: MK756317)
pSIM6	<i>oriV</i> _{pSC101ts} , Lambda Red recombinase, <i>bla</i> (Ap ^R)	(Datta <i>et al.</i> , 2006)
pT	<i>oriV</i> _{pMB1} , <i>araC</i> controlled GFP, <i>cat</i> (Cm ^R)	BBa_I746908 (IGEM)
RAD module G8-C1	<i>oriV</i> _{ColE1} , <i>bxb1</i> integrase and excisionase, <i>araC</i> , <i>cat</i> (Cm ^R)	Addgene #38208
TP114	Incl2 conjugative plasmid	DSM-4246 (DSMZ)
TP114::Kill1	TP114:: <i>tetB</i> -Kill1 after <i>FRT</i> driven deletion of <i>tetB</i> and <i>oriV</i> _{pSC101ts}	This study
TP114::Kill3	TP114:: <i>tetB</i> -Kill3 after <i>FRT</i> driven deletion of <i>tetB</i> and <i>oriV</i> _{pSC101ts}	This study
TP114:: <i>tetB</i>	TP114Δ <i>aph-III</i> :: <i>FRT</i> - <i>tetB</i> - <i>attP</i> _{bxb1}	This study
TP114:: <i>tetB</i> -Kill1	TP114:: <i>tetB</i> with inserted Killing module 1	This study
TP114:: <i>tetB</i> -Kill3	TP114:: <i>tetB</i> with inserted Killing module 3	This study

Table S3.2: Oligonucleotides used in this study

Purpose	Name	Sequence ^{a, b}	Template	Purpose
pBxB1	oBxB1-F	ATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATT <u>CACGAGGCAGAATTCAGAT</u>	RAD Module G8-C1	Amplify <i>bxb1</i> integrase
	oBxB1-R	ATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAG <u>CGCCCTGCAGGAAATAATAA</u>		
	oBxB2-F	GATTATTAATCCGGCTTTTTTATTATTTCTGCAGGGCGCTCAAAGGCGGTAATACGGTT	pSB1A3	Amplify <i>bla-oriV_{pMB1}</i>
	oBxB2-R	AAGCTAAGGATTTTTTTATCTGAAATTCTGCCTCGTGAATGGTTTCTTAGACGTCAGGT		
pKill3	oKIL1-F	AACCACCGCGGTCTCAGTGGTGTACGGTACAAACCCCGAC <u>CGACAGTAAGACGGGTAAGC</u>	pGRG36	Amplify <i>oriV_{pSC101ts}</i>
	oKIL1-R	TATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTC <u>GGCTGAAAGCGCTATTTCTT</u>	pKN02 (unpublished)	Amplify gRNA 1
	oKIL2-F	AAGTTCCTATTCTCTAGAAAAGTATAGGAACTTC <u>GTGCACACTGATTAAGCATTGGTAACAGG</u>		
	oKIL2-R	TTTCGGGACATTCAGGAGATTTTCGCCGGACGTACGCATT <u>TCAGCACACTGAGACTTGT</u>	pKN02 (unpublished)	Amplify gRNA 2
	oKIL3-F	AATGCGTACGTCCGGCGAAAATCTCCTGAATGTCCCGAAAGGGCAGAAAGATGAATGACT		
	oKIL3-R	CTAAACAGGGATTGGCTGAGACGAAAGAATCTATTATACAGAAAAATTTCTGAAAGC	pKN02 (unpublished)	Amplify gRNA 2
	oKIL4-F	TTCTGTATAATAGATTCTTTTCGTCTCAGCCAATCCCTGTTTAGAGCTAGAAATAGCAAG		
	oKIL4-R	AACCGAGTGACCAAGAGAGGATGAAGCATTGCGGAGTAGTTCAGCACACTGAGACTTGT	pKN02 (unpublished)	Amplify gRNA 3
	oKIL5-F	CTACTCGGCAAATGCTTCATCCTCTCTTGGTCACTCGGTTGGGCAGAAAGATGAATGACTGTC		
	oKIL5-R	CTAAACTATTGGCCACGTTTAAATCAGAATCTATTATACAGAAAAATTTCTGAAAGC	pKN02 (unpublished)	Amplify gRNA 3
	oKIL6-F	TTCTGTATAATAGATTCTGATTTAAACGTGGCCAATAGTTTTAGAGCTAGAAATAGCAAG		
	oKIL6-R	GTCGGAAAAGTGGCCATCATTTGACGAACTACAG <u>CCCGGGTTCAGCACACTGAGACTTGT</u>	pKD4	Amplify <i>aph-III</i>
	oKIL7-F	<u>CCCGGGCTGTAGTTCGTCAAATGATGGCCACTTTTCCGACGGTGCTGACCCCGGATGAAT</u>		
	oKIL7-R	ATCATTCTATAGTATTAAGTATTGTTTTATGGCTGATAAAAGCGCTTTTGAAGCTGGGGT	pCas9	Amplify <i>cas9</i>
	oKIL8-F	GTTCTTCGCCACCCCAGCTTCAAAGCGCTTTTATCAGCCATAAAACAATACTTAATAC		
	oKIL8-R	CACCACTGAGACCGCGGTGGTTGACCAGACAAACCACGACTCAGTCACCTCCTAGCTGAC		
pKill1	oKIL9-F	AAGTTCCTATTCTCTAGAAAAGTATAGGAACTTC <u>GTGCACACTGATTAAGCATTGGTAACAGG</u>	pKN02 (unpublished)	Amplify gRNA 1
	oKIL9-R	GTCGGAAAAGTGGCCATCATTTGACGAACTACAG <u>CCCGGGTTCAGCACACTGAGACTTGT</u>		

Table S3.2: (continued)

pREC1	oREC1-F	<u>GACGACGGCGGTCTCCGTCGTGAGGATCATCCGGCGGAGGATATTCATATGGACCATGG</u>	pKD4	<i>Amplify oriV_{R6K}</i>
	oREC1-R	TATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCTTTTGC GGCCGCAAGATCCG		
	oREC2-F	TATCCGAAGTTCCTATTCTCTAGAAAAGTATAGGAACTTCATAGCCAGTTTTGGGGAAAG	pOSIP-TT	<i>Amplify tetB</i>
	oREC2-R	TCCTGACGACGGAGACCGCCGTCGTCGACAAGCCGGCCGACTAGGTCGACGCTTGGATT		
TP114:: <i>tetB</i>	oTPT1-F	<u>ACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGCGGATCTTGCGGCCGCAAAA</u>	pREC1	<i>Amplify FRT-tetB-attB_{bx1}</i>
	oTPT1-R	GCCGATGCGCCAACCGCATTCAATAAGACTAACTACCATGGTCCATATGAATATCCTCC		
	oTPT2-F	<u>GAGCCATATTCAACGGGAAACGTC</u>	TP114:: <i>tetB</i>	Screen <i>tetB</i> insertion
	oTPT2-R	<u>CACATCGGTGAAAGCTATGCC</u>		
Screen TP114::Kill	oTPK1-F	<u>GCTCGCTTGGACTCCTGTTG</u>	pREC1	Screen :: <i>tet</i> in TP114
	oTPK1-R	<u>CGTTGGCAAGACTGGCATGAT</u>		
	oTPK2-F	<u>TTGAAGGGTAGTCCAGAAGATAACG</u>	TP114::Kill3 or 1	Screen ::Kill3/Kill1 TP114
	oTPK2-R	<u>GGTAAATGGCACTACAGGCGC</u>		
	oTPK3	<u>CCTGTTACCAATGCTTAATCAGTGAC</u>	TP114::Kill3 or 1	Screen <i>FRT</i> recombination
	oTPK4-F	<u>GTGCACTGATTAAGCATTGGTAACAGG</u>	TP114::Kill3 or 1	Screen gRNA number
	oTPK4-R	<u>GGTCAGCACCGTCGGAAAAG</u>		
Plasmid copy number	oCAT-F	CCACGACGATTTCCGGCAG	pT <i>cat</i>	qPCR pT copy number
	oCAT-R	CTCACCCAGGGATTGGCT		
	oRPO-F	<u>GAATTCGATCCGAAGGACAA</u>	<i>EcN rpoB</i>	qPCR control
	oRPO-R	<u>AAAAGATGCGGTTTCACCAC</u>		

a. Underlined sequence are priming sites.

b. Bold sequence are restriction sites.

CHAPITRE 4

HIGH-EFFICIENCY CRISPR-CAS9 DELIVERY BY ENGINEERED BACTERIA FOR PRECISE ELIMINATION OF ANTIBIOTIC RESISTANT *ENTEROBACTERICEAE*.

4.1 Présentation de l'article et contributions.

La résistance aux antibiotiques est un problème pressant qui menace de transformer des infections bénignes en maladies mortelles (Antimicrobial resistance: global report on surveillance 2014. Available online at <https://www.who.int/drugresistance/documents/surveillancereport/en/>, 2014). Le microbiote entérique est un important réservoir de gènes de résistance aux antibiotiques et un endroit propice à leur mobilité (Schaik, 2015). Les *Enterobacteriaceae*, une famille de bactéries résidentes du microbiote entérique et parfois pathogène accumule les résistances aux antibiotiques à un rythme alarmant (Yamamoto & Pop-vicas, 2014). Une grande proportion des infections causées par les entérobactéries peuvent être attribuées à *E. coli* (Poolman & Wacker, 2016; Rojas-Lopez *et al*, 2018). Un outil permettant l'élimination spécifique des *E. coli* pathogènes et/ou résistants aux antibiotiques constitue une approche thérapeutique intéressante.

Dans ce chapitre, nous avons appliqué les connaissances et outils développés lors des deux chapitres précédents afin de développer la technologie du Conjugative Probiotic (COP). Cette technologie utilise un probiotique dérivé de EcN mobilisant, à l'aide de TP114, un module CRISPR pour éliminer des *E. coli* résistants aux antibiotiques. Dans un premier temps, la spécificité et l'efficacité du système sont évaluées dans un modèle de souris bicolonisé avec deux *E. coli*, dont seulement un est ciblé. Le traitement avec une seule dose du système COP a été en mesure d'éliminer 98 % des bactéries ciblées sans affecter le profil de colonisation de la souche non ciblée. Le système de mobilisation basé sur TP114 fut ensuite amélioré par des approches d'évolution accélérée en laboratoire. La souche COP résultante de cette évolution fut

ensuite mise au défi dans un modèle de souris simulant un résistome abondant. Les souris étaient colonisées cette fois-ci avec 1000 fois plus de cellules cibles résistantes aux antibiotiques que de non-cible. Suite au traitement avec le COP évolué, 99,98 % des bactéries cibles ont pu être éliminées. De plus, la population de bactéries non ciblée a su combler la niche écologique, permettant l'augmentation de leur abondance par plusieurs ordres de magnitudes. Somme toute, cette preuve de concept a su démontrer le potentiel de la technologie COP comme outil de modification du microbiome et comme antibiotique hautement spécifique.

Pour cet article, j'ai moi-même effectué la plupart des expériences et analyses, ainsi que fourni un effort considérable d'écriture. Patricia Roy m'a assisté pour les expériences d'évolution en laboratoire. Nancy Allard a contribué aux expériences de test d'efficacité du premier système COP. Frédéric Grenier a effectué les manipulations bio-informatiques et m'a assisté dans l'analyse des données de ribotypage 16S et de mutagenèse. Vincent Burrus a aidé à l'écriture du manuscrit et a fourni des idées pour la conception des expériences. Sébastien Rodrigue a aidé à la conception des expériences, dirigé les efforts de recherche et a contribué à l'écriture du manuscrit.

Référence bibliographique : Kevin Neil, Patricia Roy, Nancy Allard, Frédéric Grenier, Vincent Burrus, and Sébastien Rodrigue. (2020) High-efficiency CRISPR-Cas9 Delivery by Engineered Bacteria for Precise Elimination of Antibiotic Resistant *Enterobacteriaceae*. *Nature Medicine*. *Under consideration*.

4.2 Page titre

High-efficiency CRISPR-Cas9 Delivery by Engineered Bacteria for Precise Elimination of Antibiotic Resistant *Enterobacteriaceae*

Authors: Kevin Neil¹, Nancy Allard^{1§}, Patricia Roy^{1§}, Frédéric Grenier¹, Vincent Burrus¹ and Sébastien Rodrigue^{1*}.

Affiliations:

¹ Département de biologie, Université de Sherbrooke, Sherbrooke, Québec, J1K 2R1, Canada.

*Correspondence to: Sebastien.Rodrigue@USherbrooke.ca.

§These authors contributed equally to the presented work.

4.3 Abstract:

Bacterial resistance to antibiotics threatens our ability to treat infectious diseases, spurring a growing interest in alternative antimicrobial technologies. One attractive strategy is to use bacterial conjugation to deliver CRISPR-*cas9* systems programmed to precisely target and eliminate antibiotic resistant bacteria. However, the success of this approach heavily depends on high *in situ* DNA transfer rates. Here, we repurposed conjugative plasmid TP114 as a high efficiency delivery vehicle for CRISPR-*cas9* in the gut microbiota. A single dose of this system was able to eliminate 98% of targeted antibiotic resistant *Escherichia coli* in the mouse digestive tract while leaving an otherwise isogenic population and the rest of the microbiota essentially untouched. Through accelerated laboratory evolution, we next generated an improved system that killed >99.9% of the targeted antibiotic resistant population *in situ*. Taken together, these results demonstrate the potential of CRISPR-Cas9 delivery by engineered bacteria as a promising strategy to fight antibiotic resistance.

4.4 Main Text:

Antibiotic resistant infections rates are rapidly increasing worldwide, and threaten to become the second most important cause of mortality by 2050¹. *Enterobacteriaceae* are natural residents of the microbiota particularly infamous for their rapid accumulation of antibiotic resistance genes^{2,3}. This group comprises pathogens such as *Escherichia coli*, *Klebsiella pneumoniae*, *Shigella* sp., and *Salmonella* sp. that can cause deadly infections when not adequately treated⁴. Among these pathogens, *E. coli* is responsible for various diseases in farm animals and accounts for almost half of all human antibiotic resistant infections attributed to enterobacteria⁵.

The use of CRISPR-Cas9 as an antimicrobial represents a promising strategy to precisely eliminate targeted bacterial populations. CRISPR-Cas9 is a protein-RNA complex that can be programmed to cleave specific DNA sequences found only in target bacteria, which compromises the integrity of their genomes and leads to death⁶. The key step to enable this type

of antimicrobial application is to deliver the CRISPR-*cas9* system to virtually all bacteria of the targeted population. Engineered bacteriophages have been proposed as delivery vehicles for CRISPR-*cas9* systems, and proof-of-concept demonstrations have been performed in bacterial cultures as well as in mouse models⁷⁻⁹. While bacteriophages offer interesting advantages such as high infectivity, they often display narrow host ranges limited to specific strains⁸. Furthermore, phage receptors on the cell surface can quickly mutate¹⁰ and environmental conditions (low pH, gastric fluid, proteases, etc.) encountered by viral particles can dramatically limit their activity in the gut¹¹. In contrast, a probiotic strain could prove more resilient in the intestinal microbiota and leverage bacterial conjugation to deliver CRISPR-*cas9* to a broader range of target cells. Through bacterial conjugation, a probiotic bacterium can establish close contacts and transfer DNA to neighbouring microbes using a type IV secretion system (T4SS)¹². Until recently, conjugative transfer rates have generally been considered to be limiting for this type of application⁷. However, our group has recently identified plasmid TP114 as a highly proficient DNA transfer machinery in the gut microbiota¹³. We thus sought to develop a genetically engineered conjugative probiotic (COP) strain that repurposes TP114 for the mobilization of CRISPR-*cas9* genes (Figure 4.1a). We evaluate the feasibility of this approach in a proof-of-concept demonstration tackling antibiotic resistant *E. coli* in the gastro-intestinal (GI) tract.

Several bacterial strains would constitute interesting bacterial chassis for the development of the COP system. Because of a long history of clinical use¹⁴ and the availability of efficient genetic manipulation tools, we selected *Escherichia coli* Nissle 1917 (EcN) as our initial testbed. Three derivatives of EcN with different antibiotic resistance profiles were used (KN01, KN02 and KN03) to facilitate donor and recipient strain discrimination in conjugation assays (Extended Data Table 4.1). As previously reported, high in situ transfer rates were observed after the successive introduction in mice of recipient strain EcN KN02 followed by the donor strain EcN KN01 + TP114, and by probing conjugation events in feces¹³ (Figure 4.1b). The CRISPR-*cas9* system was next armed with a guide RNA (gRNA) targeting the chloramphenicol acetyl-transferase (*cat*) gene inserted in the genome of EcN KN02. The resulting CRISPR-*cas9*

system formed a killing module (Kill) that was integrated in TP114 by Double Recombinase Operated Insertion of DNA (DROID)¹⁵ to generate TP114::Kill. The activity of TP114::Kill against EcN KN02 was tested in vitro on agar plates, using EcN KN03 as a control non-target strain (Extended Data Figure 4.1). While transfer frequencies of TP114::Kill in EcN KN03 were similar to the TP114 wild-type (WT) plasmid, the rate of EcN KN02 TP114::Kill transconjugant formation was >10,000-fold lower, suggesting that the CRISPR-Cas9 system eliminated EcN KN02 containing the targeted *cat* gene upon DNA transfer.

The activity of the COP strain EcN KN01 bearing TP114::Kill was next assessed in the mouse gut microbiota. Mice first received a 1:1 mixture of EcN KN02 (target) and KN03 (non-target) by gavage to test both the efficiency and specificity of killing. Twelve hours later, a second gavage was performed using Phosphate Buffered Saline (PBS), EcN KN01 + TP114, or the COP strain (Figures 4.1c, 4.1d, and 4.1e). The amount of EcN KN02 and KN03 was then quantified in feces over a period of 4 days without any additional administration of the treatments (Figures 4.1f, 4.1g, and 4.1h). No significant difference was observed in the ratio of the EcN recipient strains KN02 and KN03 for the PBS and TP114 controls (Figures 4.1i and 4.1j). However, the number of EcN KN02 target bacteria was decreased by almost two orders of magnitude in the presence of the COP (median=98.6% decrease 36 hours after the administration of the COP) (Figure 4.1k). Similar results were obtained when the COP strain was introduced as a prophylactic treatment 12 hours prior to the gavage of the target and non-target strains, leading to a specific ~30-fold depletion of the targeted KN02 strain (median=96.4% decrease 36 hours after the administration of the COP) (Extended Data Figure 4.2).

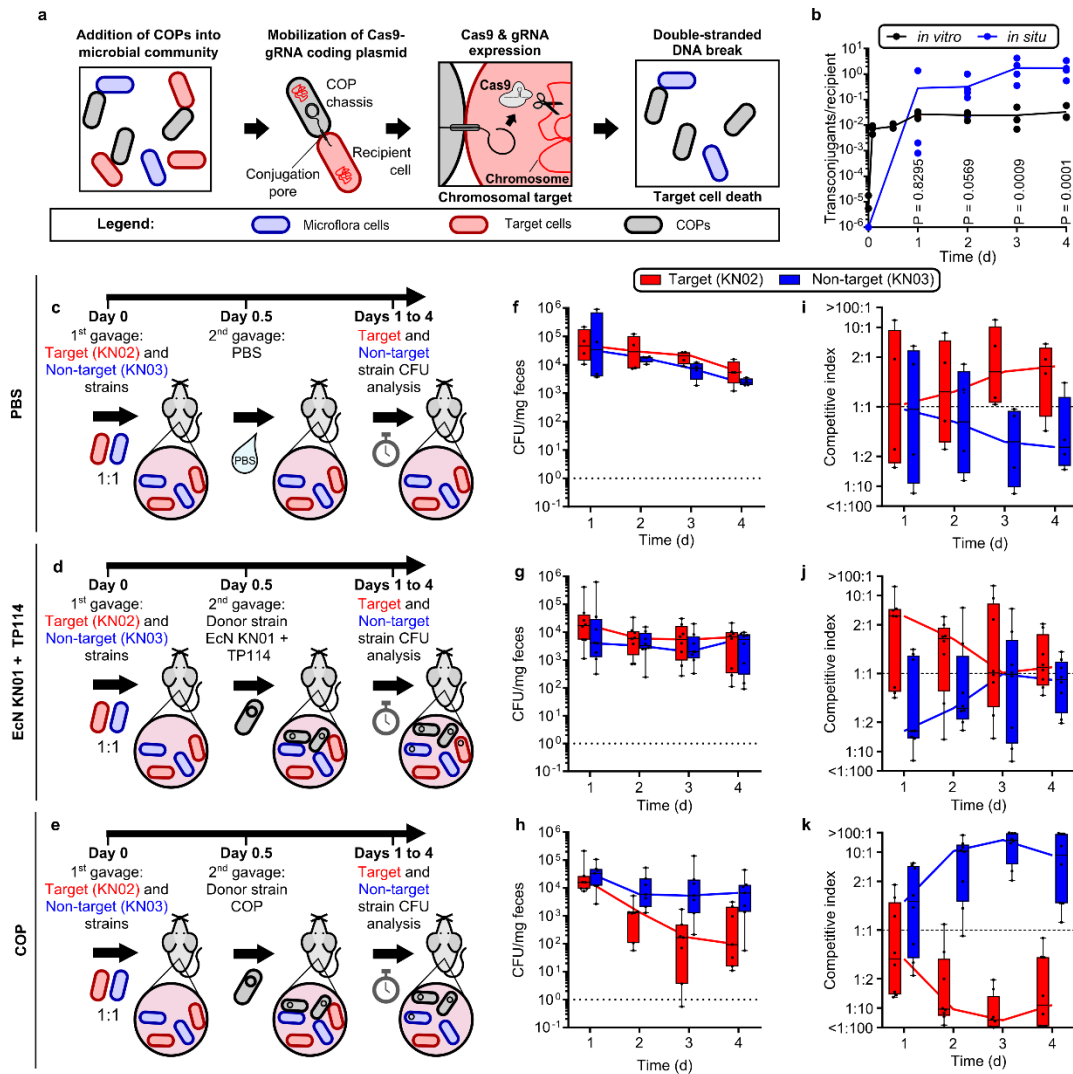


Figure 4.1: Conjugative probiotics (COPs) can transfer CRISPR-*cas9* at high-efficiency to specifically eliminate a target strain from the gut microbiota. Mode of action of the COP system (a). Transfer rates of TP114 on agar (n=3), and in the murine gut (*in situ*) (n=5) (b). Overview of the mouse model experiments. Mice were administered a 1:1 mixture of target and non-target strains and subsequently treated of PBS (c), EcN + KN01 + TP114 (d) or the COP (e). Absolute CFU abundance of the target and non-target strain in mice feces is shown after treatment with PBS (f), EcN + KN01 + TP114 (g) or the COP (h). Competitive index of the target and non-target strains were calculated from the CFU ratio between both strains for the PBS (i), EcN + KN01 + TP114 (j) or the COP (k) treated groups. For absolute quantifications,

the dotted lines indicate the highest detection limit observed (panels f, g and h). For competitive indexes, the dotted line indicates the expected competitive index at equilibrium (panels i, j and k). The box and whiskers graphs show the median with a line while boxes indicate the 10-90 percentiles, and the whiskers extends from the minimum to the maximum values. Six mice were used for each experimental groups (N=6).

To evaluate the broader impact of the COP strategy on the microbiota, 16S ribotyping was performed using feces from the previous experiments. Shannon alpha diversity was highly similar between treatments (average Krustal-Wallis p-value=0.8130), indicating that the COP had no significant effect on the microbiota diversity of each sample (Extended Data Figure 4.3a). Moreover, weighted Unifrac beta diversity for both the COP or the EcN KN01 bearing TP114 control strain were not significantly different, suggesting that the treatment with the COP did not induce major shifts in microbiota composition as compared to the control groups (Extended Data Figures 4.3b-h). The overall composition of the 10 bacterial orders detected in at least one sample from each treatment also remained similar between the controls and the COP (Extended Data Figure 4.3i). Taken together, these results show that a single dose of the COP can efficiently decrease the abundance of a target strain *in situ*, leaving the rest of the microbiota essentially untouched.

We next reasoned that Accelerated Laboratory Evolution (ALE) could be applied to the generation of more efficient variants of TP114. Plasmid MP6 was used to introduce mutations in TP114 without any strong bias¹⁶. MP6-dependent mutagenesis was performed in the diaminopimelic acid (DAP) auxotroph EcN KN01 Δ dapA donor harboring MP6 and TP114 to generate a library of donor strains containing randomly mutated copies of TP114. The donor library was then used in conjugation experiments either on agar or in broth, hence transferring the mutated copies of TP114 to EcN KN03 and leading to the loss of the EcN KN01 Δ dapA strain in absence of DAP supplementation. Transconjugants obtained in EcN KN03 were then

re-introduced into EcN KN01 Δ dapA + MP6 for 5 consecutive rounds of mutagenesis (Figure 4.2a). The transfer rates of the mutant libraries were measured at each round (Figure 4.2b) and improved by three orders of magnitude to reach nearly 100% at the end of the experiment. Mutations acquired by TP114 during ALE were identified for bulk mutant libraries as well as 6-7 isolated clones evolved for transfer on agar or in broth (eA or eB, respectively) per round of mutagenesis (Extended Data Figure 4.4, Extended Data 4.1). The exact eA and eB strains were identified with three numbers respectively indicating the mutagenesis round, the replicate, and the clone (Extended Data Figure 4.5 and 4.6). While ALE induced a variety of mutations, a gradual enrichment for mutations in the intergenic region between TP114-084 and TP114-085 was observed for eA- and eB-TP114 libraries, and was highly correlated with increased transfer rates (p-value=1.47E-42) (Extended Data Figure 4.7). eB-TP114 mutants were additionally enriched with mutations in *yaeC* (TP114-70) a predicted transcription regulator encoded by TP114. Conversely, eA-TP114 mutants displayed increased frequencies of deleterious mutations in the *pil* genes, which are required for conjugation in broth and *in situ*^{13,17}.

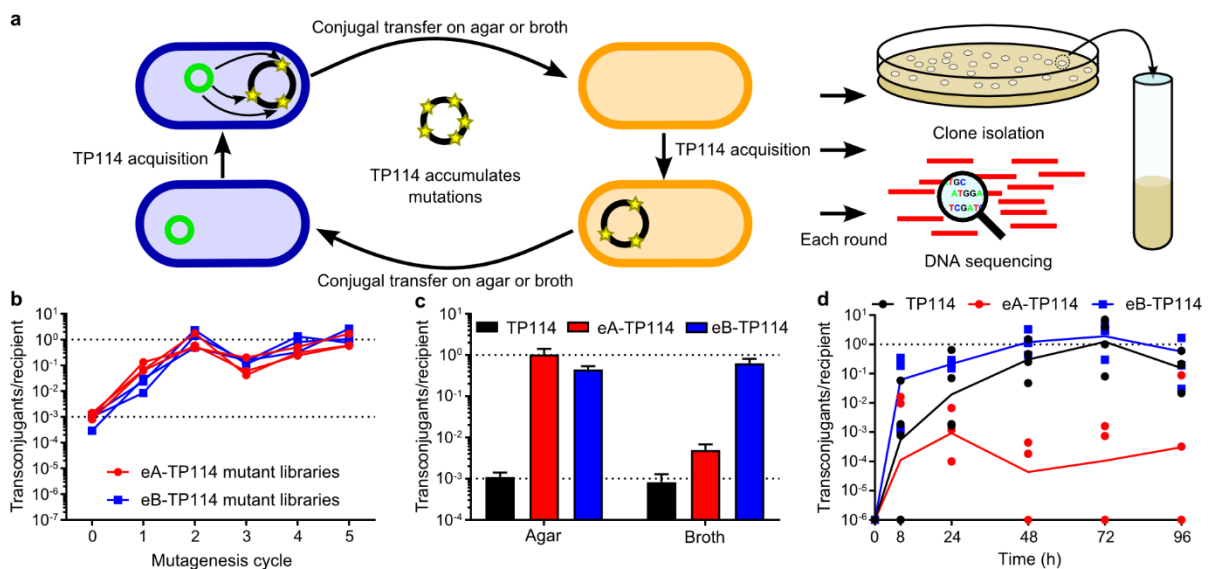


Figure 4.2: Accelerated Laboratory Evolution can boost TP114 transfer rates. Schematic representation of the experiment (a). TP114 mutant libraries transfer rates after 2 hours at 37°C as measured after each round of random mutagenesis on agar (eA) or in broth (eB) (b).

Comparison of conjugation efficiencies of final TP114 mutants obtained after 5 mutagenesis selection rounds on agar (eA-TP114) and in broth (eB-TP114). Histograms are representative of the average of the data and error bars represent standard deviation (c). Comparison of TP114, eA-TP114 and eB-TP114 transfer rates *in situ* quantified in feces (d). Four mice were used for each experimental groups (N=4).

Two representative clones, eA-TP114 (eA517) and eB-TP114 (eB527), were next selected to systematically quantify transfer frequencies on agar, in broth, and in the mouse intestinal tract. Both eA- and eB-TP114 transferred at >400-fold higher rates on agar relative to wild-type TP114. On the other hand, eA-TP114 showed only modest increase in transfer rates in broth (~6-fold increase) while eB-TP114 improved by almost three orders of magnitude (~768-fold) compared to TP114. Conjugative transfer in the mouse gut microbiota was next assessed by administering the recipient EcN KN02 strain by gavage, followed 2 hours later by the EcN KN01 donor strain carrying TP114, eA-TP114 or eB-TP114. Not surprisingly, eA-TP114 showed relatively low conjugation rates due to the presence of a mutation in the T4P, confirming its involvement in mating pair stabilization *in situ*¹³. In contrast, the eB-TP114 clone reached higher transfer efficiencies, with faster kinetics than TP114 (Figure 4.2d). For instance, an average of ~17% of the target cells had received eB-TP114 only 8 hours after gavage of the donor strain while this took between 24 and 48 hours for TP114.

Based on the higher transfer efficiency of eB-TP114 *in situ*, we sought to exploit this variant as an improved version of the conjugative machinery for the COP system. The CRISPR-*cas9* killing module was introduced in eB-TP114, producing the eB-COP strain. To challenge the eB-COP system, target (KN02) and non-target (KN03) EcN bacteria were orally administered to mice in a 1,000:1 (target:non-target) ratio. Twelve hours later, the EcN KN01 carrying eB-TP114 control strain (devoid of the CRISPR-*cas9* system) or the eB-COP strain was introduced by gavage (Figures 4.3a and 4.3b). The colonization level of the target (KN02) and non-target (KN03) strains was next followed in feces 24, 48, 72 and 96 hours after their introduction in mice (Figures 4.3c and 4.3d). While the population of EcN KN02 and KN03 remained stable in

the EcN KN01 eB-TP114 treated groups (Figures 4.3ce), a complete shift in bacterial populations was induced in eB-COP treated mice (Figures 4.3df). As little as 12 hours after the introduction of the donor strain, eB-COP was able to reduce the population of EcN KN02 relative to KN03 by ~95% compared to the control. At day 2, the antibiotic resistant strain KN02 no longer accounted for the majority of the population in eB-COP treated mice, with an average elimination of 99.5%. Four days after the single administration of eB-COP, the treatment continues to decrease the abundance of the KN02 population achieving a decrease of >4,000-fold of the mean CFU relative to control, and a clearance of >99.9% of the targeted bacteria. By eliminating the resistant *E. coli* from the gut, the eB-TP114::Kill system completely shifted the composition of the *E. coli* strain populations from a largely antibiotic resistant state to an almost totally sensitive one. The remaining <0.1% of target bacteria survived but did not acquire TP114 based on their antibiotic resistance profile, raising the possibility that they could have been recalcitrant to conjugative transfer. To investigate this possibility, we performed an *in vitro* experiment by mixing the eB-COP with EcN KN03 harboring a GFP reporter plasmid containing the targeted *cat* gene. After 24 hours, $\geq 98\%$ of EcN KN03 had lost the targeted plasmid but three GFP positive clones were recovered and re-exposed to the eB-COP (Extended Data Figure 4.8). A similar level of plasmid elimination was observed during this second treatment, suggesting that spared EcN KN03 do not possess inheritable resistance to bacterial conjugation and were likely not reached by the eB-COP strain. Residual target bacteria may therefore become actionable with further improvements to the system or to the eB-COP administration regimen.

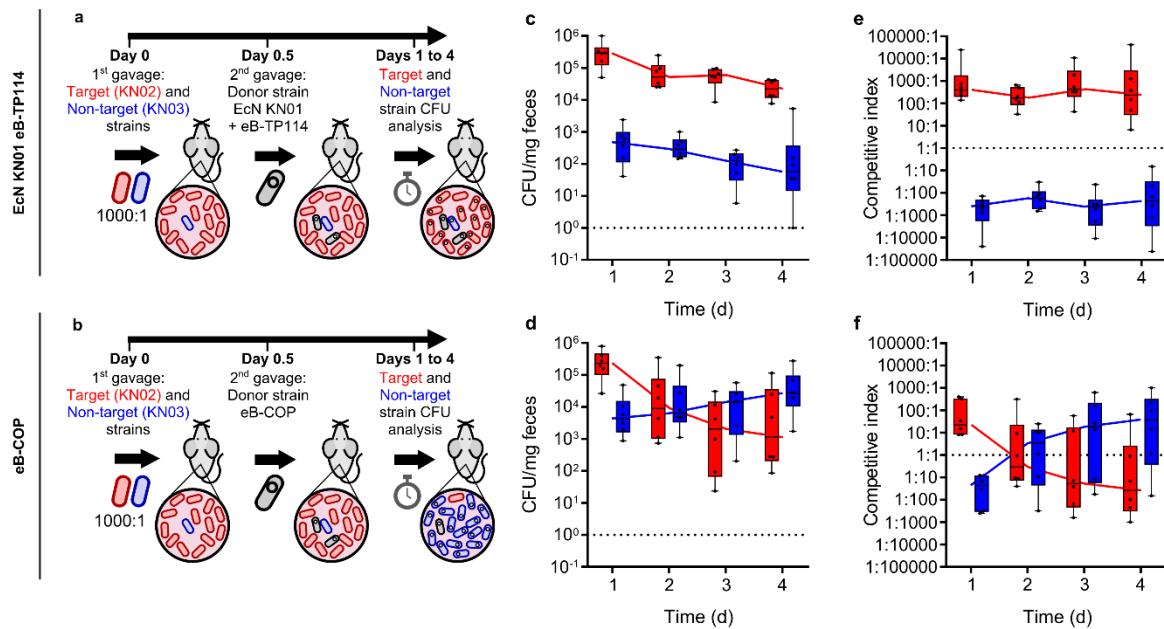


Figure 4.3: The eB-COP system can reshape bacterial populations in the gut as followed in feces. Schematic representation of the experiment, mice were first administered with a 1000:1 mix of target/non-target strains and were treated 12 hours later with either EcN KN01 + eB-TP114 (a), or the eB-COP (b). Absolute CFU abundance of the target and non-target strain was monitored in CFU of mice treated with EcN KN01 + eB-TP114 (c), or the eB-COP (d). Competitive indexes of the target and non-target strains are followed daily with EcN KN01 + TP114 devoid of a CRISPR-Cas9 system (e) and in mice treated with the eB-COP system (f). For absolute quantifications, the dotted lines indicate the highest detection limit observed (panels c and d). For competitive indices, the dotted line indicates the expected competitive index at equilibrium (panels e and f). The box and whiskers graphs show the median with a line while boxes indicate the 10-90 percentiles, and the whiskers extends from the minimum to the maximum values. Six mice were used for each experimental groups (N=6).

Our work provides a clear proof-of-concept demonstration that bacterial conjugation can be leveraged as an efficient alternative to conventional antibiotics, or as a potential strategy to revitalize conventional antibiotics by tackling the gut resistome¹⁸. The use of CRISPR-Cas9 as a killing agent provides a unique surgical precision to the treatment, ensuring that only the

targeted bacteria are eliminated. This strategy also offers the potential to perform microbiome editing to knock-down or eliminate specific microbial species. Additional work will be needed to determine or engineer the host range of mobile genetic elements that could be used to address different bacterial species with the COP approach. Future studies will also have to implement strict biocontainment measures, optimize administration regimens, and study the COP treatment dynamics for this technology to mature towards potential applications in animal or human health. At term, the COP could offer a new precision medicine strategy to prevent or treat a variety of microbiome related diseases and antibiotic resistant infections.

4.5 Methods:

4.5.1. Strains, plasmids and growth conditions.

Strains and plasmids are detailed in Extended Data Table 4.1 and are available upon request. Cells were routinely grown at 37°C for up to 18 hours in Luria broth Miller (LB) or on LB agar medium supplemented, with antibiotics when needed. Antibiotics were used at the following working concentrations: ampicillin 100 µg/mL, chloramphenicol 34 µg/mL, kanamycin 50 µg/mL, nalidixic acid 4 µg/mL, spectinomycin 100 µg/mL, streptomycin (Sm) 50 µg/mL, tetracycline 15 µg/mL. Diaminopimelic acid (DAP) auxotrophy was complemented by adding DAP at a final concentration of 57 µg/mL in the medium.

4.5.2. DNA manipulations.

The EZ10-Spin Column Plasmid Miniprep kit (Bio Basic) was used to extract small vector DNA whereas conjugative plasmids and genomic DNA (gDNA) were extracted using the Quick-gDNA miniprep (Zymo Research) according to the manufacturer's instructions. PCR amplifications were performed using Veraseq DNA polymerase (Enzymatics). Preparation of DNaseq libraries was performed using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina. Purification of DNA was performed between each step of DNaseq and plasmid

assembly by Solid Phase Reversible Immobilization (SPRI) using Agencourt Ampure DNA XP DNA binding beads (Beckman Coulter) following the manufacturer's guidelines. After purification, DNA concentration and purity were routinely assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific).

4.5.3. DNA transformation into *E. coli* by electroporation.

Routine plasmid transformations were performed by electroporation. Briefly, to prepare electrocompetent *E. coli* strains, an overnight culture of the desired host strain was sub-cultured 1/50 in 20 mL of LB broth and allowed to grow until an optical density of 0.6 at 600 nanometers (OD_{600nm}) was reached. The cells were then washed three times in sterile 10% glycerol solution and resuspended in 200 μ L of 10% glycerol. The DNA was then added to 40 μ L of electrocompetent cells, which was then transferred in a 1 mm electroporation cuvette. A pulse of 1.8 kV, 25 μ F, and 200 Ω for 50 ms was applied to the cells for electroporation. Cells were then resuspended in 1 mL of non-selective LB medium and incubated 1 hour at 37°C or 30°C (for thermosensitive plasmids) for recovery before plating on selective media.

4.5.4. *In vitro* conjugation assays.

In vitro conjugation assays routinely used EcN KN01 Δ *dapA* as the donor strain and EcN KN02 as the recipient strain. Bacterial culture were grown for no more than 18 hours prior to conjugation experiments. For each of the donor and recipient strain cultures, an adjusted volume of 500 μ L at a OD_{600nm} of 1.0 was transferred to separate sterile 1.5 mL microtube. Cells were centrifuged at 20,000 x g for one minute and washed once in 200 μ L sterile LB. The cells were then either resuspended in 2.5 μ L (for conjugation on agar) or 500 μ L (for conjugation in broth). Donor and recipient strains were next mixed at a 1:1 ratio. For conjugation on agar, the mixture was then deposited on an LB agar plate with DAP when appropriate. For conjugation in broth, the mixture was supplemented with DAP when appropriate and the microtube was put on a tube tumbler rotating mixer (VWR CAT# SBS550-2) at 20 rpm. Conjugation mixtures were

incubated at 37°C for 2 hours, unless specified otherwise, before being resuspended in sterile PBS and diluted 1/10 serially. A volume of 5 µL of each dilution was then spotted in duplicates on LB selective plates to select donors, recipients, and transconjugants. All conjugation frequencies were reported as a factor of the recipient CFUs. All conjugation experiments were repeated in at least three independent biological replicates.

4.5.5. Mouse Model.

All mice-related protocols were strictly evaluated to avoid animal suffering by the *Université de Sherbrooke Animal Care Committee*. Water and standard chow (Charles River) were provided *ad libitum* to the animals during the experiments. No more than 5 animals were housed in the same ventilated cage. All experiments used C57 BL/6 females of 16-20 g (Charles River), which were given at least a 3-days rest upon arrival. Daily evaluation of mice health and weight revealed no significant health or weight loss for all mice throughout the experiments.

Experiments involving *in situ* conjugation were based on a previously published model¹³ with minor modifications. Briefly, two days prior to gavage (Day -2), the strain(s) used for colonization were streaked from frozen stocks onto MacConkey selective plates and incubated at 37°C overnight. The same day, 1 g/L of Sm was added to the mice drinking water and changed every 3 days to maintain Sm efficacy. Adding Sm to the drinking water helped stabilize the colonization levels of EcN in mice as all introduced strains were resistant to Sm. The next day (Day -1), several colonies were inoculated in selective LB broth 5 mL pre-cultures and incubated at 37°C. Three to four hours prior to mice gavage, 200-500 µL of each strain pre-culture was transferred in 20 mL selective LB broth and allowed to grow at 37°C until an OD_{600nm} of 0.6 ± 0.1. The cells were then washed once in PBS and concentrated in a volume equivalent to 6.0 OD_{600nm} according to initial OD in LB. Each mouse then received 100 µL of the cell suspension by gavage (corresponding to ~1x10⁸ CFU). Part of the inoculum was plated and analyzed for CFU counts, which ranged between 10⁸ and 10⁹ CFU/mL for all experiments. Routinely, the recipient strain was fed to mice 12 hours before the donor strain unless specified otherwise.

Conjugation was then monitored in feces at specified time points. For each sampling, collection tubes were prepared in advance by aseptically adding a single 0.2 cm glass bead and 500 μ L of PBS to a sterile 1.5 mL tube. To normalize CFU by sample weight, tubes were weighted before and after sampling. Samples were next homogenized using a FastPrep-24 (MP) instrument for 1 minute at maximum speed. The samples were then serially diluted 1/10 in sterile PBS from 10^0 to 10^{-7} of the initial concentration, and 2.5 μ L of each dilution was spotted in technical duplicates on selective MacConkey plates. As a quality control for Sm treatment, total *Enterobacteriaceae* clearance was also followed on MacConkey plates without antibiotics for all experiments.

4.5.6. Accelerated Laboratory Evolution.

Plasmid MP6 was transformed into EcN KN01 Δ *dapA* and EcN KN01 Δ *dapA* + TP114 (Extended Data Table 4.1) by electroporation. Resulting transformants were plated on selective LB supplemented with 1% glucose to repress MP6. To generate mutant libraries, EcN KN01 Δ *dapA* + MP6 + TP114 was grown in LB containing 200 mM arabinose overnight at 37°C. The mutant libraries were then used as donors in a conjugation experiment for 2 hours at 37°C either on agar or in broth using EcN KN03 as the recipient strain. Transconjugants were spread on at least three selective LB agar plates to keep a good diversity of mutants throughout the tests. After an overnight incubation, transconjugant cell lawns were resuspended in 9 mL LB broth using a cell scraper and diluted 1/50 in LB containing adequate antibiotics before incubation at 37°C for 1 hour. They were next used as donors to transfer TP114 mutants back to KN01 Δ *dapA* + MP6 either on agar or in broth. This process was repeated for 5 successive rounds of mutagenesis. For each cycle, mutant libraries in EcN KN03 were cryopreserved for sequencing analyses. In addition, six isolated clones per round were kept for sequencing and further testing. For sequencing, isolated clones were inoculated in 200 μ L of LB selective medium and allowed to grow overnight while 200 μ L aliquots of mutant libraries were simply thawed on ice. DNA was extracted using Quick DNA Magbead Plus Kit (Zymo Research) and DNAseq libraries were prepared using the NEBNext Ultra II FS DNA Library Prep Kit. Illumina

sequencing of the mutagenesis experiment was performed at the Plateforme RNomics (<https://rnomics.med.usherbrooke.ca/>) of the Université de Sherbrooke on a NextSeq 500 instrument using single-end reads of 75 bp. Processing of the reads was inspired by the GenPipes¹⁹. Reads were first trimmed based on their quality using Trimmomatic v0.32²⁰ with the parameters TRAILING:30 and MINLEN:30. Reads quality before and after trimming was measured with FastQC v0.11.5²¹. Reads were then aligned to reference TP114 sequence from Genbank (MF521836.1) using BWA-MEM v0.7.10 with the default parameters²². The aligned reads were sorted with Picard v1.123²³ and realigned with the GATK v3.7 RealignerTargetCreator and IndelRealignerSingle²⁴. The coordinates of the realigned read mates were fixed with BVATools v1.4²⁵. Duplicates were marked with Picard v1.123²³. Haplotypes were called using GATK v3.7 Haplotype Caller and their effects were predicted using SnpEff v4.1²⁶. The intergenic region between *pilV* and *rci* was removed from analysis as it contains a shufflon, for which the DNA sequence is constantly re-arranged in the TP114 population. TP114 sequence in the Genbank record missed an A at position 52,704, which was added in the reference sequence for SNP calling analysis. Also, TP114 gene *yadA*, *yadB* and TP114-054 all showed high homology to *E. coli* Nissle1917 genome, provoking false positive SNP callings. These three genes were removed from heatmaps shown in the Extended Data Figures 4.4-4.7, but did not contained any enriched mutation profiles.

4.5.7. 16S ribotyping.

Mice feces from various treatments (PBS, EcN KN01 + TP114, COP) were sampled daily, and the DNA was purified using the ZymoBIOMICS DNA Miniprep Kit (ZymoResearch) as recommended by the manufacturer. DNA was further purified by SPRI and amplified with primers binding on the V3-V4 region of the 16S rDNA²⁷. The PCR products were purified by SPRI and barcoded for Illumina TruSeq sequencing. The resulting data were analyzed using the Qiime2 suite v2018.8²⁸ and visualized with Qiime2 view. Read counts for all detected bacterial orders in all conditions were transformed into read proportion and shown as a heatmap of its

logarithmic value. Alpha diversity (Shannon) and beta diversity (weighted Unifrac) were analyzed by built-in Qiime2 plugins.

4.5.8. Statistics and reproducibility.

Statistical significance was performed on the logarithmic value of the data using One-way ANOVA unless specified otherwise. P-values are directly indicated on the graphs were relevant and represent statistical significance of the difference between the two data groups. Differences in the data were considered significant when the P-value was below 0.05. Box and whiskers graphs represent the center line as the median, box encompasses the 25-75 percentiles, and whiskers are the minimal and maximal values for each data groups. All *in vitro* experiments were performed in biological triplicates using three independently grown cultures. For experiments involving mice, a minimum of 4 C57 BL/6 female mice were used for each sampling.

4.5.9. Data and materials availability.

Reads for 16S ribotyping (Figure 4.1) and mutagenesis experiments (Figure 4.2) are available at NCBI under BioProject: PRJNA510842 and PRJNA674373 respectively. Raw CFU counts and conjugation frequencies evaluation are available upon request.

4.6 Author contributions:

K.N. and S.R. designed the experiments; K.N. performed the experiments with assistance from P.R. and N.A.; K.N. analyzed the data; F.G. performed bioinformatic analyses for the 16S ribotyping and mutagenesis experiments; S.R. and V.B. supervised the project. K.N. and S.R. wrote the manuscript; P.R., N.A. and V.B. revised the manuscript.

4.7 Acknowledgments:

The Plateforme RNomics of Université de Sherbrooke for assistance with Illumina sequencing, as well as Compute Canada and Compute Quebec for access to bioinformatics resources and support. We also acknowledge Alfredo Menendez, Alain Lavigueur, Jean-François Millau and all lab members for thoughtful discussions. This work was supported by the Canadian Institutes of Health Research (CIHR #159817). S.R. holds a Chercheur boursier junior 2 fellowship from the *Fonds de recherche du Québec – Santé* (FRQS), and K.N. is the recipient of a graduate research scholarship from the *Fond de Recherche du Québec – Nature et Technologies* (FRQNT) and from the Natural Science and Engineering Research Council of Canada (NSERC). N.A. is supported by a doctoral scholarship from the *Université de Sherbrooke*. P.R. is supported by a scholarship from the Natural Science and Engineering Research Council of Canada (NSERC).

4.8 Competing interests:

The work presented in the manuscript is part of patent application WO2020010452A1. All of the authors of the present manuscript except for P.R. and F.G. are also co-authors of this provisional patent application. K.N. and S.R. have a financial interest in TATUM bioscience.

4.9 References:

1. Antimicrobial resistance: global report on surveillance 2014. *World Health Organ.* Available at <http://www.who.int/drugresistance/doc> (2014).
2. Nordmann, P., Naas, T. & Poirel, L. Global spread of carbapenemase producing *Enterobacteriaceae*. *Emerg. Infect. Dis.* **17**, 1791–1798 (2011).
3. Yamamoto, M. & Pop-vicas, A. E. Treatment for infections with carbapenem-resistant *Enterobacteriaceae* : what options do we still have ? *Crit. Care* **18**, 229–237 (2014).

4. Matthew E. Falagas, Giannoula S. Tansarli, Drosos E. Karageorgopoulos & Konstantinos Z. Vardakas. Deaths Attributable to Carbapenem-Resistant *Enterobacteriaceae* Infections. *Emerg. Infect. Dis.* **20**, 1170–1175 (2014).
5. Cassini, A. *et al.* Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *Lancet Infect. Dis.* **19**, 56–66 (2019).
6. Cui, L. & Bikard, D. Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res.* **44**, 4243–4251 (2016).
7. Citorik, R. J., Mimee, M. & Lu, T. K. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat. Biotechnol.* **32**, 1141–1145 (2014).
8. Yosef, I., Manor, M., Kiro, R. & Qimron, U. Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. *Proc. Natl. Acad. Sci.* **112**, 7267–7272 (2015).
9. Bikard, D. *et al.* Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat. Biotechnol.* **32**, 1146–1150 (2014).
10. Shabbir, M. A. B. *et al.* Bacteria vs. Bacteriophages: Parallel Evolution of Immune Arsenal. *Front. Microbiol.* **7**, 1292 (2016).
11. Nobrega, F. L. *et al.* Genetically manipulated phages with improved pH resistance for oral administration in veterinary medicine. *Sci. Rep.* **6**, 1–12 (2016).
12. Arutyunov, D. & Frost, L. S. F conjugation: Back to the beginning. *Plasmid* **70**, 18–32 (2013).
13. Neil, K., Allard, N., Grenier, F., Burrus, V. & Rodrigue, S. Highly efficient gene transfer in the mouse gut microbiota is enabled by the IncI2 conjugative plasmid TP114. *Commun. Biol.* **3**, 1–9 (2020).
14. Jacobi, C. A. & Malfertheiner, P. *Escherichia coli* Nissle 1917 (Mutaflor): new insights into an old probiotic bacterium. *Dig. Dis.* **29**, 600–7 (2011).
15. Neil, K., Allard, N., Jordan, D. & Rodrigue, S. Assembly of large mobilizable genetic cargo by double recombinase operated insertion of DNA (DROID). *Plasmid* **104**, 102419 (2019).
16. Badran, A. H. & Liu, D. R. Development of potent *in vivo* mutagenesis plasmids with broad mutational spectra. *Nat. Commun.* **6**, 1–10 (2015).

17. Ishiwa, A. & Komano, T. The lipopolysaccharide of recipient cells is a specific receptor for PilV proteins, selected by shufflon DNA rearrangement, in liquid matings with donors bearing the R64 plasmid. *Mol. Gen. Genet.* **263**, 159–164 (2000).
18. Schaik, W. Van. The human gut resistome. *Phil. Trans. R. Soc. B* **370**, 1–9 (2015).
19. Bourgey, M. *et al.* GenPipes: An open-source framework for distributed and scalable genomic analyses. *Gigascience* **8**, 1–19 (2019).
20. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
21. Andrew, S. FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.bab> (2010).
22. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN] (2013).
23. Available online at <https://broadinstitute.github.io/picard/>.
24. Mc Kenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
25. Available online at <https://bitbucket.org/mugqic/bvatools>.
26. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. **6**, 80–92 (2012).
27. Thijs, S. *et al.* Comparative Evaluation of Four Bacteria-Specific Primer Pairs for 16S rRNA Gene Surveys. *Front. Microbiol.* **8**, 1–15 (2017).
28. Rideout, J. R. *et al.* QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ Prepr.* (2018).

4.10 Extended Data Document

High-efficiency CRISPR-Cas9 Delivery by Engineered Bacteria for Precise Elimination of Antibiotic Resistant *Enterobacteriaceae*

Authors: Kevin Neil¹, Nancy Allard^{1§}, Patricia Roy^{1§}, Frédéric Grenier¹, Vincent Burrus¹ and Sébastien Rodrigue^{1*}.

Affiliations:

¹ Département de biologie, Université de Sherbrooke, Sherbrooke, Québec, J1K 2R1, Canada.

*Correspondence to: Sebastien.Rodrigue@USherbrooke.ca

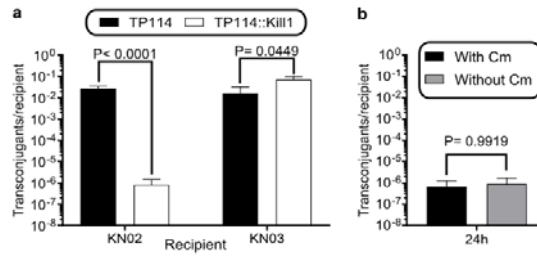
§These authors contributed equally to the presented work.

This file includes:

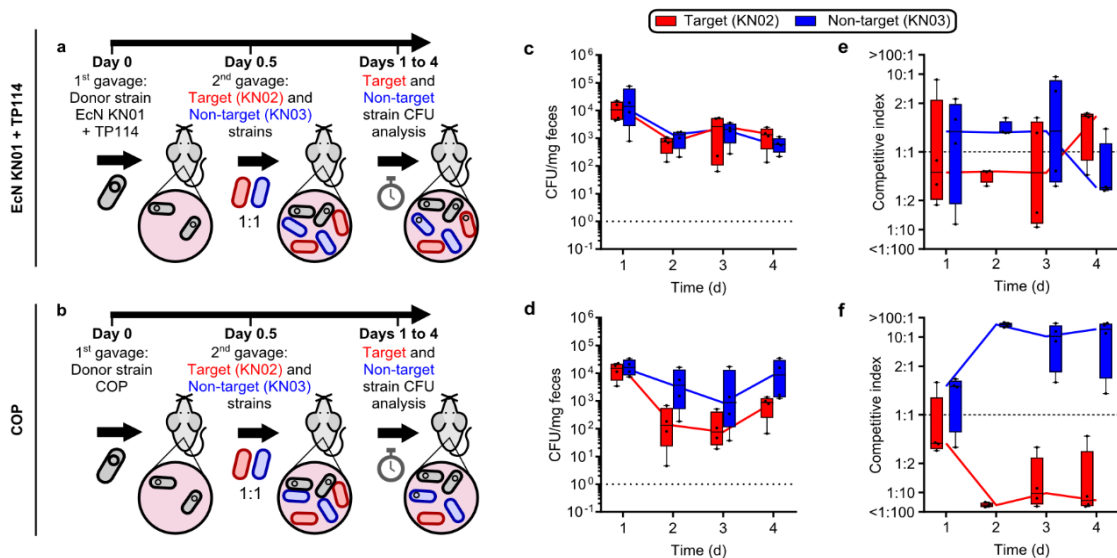
Supplementary Figures. 1 to 8

Supplementary Table 1

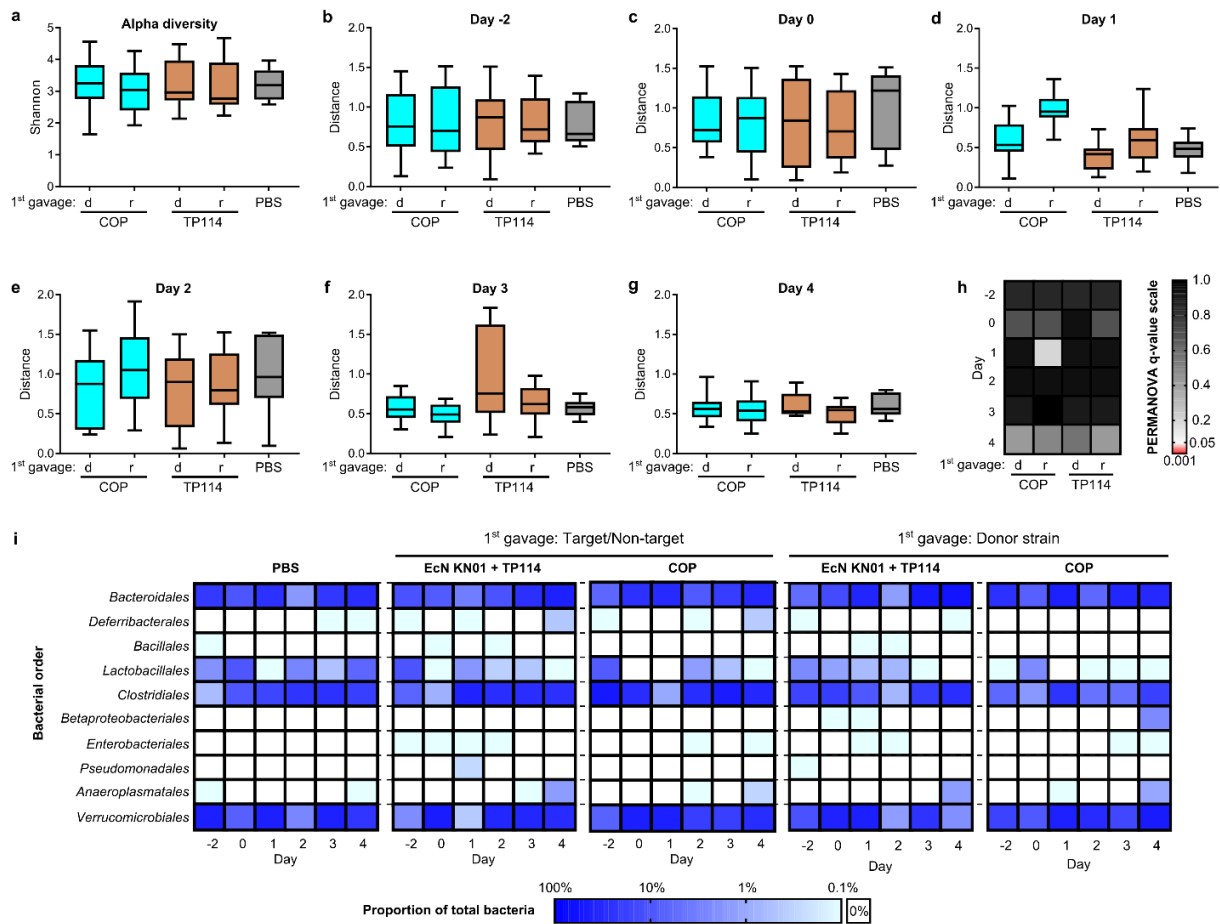
Supplementary References 1 to 5



Extended Data Figure 4.1. COPs can selectively eliminate a target strain from a mixture *in vitro*. **a**, Transfer frequency of TP114::Kill and TP114 towards a 1:1 mix of target (KN02) and non-target (KN03) recipient bacteria (n=3). **b**, Selection of TP114::Kill transconjugants with or without chloramphenicol (Cm) allows distinction between killing and Cm sensitization (n=3). **a-b**, Error bars show standard deviation of the mean of at least 3 biological replicates.

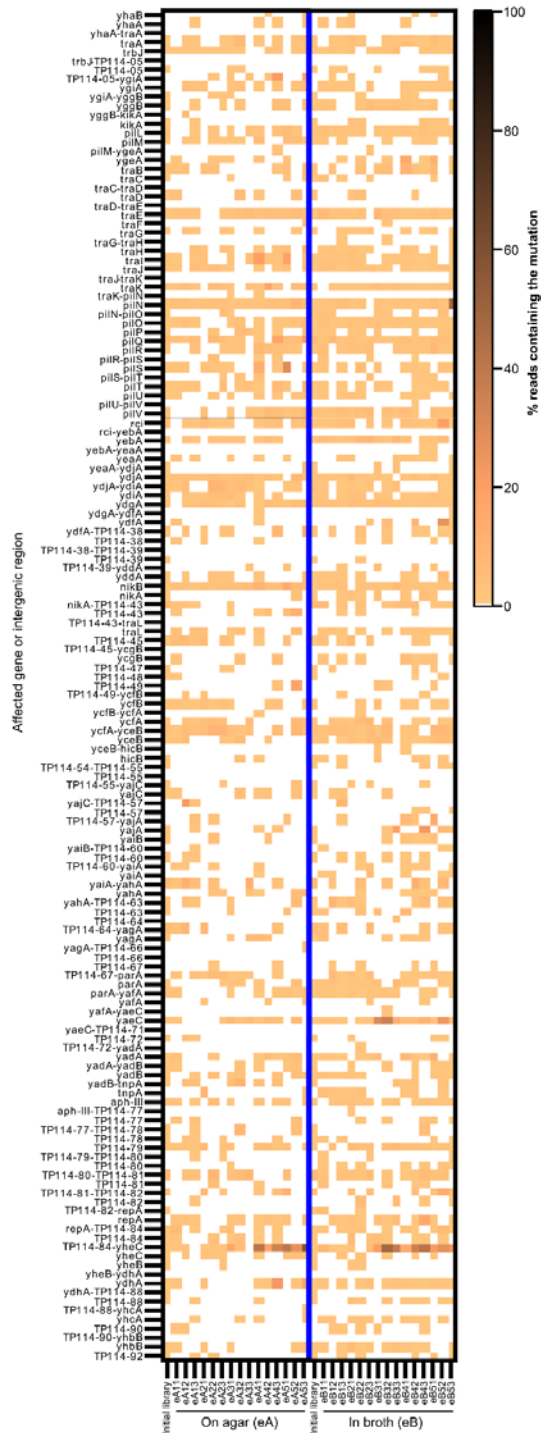


Extended Data Figure 4.2. COPs can protect the microbiota from colonization of a specific strain. Mouse model experiment, mice received a dose of an EcN KN01 + TP114 (a) or the COP (b) 12 hours prior to the introduction of the target/non-target 1:1 mix. CFU abundance of the target and non-target strains for the mice groups given EcN KN01 + TP114 (c) or the COP (d) was followed in feces for up to 4 days post gavage (n=4). Competitive indexes calculated from the CFU ratio between both strains in mice treated prophylactically with either EcN KN01 + TP114 (e) or the COP (f) (n=4). For panels b and e, the dotted line indicates the highest detection limit based on feces weight (10⁰). For panels c and f, the dotted line highlights the expected competitive index at equilibrium. For all box and whiskers plots, medians are shown as a bar, boxes represent the 10-90% percentile of the data and whiskers extend to the minimal and maximal values.



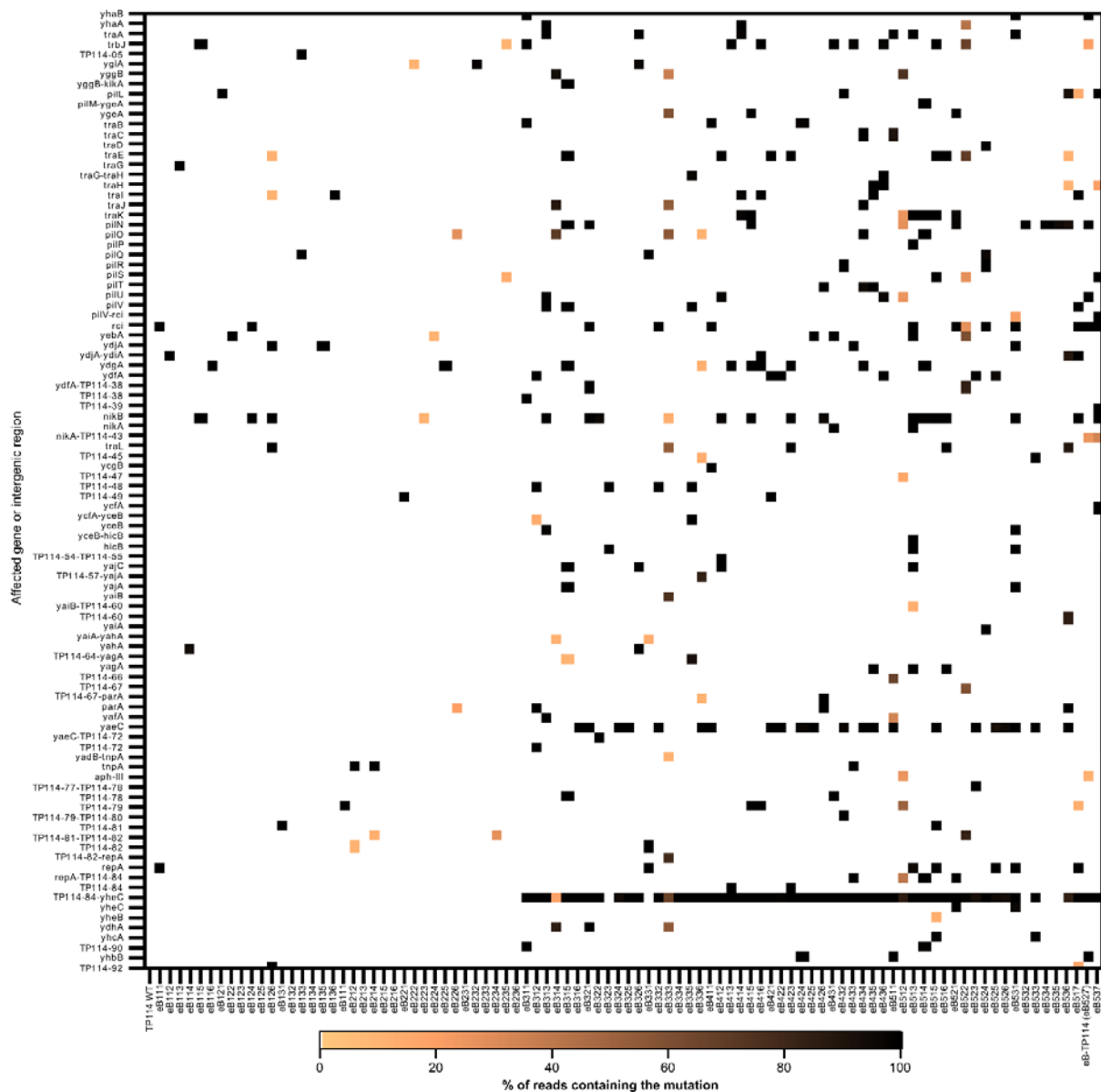
Extended Data Figure 4.3. COP treatment has no major effect on microbiota composition.

a, Microbiota diversity for each treatment groups (see Figure 1 and Extended Data Figure 2) from day 1 to 4 was evaluated using Shannon's index. Beta diversity was calculated by weighted unifrac comparing each group to the PBS control at day -2 (**b**), 0 (**c**), 1 (**d**), 2 (**e**), 3 (**f**), and 4 (**g**). **h**, The corrected significance of Beta diversity divergence to the PBS control is shown as a heatmap. **a-h**, days are given relative to the first gavage on day 0. Day -2 represents the beginning of the streptomycin treatment. Mice groups are divided by treatment types (COP, TP114 or PBS) and strains used for the first gavage, i.e. donor (d), recipient (r) or phosphate buffered saline (PBS). All groups contain at least 4 mice. **i**, Relative abundance of 10 bacterial orders in mice groups treated with PBS, EcN KN01 + TP114 or the COP treated mice groups by 16S ribotyping (n=4).

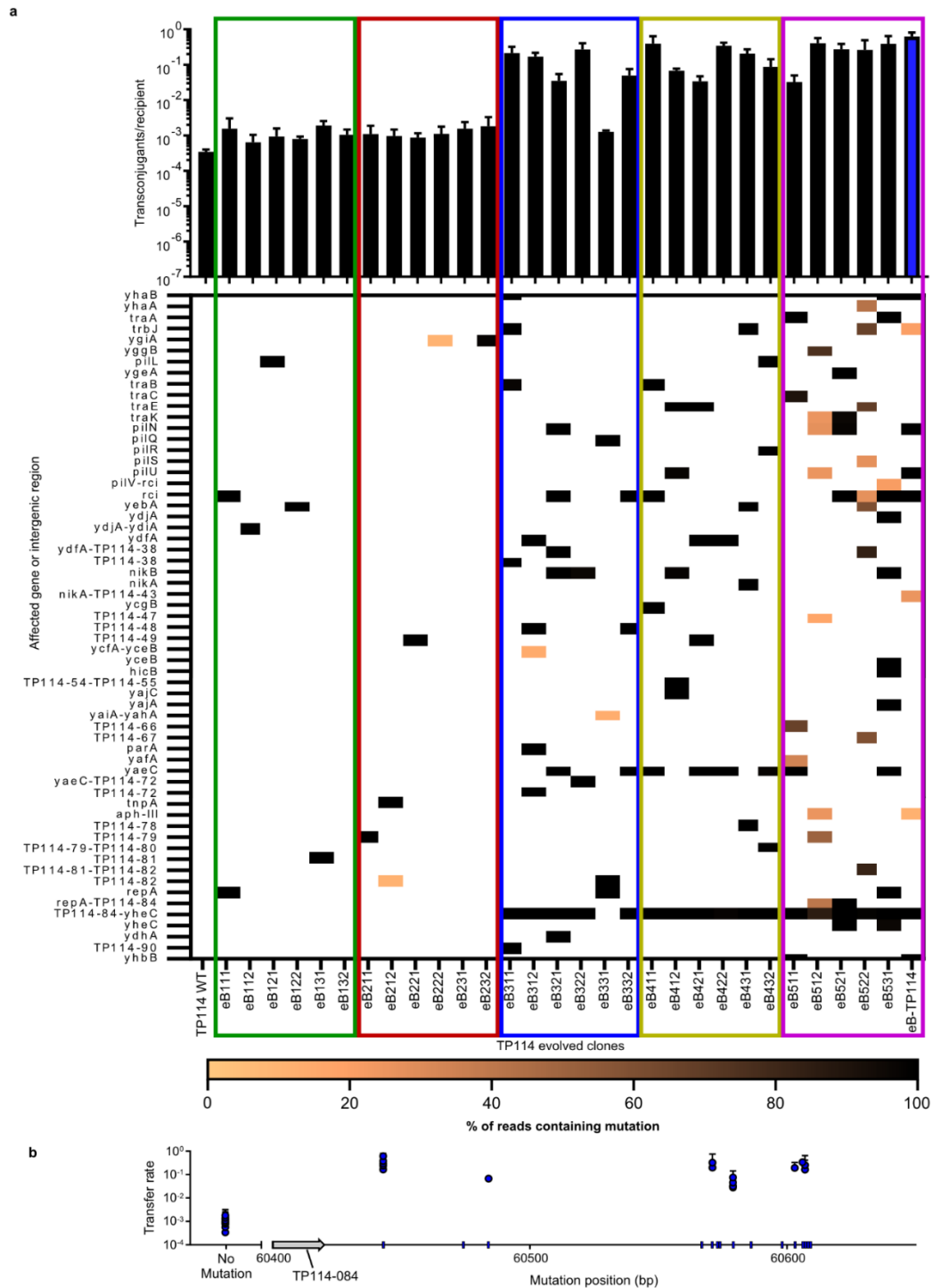


Extended Data Figure 4.4. Accelerated laboratory evolution selected for mutations in specific regions of TP114. Mutation heatmap showing the proportion of reads detecting a

mutation in a specific gene or intergenic region in mutant libraries selected for conjugation on agar and in broth. Mutant libraries are identified as either evolved on agar (eA) or in broth (eB), followed by a first digit which specify the number of mutagenesis rounds and a second digit specifying the library replicate in a eAXX, eBXX fashion.

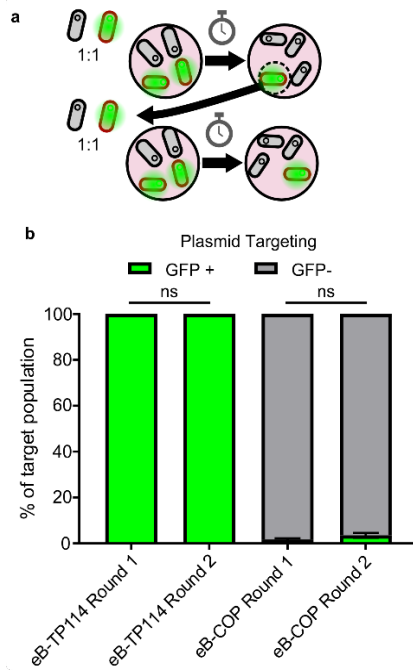


Extended Data Figure 4.6. Isolated clones for mutagenesis rounds in broth present enriched mutations in specific regions of TP114. Mutation heatmap showing the proportion of reads detecting a sequence variant in a specific gene or intergenic region in isolated clones. The exact clones obtained from evolution experiments selected in broth were identified with the letters eB, followed by three numbers respectively indicating the mutagenesis round, the replicate, and the clone (eBXXX).



Extended Data Figure 4.7. Correlation between mutation mapping and transfer rates of TP114 evolved clones. a, The transfer rates in broth of wildtype TP114 was compared to 30

clones of TP114 evolved for conjugation in broth (eB) and isolated at different mutagenesis cycles (n=3, biological triplicate). Presence of any mutation in TP114 genes is shown in the heatmap, with the color scale representing the proportion of reads confirming the presence of a variant. **b**, The transfer rate of each clone shown in panel a is shown relative to the position of the mutation in the intergenic region between TP114-084 and TP114-085 affecting the clone, if any. The exact clones obtained from evolution experiments selected in broth were identified with the letters eB, followed by three numbers respectively indicating the mutagenesis round, the replicate, and the clone (eBXXX).



Extended Data Figure 4.8. The remaining fraction of target bacteria after eB-COP do not display inheritable resistance to the system. Overview of the re-exposition experiment (a). A recipient strain carrying a GFP plasmid encoding the targeted *cat* gene was mixed with the eB-COP strain. Then three residual GFP positive colonies were picked and re-challenged with a second round of eB-COP. Plasmid elimination, as measured by counting the proportion of GFP positive and negative colonies, was evaluated for the two successive rounds of COP treated (b). As a control, the same strain was exposed to KN01 carrying eB-TP114. All experiments were performed in biological triplicates.

Extended Data Table 4.1: List of strains and plasmids used in this study.

Strain or plasmid	Relevant phenotype or genotype*	Source/Reference
<i>E. coli</i>		
COP	EcN KN01 + TP114::Kill	This study
eB-COP	EcN KN01 + eB-TP114::Kill	This study
EC100D <i>pir</i> ⁺	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) φ80 <i>dlacZ</i> ΔM15 Δ <i>lacX74 recA1 endA1 araD139</i> Δ(<i>ara, leu</i>)7697 <i>galU galK</i> λ- <i>rpsL nupG pir</i> ⁺ (DHFR)	#ECP09500 (Lucigen)
KN01	Sm ^r Sp ^r Nissle 1917	Neil <i>et al.</i> 2020 ¹
KN01Δ <i>dapA</i>	Δ <i>dapA</i> KN01	Neil <i>et al.</i> 2020 ¹
KN02	Sm ^r Cm ^r Nissle 1917	Neil <i>et al.</i> 2020 ¹
KN03	Sm ^r Tc ^r Nissle 1917	Neil <i>et al.</i> 2020 ¹
MG1655Nx ^R	K-12 F ⁻ λ- <i>ilvG</i> - <i>rfb-50 rph-1 Nx</i> ^r	Carraro <i>et al.</i> 2014 ²
Nissle 1917	Wildtype probiotic strain	DSM-6601 (DSMZ)
Conjugative plasmid		
TP114	Incl2, Km ^r	DSM-4246 (DSMZ)
eA-TP114	Evolved by accelerated laboratory evolution for 5 cycles on agar	This study
eB-TP114	Evolved by accelerated laboratory evolution for 5 cycles in broth	This study
Plasmid		
pBxB1	<i>oriV</i> _{pMB1} , <i>bx1</i> integrase, <i>bla</i> (Ap ^r)	Neil <i>et al.</i> 2019 ³
pE-FLP	<i>oriV</i> _{pSC101ts} , <i>flp</i> , Ap ^r	Addgene #45978
pKill	<i>oriV</i> _{pSC101ts} , <i>attP</i> _{bx1} , FRT, 1 gRNA vs <i>cat</i> , <i>aph-IIIa</i> (Km ^r), <i>cas9</i>	Neil <i>et al.</i> 2019 ³
MP6	Mutagenesis plasmid	Bradán <i>et al.</i> 2015 ⁴
pREC1	<i>oriV</i> _{REK} , <i>attB</i> _{bx1} , <i>tetB</i> (Tc ^r), FRT	Neil <i>et al.</i> 2019 ³
pSIM6	<i>oriV</i> _{pSC101ts} , Lambda Red recombinase, Ap ^r	Datsenko and Wanner 2000 ⁵
TP114::Kill	TP114:: <i>tetB</i> -Kill1 after FRT driven deletion of <i>tetB</i> and <i>oriV</i> _{pSC101ts}	Neil <i>et al.</i> 2019 ³
TP114:: <i>tetB</i>	TP114Δ <i>aph-III</i> :: <i>tetB</i>	Neil <i>et al.</i> 2019 ³
TP114:: <i>tetB</i> -Kill	TP114:: <i>tetB</i> with inserted Killing module 1	Neil <i>et al.</i> 2019 ³
eB-TP114::Kill	eB-TP114:: <i>tetB</i> -Kill1 after FRT driven deletion of <i>tetB</i> and <i>oriV</i> _{pSC101ts}	This study
eB-TP114:: <i>tetB</i>	eB-TP114Δ <i>aph-III</i> :: <i>tetB</i>	This study
eB-TP114:: <i>tetB</i> -Kill1	eB-TP114:: <i>tetB</i> with inserted Killing module 1	This study

*Abbreviations: Ap^r: Ampicillin resistance, Cm^r: Chloramphenicol resistance, Km^r: kanamycin resistance, Nx^r: Nalidixic acid resistance, Sp^r: Spectinomycin resistance, Sm^r: Streptomycin

resistance, Su^r: Sulfamethoxazole resistance, Tc^r: Tetracycline resistance and Tm^r: Trimethoprim resistance.

Extended Data 4.1: Position and putative impact of mutations found by sequencing accelerated evolution mutant libraries and clones.

Available online at: https://drive.google.com/drive/folders/1Z_HsbiMFznNgJiZIW1uK90Obc-Ihx9ka?usp=sharing

4.11 Supplementary references:

1. Neil, K., Allard, N., Grenier, F., Burrus, V. & Rodrigue, S. Highly efficient gene transfer in the mouse gut microbiota is enabled by the IncI2 conjugative plasmid TP114. *Commun. Biol.* **3**, 1–9 (2020).
2. Carraro, N., Matteau, D., Luo, P. & Burrus, V. The Master Activator of IncA / C Conjugative Plasmids Stimulates Genomic Islands and Multidrug Resistance Dissemination. *PLoS Genet.* **10**, e1004714 (2014).
3. Neil, K., Allard, N., Jordan, D. & Rodrigue, S. Assembly of large mobilizable genetic cargo by double recombinase operated insertion of DNA (DROID). *Plasmid* **104**, 102419 (2019).
4. Badran, A. H. & Liu, D. R. Development of potent *in vivo* mutagenesis plasmids with broad mutational spectra. *Nat. Commun.* **6**, 1–10 (2015).
5. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* **97**, 6640–6645 (2000).

CHAPITRE 5

DISCUSSION GÉNÉRALE ET CONCLUSION

5.1 Discussion et conclusion générale

Dans le cadre de mes études doctorales, j'ai cherché à démontrer qu'un outil basé sur le phénomène de conjugaison bactérienne pourrait permettre de contrôler les populations microbiennes du microbiote entérique. Le développement du système COP a contribué de plusieurs manières à faire avancer les connaissances fondamentales. Le criblage de plasmides conjugatifs *in situ* nous a permis d'évaluer la contribution de plusieurs familles de plasmides conjugatifs à la mobilité des gènes dans le microbiote entérique. Près de la moitié des plasmides n'ont pas été en mesure d'effectuer du transfert d'ADN dans le microbiote entérique. Ce criblage nous a permis d'identifier TP114 comme une machinerie de transfert très efficace dans le milieu intestinal. En étudiant le fonctionnement de TP114, nous avons également été en mesure d'identifier la stabilisation de la paire de conjugaison comme un mécanisme contribuant grandement à la conjugaison bactérienne dans le microbiote entérique. Ce mécanisme pourrait d'ailleurs avoir des implications dans la sélection de la bactérie réceptrice, déterminant ainsi le spectre d'hôte effectif du plasmide conjugatif (Ishiwa & Komano, 2000, 2003).

L'insertion de CRISPR-*cas9* dans TP114 nous a également permis de mettre sur pied une nouvelle méthode d'assemblage d'ADN. Cette méthode trouve des applications directes dans la fusion de modules thérapeutiques à des plasmides conjugatifs pour leur dissémination dans le microbiote entérique. Outre la livraison de modules CRISPR-*cas9*, la conjugaison bactérienne pourrait alors être utilisée comme thérapie génique pour le microbiote entérique, supplémentant ainsi les bactéries résidentes avec de nouvelles fonctions. Par exemple, des gènes de dégradation du lactose (lactases) pourraient être introduits dans les bactéries du microbiote entérique afin de diminuer l'intolérance au lactose (comme suggéré dans le brevet abandonné concernant l'utilisation de bactéries modifiées pour exprimer la lactase humaine : US20120321718A1). La

méthode DROID a également un potentiel encore inexploré pour l'assemblage d'ADN. En effet, l'intégrase Bxb1 utilisée dans la première étape du protocole recombine des sites *attB/attP*. Il fut démontré qu'en modifiant deux nucléotides centraux à ces sites, on pouvait obtenir plusieurs paires de sites *attP/attB* orthogonaux permettant la fusion simultanée de plusieurs fragments d'ADN via des combinaisons des variants des sites *attP/attB* correspondants (Ghosh *et al*, 2008). En découpant les constructions en parties d'ADN sélectionnable, il serait en théorie possible de fusionner plusieurs plasmides via ces sites orthogonaux soit simultanément ou par étape. L'utilisation de recombinases et d'intégrases permet également en théorie de manipuler de très grosses molécules d'ADN. La méthode DROID pourrait donc être facilement adaptée à l'intégration stable de gros fragments d'ADN dans le génome de *E. coli*. Cette application du DROID est particulièrement intéressante pour l'intégration de systèmes conjugatifs dans le chromosome de la souche donneuse dans une optique de bioconfinement et d'amélioration de la stabilité de la souche.

Au terme de mes études, j'ai été en mesure de développer l'outil COP jusqu'à une phase de preuve de concept, où j'ai démontré sa grande efficacité (>99 %) à éliminer des bactéries résistantes aux antibiotiques dans l'intestin d'un modèle de souris. En adaptant des méthodes d'évolution accélérée en laboratoire, nous avons pu augmenter considérablement la fréquence de transfert de TP114 tant *in vitro* que *in situ*. En utilisant cet outil amélioré, nous avons ensuite été en mesure de complètement renverser le profil de résistance d'une population bactérienne au sein du microbiote entérique. Bien que la technologie COP ait mûri jusqu'à l'étape de preuve de concept, plusieurs aspects de la technologie COP devront cependant être étudiés avant que cette technologie ne puisse être utilisée dans une application thérapeutique.

5.2 Comprendre TP114 pour mieux le contrôler

Le plasmide conjugatif TP114 est un outil très efficace pour transférer de l'ADN dans le microbiote entérique (Neil *et al*, 2020). Malheureusement, avant le début de mes travaux sur TP114, les plasmides de sa famille (IncI₂) étaient très peu étudiés (Bradley *et al*, 1980; Kim &

Komano, 1992). Bien que plusieurs études intéressantes ont été menées sur la famille IncI₁ et le complexe-I (l'ensemble des familles de plasmide encodant un pilus accessoire reconnu par le bactériophage I), très peu d'informations sont disponibles sur la famille IncI₂ (Praszkier & Pittard, 2005; Ishiwa & Komano, 2003; Kim & Komano, 1992; Bradley, 1984; Ishiwa & Komano, 2004; Brouwer *et al*, 2015). Malgré leurs noms similaires, les familles IncI₁ et IncI₂ partagent très peu de régions d'homologies et ont des archétypes de T4SS différents (Tableau 1.2, Supplementary Figure 2.6) (Neil *et al*, 2020). Il est donc difficile d'inférer au IncI₂ les découvertes faites chez les IncI₁. Les travaux réalisés au chapitre 2 dressent une base solide pour l'étude de TP114. L'expérience d'HDTM combinée aux prédictions bio-informatiques de la fonction des gènes et de leur conservation chez les IncI₂ nous ont permis d'avoir un aperçu global de l'importance de chaque gène à la conjugaison *in situ*. Ces connaissances pourront nous permettre de concevoir des versions réduites de TP114 possédant seulement les gènes nécessaires à son transfert afin de réduire la charge du plasmide pour sa cellule hôte. En effet, plusieurs gènes présents sur TP114 ne sont pas nécessaires à son transfert ou à son utilisation dans le système COP. Ces gènes peuvent encoder pour des fonctions tout de même importantes pour le plasmide, ou être simplement un vestige d'événements fortuits (e.g. insertion de transposon). Or, plusieurs fonctions importantes pour TP114 dans la nature ne sont pas nécessaires au système COP. Par exemple, TP114 encode pour un homologue de *finO*, connu dans la littérature comme une arme des plasmides IncI pour réprimer les plasmides du groupe IncF dans une situation de compétition (Getino & de la Cruz, 2018). Ceci n'est pas nécessaire dans notre système, mais est un avantage clair pour un plasmide conjugatif sauvage. Un premier pas vers la conception d'une version minimale de TP114 serait de créer une banque de mutant de délétion de tous les gènes de TP114 à la manière de la collection Keio (Baba *et al*, 2006). Cette banque de mutant permettrait de confirmer les données d'HDTM et de concevoir des délétions plus larges sur TP114 afin de réduire sa complexité.

Une meilleure compréhension de TP114 nous permettra également de modifier son efficacité à effectuer du transfert conjugatif. En effet, en identifiant les éléments de régulations clés de TP114, il devrait être possible de les modifier à notre avantage. En contrôlant l'expression des

régulateurs chez les plasmides conjugatifs, il est possible d'augmenter la fréquence de transfert de façon significative (Carraro *et al*, 2014). Les expériences de mutagenèse par transposon présentées au chapitre 2 ne nous ont pas permis d'identifier un gène qui, lorsqu'interrompu par un transposon, semble améliorer le taux de transfert de TP114. Ceci suggère que la régulation de TP114 n'est pas dépendante d'un seul gène agissant comme répresseur global du transfert. L'analyse bio-informatique de la fonction des gènes a identifié 5 gènes homologues à des facteurs de transcriptions chez TP114. Par contre, aucun de ces gènes n'était muté dans la version eB-TP114 évolué au chapitre 4. En effet, la dérégulation du transfert chez eB-TP114 semble être fortement liée à la présence de mutations dans la région intergénique entre le gène TP114-084 et TP114-085. Cette région intergénique pourrait moduler l'expression des gènes en aval TP114-085 à TP114-005, suggérant que les éléments de régulation les plus importants sur TP114 se trouveraient dans ce locus. Alternativement, il est possible que la régulation de TP114 soit exercée directement par des facteurs de transcription provenant de l'hôte bactérien. Comme démontré au chapitre 2, la régulation de TP114 semble être liée à la concentration d'oxygène ambiante. Les gènes responsables du transfert de TP114 sont donc réprimés en présence de fortes concentrations en oxygène, puis activés en absence d'oxygène. Il est donc possible que la régulation de TP114 passe par des régulateurs globaux du changement de métabolisme entre les conditions aérobiques et anaérobiques telles que Fnr, ArcA, NarL et NarP (Tolla & Savageau, 2010; Brunet *et al*, 2011). On trouve d'ailleurs un site de liaison Fnr dans la région intergénique entre le gène TP114-084 et TP114-085. Il se peut que les mutations présentes chez les clones suractivés empêchent la régulation du plasmide par l'hôte bactérien. Élucider le réseau de régulation de TP114 nous permettra peut-être d'augmenter davantage les performances de ce plasmide *in situ* et d'obtenir de meilleures efficacités d'élimination, comme démontré par les expériences d'évolutions dirigées.

Parallèlement, il sera rapidement important d'identifier les unités transcriptionnelles chez TP114. Bien que des outils performants existent pour identifier *in silico* des terminateurs (de Hoon *et al*, 2005), des expériences de 5'RACE seront nécessaires afin de prédire la localisation des promoteurs et de constituer les opérons (Rapid amplification of 5' complementary DNA

ends (5' RACE), 2005). L'identification des débuts de transcription nous permettra dans un premier temps de vérifier la liaison de différents gènes régulateurs candidats sur les régions régulatrices grâce à des techniques tel le ChIP (Carraro *et al*, 2014). De plus, les promoteurs présents sur TP114 pourraient être remplacés par des promoteurs plus forts ou modulés par d'autres signaux. Par exemple, il pourrait être avantageux de pouvoir éteindre l'expression des gènes de transfert durant les étapes de production pour éviter le fardeau métabolique et éviter les mutations délétères. Alternativement, en connaissant les unités de transcription sur TP114, il serait possible de réformer le plasmide et de le diviser en unités fonctionnelles. Par exemple, nous pourrions regrouper les gènes responsables de la synthèse du T4SS ensemble, ceux du T4P ensemble, etc. Ceci faciliterait ensuite la modulation de leur expression et pourrait contribuer au bioconfinement de la souche.

5.3 Bioconfinement de la souche COP

Bien que très peu discuté dans les chapitres précédents, un aspect important concernant l'utilisation de bactéries génétiquement modifiées comme médicament amène d'importants enjeux de bioconfinement de la souche utilisée (Charbonneau *et al*, 2020). En effet, ces organismes doivent être confinés pour éviter leur survie dans des contextes naturels et l'exposition accidentelle de patients sains à des traitements impliquant ces organismes modifiés. Généralement, les bactéries-médicament doivent être exemptes de matériel génétique mobilisable pour éviter la fuite de gènes synthétiques dans des bactéries naturelles (Charbonneau *et al*, 2020). À cet égard, le système COP requiert des mesures de bioconfinement particulières. On peut voir le bioconfinement du système COP sur deux niveaux : (1) la restriction de la survie de la bactérie à des conditions de laboratoire et (2) la restriction de la persistance de l'élément génétique transféré dans les bactéries sauvages. Pour le confinement de la souche OGM, l'auxotrophie est souvent utilisée et semble être jugée suffisante par la FDA (Food and Drug Administration) (Leventhal *et al*, 2020; Steidler *et al*, 2003). Les auxotrophies les plus populaires sont générés en retirant les gènes *dapA* et *thiA* seuls ou en combinaison (Leventhal *et al*, 2020). L'auxotrophie générée par la délétion du gène *dapA* est particulièrement

intéressante, puisqu'elle rend la souche dépendante pour sa survie d'un apport exogène en DAP, une molécule retrouvée en très faible quantité dans la nature (Richaud *et al*, 1986; Allard *et al*, 2015). Il a également été démontré que l'auxotrophie au DAP a peu d'influence sur l'efficacité de conjugaison de la souche donneuse, du moins en condition de laboratoire (Allard *et al*, 2015). Cette méthode promet de limiter la survie de la souche donneuse dans l'environnement, mais ne limite pas la persistance de l'élément génétique mobilisé dans le cas du système COP.

Le système COP actuel exploite TP114 pour mobiliser CRISPR-*cas9* en *cis*. Or, TP114 possède des systèmes de partitionnement et de stabilisation lui permettant de persister dans l'environnement (gènes identifiés *in silico* au chapitre 2). Une première étape vers le bioconfinement de la souche serait de limiter les gènes transférés au minimum. Pour ce faire, le module CRISPR doit être relocalisé vers un plasmide possédant l'*oriT* de TP114, permettant ainsi une mobilisation dite en *trans*. Il est bien connu que la relocalisation de l'*oriT* sur un plasmide tiers permet sa mobilisation par la machinerie de transfert du plasmide conjugatif (Ferrières *et al*, 2010). Une première étape de confinement peut donc être réalisée en retirant l'*oriT* de TP114 et en le relocalisant sur un plus petit plasmide qui serait mobilisé en *trans*. De plus, la réplication du plasmide mobilisable peut être restreinte à la souche donneuse. La méthode la plus simple implique l'utilisation de la machinerie de réplication dérivée du plasmide R6K (Azad *et al*, 1994). Dans un tel système, le gène *pir*, soit la protéine d'initiation de la réplication, est insérée dans le génome de la souche donneuse et contrôle la réplication du plasmide navette. Lorsque mobilisé, le plasmide navette est incapable de se répliquer et est rapidement perdu. Par contre, ses gènes peuvent être exprimés, ce qui permet au système CRISPR-Cas9 d'exercer son activité thérapeutique. Il est tout de même possible que la souche réceptrice contienne un plasmide de la même famille que R6K et puisse compléter la réplication du plasmide navette. Par contre, le plasmide sera tout de même éventuellement perdu selon les principes d'incompatibilité de réplication. Un bioconfinement accru du système de mobilisation peut également être atteint via la réforme du plasmide conjugatif en unité fonctionnelle minimale comme discuté plus haut. Ces unités fonctionnelles peuvent ensuite être

insérées à différents endroits dans le génome de la bactérie donneuse afin d'éviter la reconstitution du plasmide conjugatif par recombinaison.

Ces modifications permettent de diminuer les risques de persistance des gènes recombinants dans la nature, mais n'éliminent pas complètement tous les risques. À terme, le bioconfinement des organismes médicaments devrait être fait par recodage du génome entier (Mandell *et al*, 2015). Plusieurs stratégies ont déjà été proposées afin de bioconfiner les souches OGM par recodage du génome. Par exemple, l'élimination de tous les codons stop « TAG » du génome suivie de leur attribution à un acide aminé synthétique permet de limiter la persistance de la souche modifiée dans l'environnement (Mandell *et al*, 2015). Cette modification limite également sa capacité d'interprétation de l'ADN naturel qui pourrait entrer dans la cellule par transfert horizontal. Par contre, cette méthode ne permet pas de limiter l'expression des transgènes par les bactéries naturelles à moins d'introduire dans ces gènes les acides aminés non-naturels. Ces modifications pourraient par contre avoir des conséquences difficiles à anticiper sur l'efficacité des transgènes. Alternativement, il fut récemment proposé de recoder le génome d'*E. coli* afin de réattribuer les codons associés à un ARNt de la leucine vers l'arginine et vice versa (Fujino *et al*, 2020). Cette stratégie permet à la fois que l'ADN sauvage entrant et que l'ADN modifié sortant de la bactérie recodée ne soit pas compatible. En effet, avec un génome recodé inversant les codons arginine et leucine, l'ADN des bactéries sauvages entrant dans la bactérie confinée serait mal interprété et mènerait à des protéines où les acides aminés leucine et arginine seraient inversés et donc non-fonctionnels en grande majorité. De la même façon, si l'ADN de notre souche modifiée se retrouve transformé chez une bactérie sauvage, cette dernière ne sera pas en mesure de l'interpréter. Cependant, réaliser un tel exploit demeure techniquement très difficile et dispendieux.

5.4 Facteurs limitant l'efficacité d'élimination du système COP

Afin d'évaluer l'applicabilité de la technologie COP dans différentes pathologies, il faut d'abord établir les limites de ce système. Les expériences de preuve de concept ayant été faite en utilisant

E. coli Nissle1917 comme bactérie donneuse et réceptrice, peu de barrières au transfert conjugatif étaient présentes. Nous savons donc pour l'instant que le système COP peut être appliqué au traitement d'infections à *E. coli*. Par contre, pour plusieurs applications, tel le traitement du résistome entérique, l'élimination d'*E. coli* seule est insuffisante (Schaik, 2015). Il est donc souhaitable d'évaluer le spectre d'hôte de TP114 afin de délimiter le spectre d'action du système COP.

Le spectre d'hôte de TP114 est limité de trois façons : (1) par la capacité du T4P à se lier sur la réceptrice afin de stabiliser la paire de conjugaison (Neil *et al*, 2020), (2) par l'aptitude du T4SS à transférer de l'ADN à la cellule réceptrice et (3) par la capacité du plasmide TP114 à se répliquer chez la cellule réceptrice (Jain & Srivastava, 2013). Chacun de ces trois cas de figure peut être investigué et discriminé par transfert conjugatif sur agar et en bouillon :

(1) L'incapacité du T4P à reconnaître la cellule cible mènera à des taux de conjugaison beaucoup plus faibles en bouillon comparativement à un support solide sur agar. Contrairement au T4SS, les cibles moléculaires du T4P ainsi que la ou les protéines responsables de son attachement à la cellule cible sont relativement bien caractérisées chez le plasmide IncI₁ R64 (Ishiwa & Komano, 2000, 2003, 2004). Chez R64, comme chez la plupart des plasmides du complexe-I, l'extrémité du T4P responsable de l'adhésion est encodée par un gène hautement conservé appelé *pilV*. Le gène *pilV* de R64 est une adhésine capable de lier des combinaisons de sucres trouvés à la surface des LPS bactériens. Afin d'élargir leur spectre d'hôte, les plasmides du complexe-I possèdent un locus spécialisé appelé « shufflon » qui réorganise la partie C-terminale de *pilV* (Brouwer *et al*, 2015). Ces réarrangements permettent alors d'exprimer différents variants de *pilV* pouvant reconnaître différentes structures moléculaires (Ishiwa & Komano, 2003). Le shufflon peut être immobilisé en retirant la shufflase souvent retrouvée en aval du shufflon. Tout comme la plupart des plasmides du complexe I, le gène *pilV* de TP114 est réorganisé par un shufflon, produisant ainsi différents variants de l'adhésine. En caractérisant les structures reconnues par chaque variant de *pilV*, il serait possible de prédire le spectre d'hôte de TP114 à l'aide d'outils bio-informatiques. Le spectre d'hôtes de TP114 pourrait également

être modifié si nous disposons de suffisamment de connaissances sur la protéine *pilV* pour modifier sa spécificité. À terme, il pourrait être possible de contrôler le spectre d'hôtes de TP114 avec suffisamment de précision pour compléter ou remplacer l'utilisation de CRISPR-Cas9. Des mécanismes de reconnaissances et de stabilisation très performants pourraient même permettre l'utilisation d'une méthode d'élimination non spécifique des bactéries cibles plutôt que d'une approche dépendante des séquences d'ADN basée sur CRISPR-Cas9. Par contre, restreindre la reconnaissance de la souche ciblée de cette façon rendrait le système plus susceptible à des phénomènes de résistance au transfert, puisque l'unique cible de l'adhésine pourrait muter de la même manière qu'un récepteur de phage. Pour réduire la fréquence des phénomènes de résistance, il serait possible de prendre avantage du shufflon présent sur TP114 pour ré-arranger l'adhésine PilV de sorte de la faire reconnaître plusieurs structures présentes à la surface d'une même souche de bactérie.

(2) L'inaptitude du T4SS à transférer de l'ADN ne produira pas ou peu de transconjugants pour des transferts tant sur agar qu'en bouillon. En effet, une fois la bactérie réceptrice à proximité de la bactérie donneuse via la stabilisation de la paire de conjugaison soit par la machinerie du plasmide ou par un support solide, le T4SS doit être en mesure de traverser les membranes des cellules réceptrices afin d'y transférer de l'ADN. Le mécanisme de reconnaissance de la bactérie cible par le T4SS chez TP114 reste à investiguer. Cependant, des données préliminaires générées au laboratoire semblent montrer que le T4SS de TP114 peut initier le transfert d'ADN vers un plus large spectre de bactéries que ceux reconnus par son T4P. Il est donc raisonnable de supposer que le T4SS ne sera pas l'élément limitant immédiat pour le spectre d'action du système COP. Toutefois, la capacité du T4SS à transférer de l'ADN dans une cellule réceptrice peut être limitée via un phénomène nommé « exclusion plasmidique » (Garcillán-Barcia & De la Cruz, 2008). En effet, afin d'éviter le transfert redondant d'un même plasmide plusieurs fois dans une même réceptrice, les plasmides conjugatifs possèdent des protéines responsables de bloquer l'entrée de copies additionnelles de lui-même dans une cellule hôte (Garcillán-Barcia & De la Cruz, 2008). Ces protéines, appelées protéines d'exclusion, bloquent le transfert d'ADN via deux mécanismes distincts, soit par exclusion de surface ou par exclusion d'entrée

(Arutyunov & Frost, 2013). L'exclusion de surface prévient la reconnaissance de la cellule hôte par une autre cellule donneuse (Arutyunov & Frost, 2013). L'exclusion d'entrée quant à elle, bloque le transfert d'ADN en interagissant avec le T4SS via un mécanisme encore mal compris (Humbert *et al*, 2019). Comme les plasmides IncI₂ peuvent être naturellement présents dans les souches ciblées, comprendre l'exclusion de TP114 pourrait nous permettre de la contourner.

(3) Tout comme pour l'incapacité du T4SS à transférer de l'ADN, l'incapacité de TP114 de se répliquer une fois transférée chez une réceptrice ne produira pas (ou peu) de transconjugants tant sur agar qu'en bouillon. Par contre, nous pouvons discriminer entre les deux phénomènes en clonant les gènes responsables de la répllication de TP114 dans un petit plasmide et en le transformant par électroporation dans les cellules n'étant pas capables de recevoir le plasmide. Si l'électroporation est un succès, le transfert est probablement limité par le T4SS, sinon la répllication est probablement en cause. Alternativement, on peut cloner l'*oriT* de TP114 sur un plasmide connu pour sa répllication dans l'organisme d'intérêt. Si le plasmide synthétique peut être mobilisé, mais pas TP114, la répllication du plasmide est en faute. Discriminer entre les deux phénomènes a d'importantes implications pour le système COP puisque le bon fonctionnement du T4SS est essentiel au succès du traitement alors que la répllication de TP114 n'est pas nécessaire. En effet, la répllication du plasmide, une fois transféré, n'est pas essentielle à l'expression du module CRISPR-*cas9*. Une mesure de bioconfinement intéressante implique d'ailleurs la limitation de la répllication du plasmide transféré par séquestration de la protéine initiateur de la répllication dans le génome de la souche donneuse.

Outre les facteurs limitant le spectre d'hôte de TP114, plusieurs mécanismes peuvent influencer l'efficacité de transfert de TP114. La comptabilité des systèmes de restriction-modification est une barrière importante au transfert horizontal de gène (Vasu & Nagaraja, 2013). Bien qu'il soit possible par synthèse d'ADN d'éviter un grand nombre de sites de restriction, plusieurs systèmes de restriction (particulièrement de type I, III et IV) demeurent peu caractérisés et requièrent le séquençage des génomes des organismes ciblés (Johnston *et al*, 2019). Afin de surmonter ce problème, une base de données qui répertorie les sites de restriction-modifications

connus jusqu'à présent a été mise sur pied (Roberts *et al*, 2015). À terme, une telle base de données devrait nous permettre de concevoir des plasmides possédant un nombre minimal, voir aucun site de restriction. Une stratégie alternative consiste en l'expression de systèmes anti-restrictions (Shabbir *et al*, 2016). En effet, plusieurs bactériophages et plasmides conjugatifs portent des systèmes anti-restriction ayant des mécanismes variés comme le masquage de l'ADN, la méthylation des sites de restriction, ou la liaison aux facteurs de restriction (Liang *et al*, 2017; Wilkins, 2002). Une stratégie intéressante pour éviter que la restriction diminue l'efficacité du système COP serait d'utiliser un promoteur actif en simple brin tel le promoteur *ssi3* (Wilkins, 2002). En effet, les plasmides conjugatifs sont transférés sous forme simple brin dans la cellule réceptrice et synthétise le deuxième brin seulement après le transfert complet de l'ADN plasmidique. De ce fait, l'utilisation de promoteurs actifs en simple brin tel *ssi3* permettrait l'expression hâtive de CRISPR-Cas9 avant même que l'ADN ne soit double brin et susceptible aux systèmes de restriction-modification.

Enfin, le système COP, comme tout autre médicament basé sur des organismes vivants, est assujéti à des mutations. Si une mutation délétère affecte le module CRISPR-*cas9*, il est possible que le système COP puisse transférer l'ADN encodant le système CRISPR-Cas9, mais que ce dernier ne soit plus en mesure de cliver la séquence cible. Le taux de mutation chez *E. coli* K-12 atteint une fréquence moyenne de 10^{-11} /pb/réplication (Badran & Liu, 2015). À ce taux, la probabilité qu'une mutation se trouve dans l'ADN de TP114 ou du modèle CRISPR est de $\sim 7 \times 10^{-7}$ /réplication. Bien que cette probabilité soit faible, elle pourrait devenir problématique lors de la production de grande quantité de COP au long terme. Ce point doit donc être mitigé par des contrôles de qualité stricts et par l'utilisation de souches ayant un génome plus stable comme MDS42 (Csörgo *et al*, 2012; Pósfai *et al*, 2006).

5.5 Perspective

À terme, la technologie COP pourrait devenir un outil de grande importance pour la manipulation du microbiome et la lutte contre les phénomènes de résistances aux antibiotiques

tant en prévention qu'en contexte d'infection. La mobilisation de système CRISPR-Cas permet également de manière inédite de retirer des toxines de bactéries pathogènes. Il serait particulièrement intéressant d'évaluer le potentiel de la technologie pour le traitement d'infections à *E. coli* produisant la Shiga toxine. Mon travail sur la technologie COP a fait naître un axe de recherche complet au sein du laboratoire du Pr Sébastien Rodrigue. Plusieurs projets connexes à mon projet de doctorat sont maintenant menés par d'autres étudiants gradués afin de faire mûrir davantage cette technologie.

Les travaux présentés ci-haut ont également fait l'objet d'un dépôt de demande de brevet (WO2020010452A1). Cette demande de brevet, qui porte sur l'utilisation d'un plasmide capable de stabiliser la paire de conjugaison pour la mobilisation de gène vers des bactéries associées à un humain ou un animal, offre une position stratégique intéressante. C'est via cette stratégie de propriété intellectuelle, ainsi que la forte traction qui nous a valu deux prix à des compétitions nationales de présentation d'affaires en 2019, que nous avons pu (Sébastien Rodrigue, Jean-François Millau et moi) gagner de la crédibilité et fonder TATUM bioscience. Depuis, TATUM bioscience a été en mesure de sécuriser sa première levée de fond en phase de pré-amorçage, a reçu plusieurs subventions, recruté des employés qualifiés et déposé une nouvelle demande de brevet provisoire. TATUM bioscience vise à être un acteur majeur du développement de médicament via sa plateforme d'ingénierie de bactéries probiotiques. TATUM bioscience a pour mission de répondre à des besoins médicaux complexes grâce à des technologies de bactéries médicaments modulaires permettant d'attaquer les pathologies sous des angles inédits.

Le projet COP se poursuit donc sur deux fronts, soit au sein du laboratoire du Pr Sébastien Rodrigue ainsi que dans la compagnie TATUM bioscience. Suite à l'obtention de mon doctorat, j'aurai la chance de diriger le développement de cette technologie au sein de TATUM bioscience et de continuer de collaborer avec le laboratoire de Pr Sébastien Rodrigue. Ces efforts contribueront à l'essor des biothérapeutiques vivants, une nouvelle classe de médicaments de fort potentiel.

BIBLIOGRAPHIE

- Abedon ST, Kuhl SJ, Blasdel BG & Kutter EM (2011) Phage treatment of human infections. *Bacteriophage* 1: 66–85
- Al-Asmakh M & Zadjali F (2015) Use of germ-free animal models in microbiota-related research. *J Microbiol Biotechnol* 25: 1583–1588
- Allard N, Garneau D, Poulin-Laprade D, Burrus V, Brzezinski R & Roy S (2015) A diaminopimelic acid auxotrophic *Escherichia coli* donor provides improved counterselection following intergeneric conjugation with actinomycetes. *Can J Microbiol* 61: 565–74
- Allen HK, Moe LA, Rodbumrer J, Gaarder A & Handelsman J (2009) Functional metagenomics reveals diverse B-lactamases in a remote Alaskan soil. *ISME J* 3: 243–251
- Antimicrobial resistance: global report on surveillance 2014. Available online at <https://www.who.int/drugresistance/documents/surveillancereport/en/> (2014) WHO
- Arutyunov D & Frost LS (2013) F conjugation: Back to the beginning. *Plasmid* 70: 18–32
- Aviv G, Rahav G & Gal-mor O (2016) Horizontal Transfer of the *Salmonella enterica* Serovar Infantis Resistance and Virulence Plasmid pESI to the Gut Microbiota of. *Am Soc Microbiol* 7: 1–12
- Azad AK, Coote JG & Parton R (1994) Construction of conjugative shuttle and suicide vectors for *Pasteurella haemolytica* and *P. multocida*. *Gene* 145: 81–85
- Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL & Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* 2: doi:10.1038/msb4100050
- Bäckhed F, Manchester JK, Semenkovich CF & Gordon JI (2007) Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc Natl Acad Sci* 104: 979–984
- Badran AH & Liu DR (2015) Development of potent in vivo mutagenesis plasmids with broad mutational spectra. *Nat Commun* 6: 1–10
- Bassis CM, Tang AL, Young VB & Pynnonen MA (2014) The nasal cavity microbiota of healthy adults. *Microbiome* 2: 27–32
- Becker N, Kunath J, Loh G & Blaut M (2011) Human intestinal microbiota: Characterization of a simplified and stable gnotobiotic rat model. *Gut Microbes* 2: 25–33

- Belizário JE & Napolitano M (2015) Human microbiomes and their roles in dysbiosis, common diseases, and novel therapeutic approaches. *Front Microbiol* 6: 1–16
- Bennett S (2004) Solexa Ltd. *Pharmacogenomics* 5: 433–438
- Bibiloni R (2012) Rodent models to study the relationships between mammals and their bacterial inhabitants. *Gut Microbes* 3: 536–543
- Bik EM, Eckburg PB, Gill SR, Nelson KE, Purdom EA, Francois F, Perez-Perez G, Blaser MJ & Relman DA (2006) Molecular analysis of the bacterial microbiota in the human stomach. *Proc Natl Acad Sci U S A* 103: 732–737
- Bikard D, Euler C, Jiang W, Nussenzweig PM, Gregory W, Duportet X, Fischetti VA & Marraffini LA (2014) Development of sequence-specific antimicrobials based on programmable CRISPR-Cas nucleases. *Nat Biotechnol* 32: 1146–1150
- Bird AW, Erler A, Fu J, Hériché JK, Maresca M, Zhang Y, Hyman AA & Stewart AF (2012) High-efficiency counterselection recombineering for site-directed mutagenesis in bacterial artificial chromosomes. *Nat Methods* 9: 103–109
- Bondy-Denomy J, Pawluk A, Maxwell KL & Davidson AR (2013) Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493: 429–432
- Bonnet J, Subsoontorn P & Endy D (2012) Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proc Natl Acad Sci* 109: 8884–8889
- Braat H, Rottiers P, Hommes DW, Huyghebaert N, Remaut E, Remon JP, van Deventer SJH, Neiryneck S, Peppelenbosch MP & Steidler L (2006) A Phase I Trial With Transgenic Bacteria Expressing Interleukin-10 in Crohn's Disease. *Clin Gastroenterol Hepatol* 4: 754–759
- Bradley D, Taylor DE & Cohen DR (1980) Specification of Surface Mating Systems Among Conjugative Drug Resistance Plasmids in *Escherichia coli* K-12. *J Bacteriol* 143: 1466–1470
- Bradley DE (1984) Characteristics and function of thick and thin conjugative pili determined by transfer-derepressed plasmids of incompatibility groups I1, I2, I5, B, K and Z. *J Gen Microbiol* 130: 1489–1502
- Brand MW, Wannemuehler MJ, Phillips GJ, Proctor A, Overstreet AM, Jergens AE, Orcutt RP & Fox JG (2015) The altered Schaedler flora: Continued applications of a defined murine microbial community. *ILAR J* 56: 169–178
- Braun-Fahrlander C, Reidler J, Herz U, Eder W, Waser M, Grize L, Maisch S, Carr D, Gerlach F, Bufe A, *et al* (2002) ENVIRONMENTAL EXPOSURE TO ENDOTOXIN AND ITS RELATION TO ASTHMA IN SCHOOL-AGE CHILDREN. *New Engl J Med* 347: 869–877

- Brouwer MSM, Tagg KA, Mevius DJ, Iredell JR, Bossers A, Smith HE & Partridge SR (2015) IncI shufflons: Assembly issues in the next-generation sequencing era. *Plasmid* 80: 111–117
- Brunet YR, Bernard CS, Gavioli M, Llobès R & Cascales E (2011) An epigenetic switch involving overlapping fur and DNA methylation optimizes expression of a type VI secretion gene cluster. *PLoS Genet* 7
- Bry L, Falk PG, Midtvedt T & Gordon JI (1996) A model of host-microbial interactions in an open mammalian ecosystem. *Science* 273: 1380–1383
- Buelow E, Gonzalez TB, Versluis D, Oostdijk EAN, Ogilvie LA, van Mourik MSM, Oosterink E, van Passel MWJ, Smidt H, D'Andrea MM, *et al* (2014) Effects of selective digestive decontamination (SDD) on the gut resistome. *J Antimicrob Chemother* 69: 2215–2223
- Cachat E & Davies JA (2011) Application of Synthetic Biology to Regenerative Medicine. *J Bioeng Biomed Sci* S2: 003
- Cameron DE, Bashor CJ & Collins JJ (2014) A brief history of synthetic biology. *Nat Rev Microbiol* 12: 381–390
- Carpenedo RL, Bratt-Leal AM, Marklein RA, Seaman SA, Bowen NJ, McDonald JF & McDevitt TC (2009) Homogeneous and organized differentiation within embryoid bodies induced by microsphere-mediated delivery of small molecules. *Biomaterials* 30: 2507–2515
- Carraro N, Matteau D, Luo P & Burrus V (2014) The Master Activator of IncA / C Conjugative Plasmids Stimulates Genomic Islands and Multidrug Resistance Dissemination. *PLoS Genet* 10: e1004714
- Casals-Pascual C, Vergara A & Vila J (2018) Intestinal microbiota and antibiotic resistance: Perspectives and solutions. *Hum Microbiome J* 9: 11–15
- Cash HL, Whitham C V, Behrendt CL & Hooper L V (2006) Symbiotic bacteria direct expression of an intestinal bacterial Lectin. *Science* 313: 1126–1130
- Chabé M, Lokmer A & Ségurel L (2017) Gut Protozoa: Friends or Foes of the Human Gut Microbiota? *Trends Parasitol* 33: 925–934
- Chan SH, Stoddard BL & Xu SY (2011) Natural and engineered nicking endonucleases - From cleavage mechanism to engineering of strand-specificity. *Nucleic Acids Res* 39: 1–18
- Charbonneau MR, Isabella VM, Li N & Kurtz CB (2020) Developing a new class of engineered live bacterial therapeutics to treat human diseases. *Nat Commun* 11: 1–11
- Charpentier E, Richter H, van der Oost J & White MF (2015) Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol Rev* 39: 428–441

- Chen JS, Ma E, Harrington LB, Tian X & Doudna JA (2018) CRISPR-Cas12a target binding unleashes single-stranded DNase activity. *Science* 360: 436–439
- Chen L, Chavda KD, Laham N Al, Melano RG, Jacobs MR, Bonomo RA & Kreiswirth BN (2013) Complete Nucleotide Sequence of a *blaKPC*-Harboring IncI2 Plasmid and Its Dissemination in New Jersey and New York Hospitals. *Antimicrob Agents Chemother* 57: 5019–5025
- Chowdhury S, Castro S, Coker C, Hinchliffe TE, Arpaia N & Danino T (2019) Programmable bacteria induce durable tumor regression and systemic antitumor immunity. *Nat Med* 25: 1057–1063
- Citorik RJ, Mimee M & Lu TK (2014) Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat Biotechnol* 32: 1141–1145
- Cohen SN, Chang ACY, Boyert HW & Helling RB (1973) Construction of Biologically Functional Bacterial Plasmids *In Vitro*. *Proc Natl Acad Sci* 70: 3240–3244
- Couturier M, Bex F, Bergquist PL & Maas WK (1988) Identification and classification of bacterial plasmids. *Microbiol Rev* 52: 375–395
- Cox LM, Yamanishi S, Sohn J, Alekseyenko A V., Leung JM, Cho I, Kim SG, Li H, Gao Z, Mahana D, *et al* (2014) Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. *Cell* 158: 705–721
- Crepin VF, Collins JW, Habibzay M & Frankel G (2016) *Citrobacter rodentium* mouse model of bacterial infection. *Nat Protoc* 11: 1851–1876
- Cronin M, Akin AR, Collins SA, Meganck J, Kim J, Baban CK, Joyce SA, Dam GM Van, Zhang N, Sinderen D Van, *et al* (2012) High Resolution *In Vivo* Bioluminescent Imaging for the Study of Bacterial Tumour Targeting. *PLoS One* 7: e30940
- Csörgo B, Fehér T, Tímár E, Blattner FR & Pósfai G (2012) Low-mutation-rate , reduced-genome *Escherichia coli*: an improved host for faithful maintenance of engineered genetic constructs. *Microb Cell Fact* 11: 1–13
- Cui L & Bikard D (2016) Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res* 44: 4243–4251
- D'Costa VM, McGrann KM, Hughes DW & Wright GD (2006) Sampling the Antibiotic Resistome. *Science* 311: 374–377
- Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* 97: 6640–6645

- Datta S, Costantino N & Court DL (2006) A set of recombineering plasmids for gram-negative bacteria. *Gene* 379: 109–115
- Davies J & Davies D (2010) Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev* 74: 417–433
- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J & Charpentier E (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471: 602–607
- Dinan TG, Stanton C & Cryan JF (2013) Psychobiotics: A novel class of psychotropic. *Biol Psychiatry* 74: 720–726
- Doucet-Populaire F, Trieu-Cuot P, Dosbaa I, Andremont A & Courvalin P (1991) Inducible transfer of conjugative transposon Tn1545 from *Enterococcus faecalis* to *Listeria monocytogenes* in the digestive tracts of gnotobiotic mice. *Antimicrob Agents Chemother* 35: 185–187
- Doucet-Populaire F, Trieu-Cuot P, Andremont A & Courvalin P (1992) Conjugal Transfer of Plasmid DNA from *Enterococcus faecalis* to *Escherichia coli* in Digestive Tracts of Gnotobiotic Mice. *Antimicrob Agents Chemother* 36: 502–504
- Drell T, Lillsaar T, Tummeleht L, Simm J, Aaspõllu A, Väin E, Saarma I, Salumets A, Donders GGG & Metsis M (2013) Characterization of the Vaginal Micro- and Mycobiome in Asymptomatic Reproductive-Age Estonian Women. *PLoS One* 8: e54379
- Duan F, Curtis KL & March JC (2008) Secretion of insulinotropic proteins by commensal bacteria: Rewiring the gut to treat diabetes. *Appl Environ Microbiol* 74: 7437–7438
- Duan F & March JC (2010) Engineered bacterial communication prevents *Vibrio cholerae* virulence in an infant mouse model. *Proc Natl Acad Sci* 107: 11260–11264
- Dupuis MÈ, Villion M, Magadán AH & Moineau S (2013) CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nat Commun* 4: 1–7
- Dusko Ehrlich S (2010) Metagenomics of the intestinal microbiota: potential applications. *Gastroenterol Clin Biol* 34: S23–S28
- Dwyer MA & Hellinga HW (2004) Periplasmic binding proteins: A versatile superfamily for protein engineering. *Curr Opin Struct Biol* 14: 495–504
- Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE & Relman DA (2005) Diversity of the Human Intestinal Microbial Flora. *Science* 308: 1635–1638

- Ellekilde M, Selfjord E, Larsen CS, Jakesevic M, Rune I, Tranberg B, Vogensen FK, Nielsen DS, Bahl MI, Licht TR, *et al* (2015) Transfer of gut microbiota from lean and obese mice to antibiotic-treated mice. *Sci Rep* 4: 5922
- Engler C, Kandzia R & Marillonnet S (2008) A One Pot , One Step , Precision Cloning Method with High Throughput Capability. *PLoS One* 3: e3647
- English MA, Soenksen LR, Gayet RV, de Puig H, Angenent-Mari NM, Mao AS, Nguyen PQ & Collins JJ (2019) Programmable CRISPR-responsive smart materials. *Science* 365: 780–785
- Featherstone C & Jackson SP (1999) DNA double-strand break repair. *Curr Biol* 9: 759–761
- Ferrières L, Hémerly G, Nham T, Guérout AM, Mazel D, Beloin C & Ghigo JM (2010) Silent mischief: Bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *J Bacteriol* 192: 6418–6427
- Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Becker J, *et al* (2013) Topographic diversity of fungal and bacterial communities in human skin. *Nature* 498: 367–370
- Firth N, Ippen-ihler K & Skurray R a (1996) Structure and Function of the F Factor and Mechanism of Conjugation. *Cell Mol Biol*: 2377–2401
- Forbes NS (2010) Engineering the perfect (bacterial) cancer therapy. *Nat Rev Cancer* 10: 785–794
- Forslund K, Sunagawa S, Coelho LP & Bork P (2014) Metagenomic insights into the human gut resistome and the forces that shape it. *BioEssays* 36: 316–329
- Forslund K, Sunagawa S, Kultima JR, Mende DR, Arumugam M, Typas A & Bork P (2013) Country-specific antibiotic use practices impact the human gut resistome. *Genome Res* 23: 1163–1169
- Frantz LAF, Mullin VE, Pionnier-Capitan M, Lebrasseur O, Ollivier M, Perri A, Linderholm A, Mattiangeli V, Teasdale MD, Dimopoulos EA, *et al* (2016) Genomic and archaeological evidence suggests a dual origin of domestic dogs. *Science* 352: 1228–1231
- Fu YHF, Tsai MM, Luo Y & Deonier RC (1991) Deletion analysis of the F plasmid *oriT* locus. *J Bacteriol* 173: 1012–1020
- Fujino T, Tozaki M & Murakami H (2020) An Amino Acid-Swapped Genetic Code. *ACS Synth Biol* 9: 2703–2713
- Furuya EY & Lowy FD (2006) Antimicrobial-resistant bacteria in the community setting. *Nat Rev Microbiol* 4: 36–45

- Ganeshan SD & Hosseinidoust Z (2019) Phage therapy with a focus on the human microbiota. *Antibiotics* 8: 1–19
- Gao Z, Tseng C -h., Pei Z & Blaser MJ (2007) Molecular analysis of human forearm superficial skin bacterial biota. *Proc Natl Acad Sci* 104: 2927–2932
- García-Quintanilla M, Ramos-Morales F & Casadesús J (2008) Conjugal transfer of the *Salmonella enterica* virulence plasmid in the mouse intestine. *J Bacteriol* 190: 1922–1927
- Garcillán-Barcia MP, Alvarado A & De la Cruz F (2011) Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol Rev* 35: 936–956
- Garcillán-Barcia MP, Francia MV & De La Cruz F (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* 33: 657–687
- Garcillán-Barcia MP & De la Cruz F (2008) Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid* 60: 1–18
- Garneau JE, Dupuis MÈ, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH & Moineau S (2010) The CRISPR/cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468: 67–71
- Gasiunas G, Barrangou R, Horvath P & Siksnys V (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* 109: 2579–2586
- Genetically Engineered Crops: Experiences and Prospects (2016)
- Getino M & de la Cruz F (2018) Natural and Artificial Strategies to Control the Conjugative Transmission of Plasmids. *Microb Spectr* 6: MTBP-0015-2016
- Geu-flores F, Nour-eldin HH, Nielsen MT & Halkier BA (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res* 35: e55
- Ghannoum MA, Jurevic RJ, Mukherjee PK, Cui F, Sikaroodi M, Naqvi A & Gillevet PM (2010) Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. *PLoS Pathog* 6: e1000713
- Ghosh P, Bibb LA & Hatfull GF (2008) Two-step site selection for serine-integrase-mediated excision: DNA-directed integrase conformation and central dinucleotide proofreading. *Proc Natl Acad Sci U S A* 105: 3238–3243
- Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA & Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6: 343–345

- Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM & Nelson KE (2006) Metagenomic analysis of the human distal gut microbiome. *Science* 312: 1355–1359
- Gitzinger M, Kemmer C, El-Baba MD, Weber W & Fussenegger M (2009) Controlling transgene expression in subcutaneous implants using a skin lotion containing the apple metabolite phloretin. *Proc Natl Acad Sci U S A* 106: 10638–10643
- Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G & Gordon JI (2011) Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci* 108: 6252–6257
- Gootenberg JS, Abudayyeh OO, Kellner MJ, Joung J, Collins JJ & Zhang F (2018) Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a and Csm6. *Science* 360: 439–444
- Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, Verdine V, Donghia N, Daringer NM, Freije CA, *et al* (2017) Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 356: 438–442
- Greber D, El-Baba MD & Fussenegger M (2008) Intronic encoded siRNAs improve dynamic range of mammalian gene regulation systems and toggle switch. *Nucleic Acids Res* 36: e101
- Green AA, Silver PA, Collins JJ & Yin P (2014) Toehold switches: *De-novo*-designed regulators of gene expression. *Cell* 159: 925–939
- Grimm V, Radulovic K & Riedel CU (2015) Colonization of C57BL/6 mice by a potential probiotic *Bifidobacterium bifidum* strain under germ-free and specific pathogen-free conditions and during experimental colitis. *PLoS One* 10: 1–14
- Gumpert H, Kubicek-Sutherland JZ, Porse A, Karami N, Munck C, Linkevicius M, Adlerberth I, Wold AE, Andersson DI & Sommer MOA (2017) Transfer and Persistence of a Multi-Drug Resistance Plasmid *in situ* of the Infant Gut Microbiota in the Absence of Antibiotic Treatment. *Front Microbiol* 8: 1–10
- Hall BG & Barlow M (2004) Evolution of the serine β -lactamases: Past, present and future. *Drug Resist Updat* 7: 111–123
- Hansen AK (1995) Antibiotic treatment of nude rats and its impact on the aerobic bacterial flora. *Lab Anim* 29: 37–44
- Hansen CHF, Krych L, Nielsen DS, Vogensen FK, Hansen LH, Sørensen SJ, Buschard K & Hansen AK (2012) Early life treatment with vancomycin propagates *Akkermansia muciniphila* and reduces diabetes incidence in the NOD mouse. *Diabetologia* 55: 2285–2294

- Heler R, Samai P, Modell JW, Weiner C, Gregory W, Bikard D & Marraffini LA (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519: 199–202
- Hill WG & Caballero A (1992) Artificial selection experiments. *Annu Rev Ecol Syst* 23: 287–310
- Hirokawa Y, Kawano H, Tanaka-Masuda K, Nakamura N, Nakagawa A, Ito M, Mori H, Oshima T & Ogasawara N (2013) Genetic manipulations restored the growth fitness of reduced-genome *Escherichia coli*. *J Biosci Bioeng* 116: 52–58
- Hochhut B, Marrero J & Waldor MK (2000) Mobilization of plasmids and chromosomal DNA mediated by the SXT element, a constin found in *Vibrio cholerae* O139. *J Bacteriol* 182: 2043–2047
- Hoffmann C, Dollive S, Grunberg S, Chen J, Li H, Wu GD, Lewis JD & Bushman FD (2013) Archaea and Fungi of the Human Gut Microbiome: Correlations with Diet and Bacterial Residents. *PLoS One* 8: e66019
- de Hoon MJL, Makita Y, Nakai K & Miyano S (2005) Prediction of Transcriptional Terminators in *Bacillus subtilis* and Related Species. *PLoS Comput Biol* 1: e25
- Howard MT, Nelson WC & Matson SW (1995) Stepwise assembly of a relaxosome at the F plasmid origin of transfer. *J Biol Chem* 270: 28381–28386
- Hu Y, Peng J, Tai N, Hu C, Zhang X, Wong FS & Wen L (2015) Maternal Antibiotic Treatment Protects Offspring from Diabetes Development in Nonobese Diabetic Mice by Generation of Tolerogenic APCs. *J Immunol* 195: 4176–4184
- Hu Y, Yang X, Qin J, Lu N, Cheng G, Wu N, Pan Y, Li J, Zhu L, Wang X, *et al* (2013) Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* 4: DOI: 10.1038/ncomms3151
- Hughes RA & Ellington AD (2017) Synthetic DNA synthesis and assembly: Putting the synthetic in synthetic biology. *Cold Spring Harb Perspect Biol* 9: a023812
- Humbert M, Huguet KT, Coulombe F & Burrus V (2019) Entry exclusion of conjugative plasmids of the IncA, IncC, and related untyped incompatibility groups. *J Bacteriol* 201: e00731-18
- Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, Chinwalla AT, Creasy HH, Earl AM, FitzGerald MG, Fulton RS, *et al* (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486: 207–214
- Hutton ML, Mackin KE, Chakravorty A & Lyras D (2014) Small animal models for the study of *Clostridium difficile* disease pathogenesis. *FEMS Microbiol Lett* 352: 140–149

Hwang IY, Tan MH, Koh E, Ho CL, Poh CL & Chang MW (2014) Reprogramming microbes to be pathogen-Seeking killers. *ACS Synth Biol* 3: 228–237

Hwang S & Maxwell KL (2019) Meet the Anti-CRISPRs: Widespread Protein Inhibitors of CRISPR-Cas Systems. *Cris J* 2: 23–30

Hyman RW, Fukushima M, Diamond L, Kumm J, Giudice LC & Davis RW (2005) Microbes on the human vaginal epithelium. *Proc Natl Acad Sci* 102: 7952–7957

Igimi S, Ryu CH, Park SH, Sasaki Y, Sasaki T & Kumagai S (1996) Transfer of conjugative plasmid pAM beta 1 from *Lactococcus lactis* to mouse intestinal bacteria. *Lett Appl Microbiol* 23: 31–35

Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, Tolonen AC, Gianoulis TA, Goodman DB, Reppas NB, *et al* (2011) Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* 333: 348–353

Isabella VM, Ha BN, Castillo MJ, Lubkowicz DJ, Rowe SE, Millet YA, Anderson CL, Li N, Fisher AB, West KA, *et al* (2018) Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. *Nat Biotechnol* 36: 857–867

Ishiwa A & Komano T (2000) The lipopolysaccharide of recipient cells is a specific receptor for PilV proteins, selected by shufflon DNA rearrangement, in liquid matings with donors bearing the R64 plasmid. *Mol Gen Genet* 263: 159–164

Ishiwa A & Komano T (2003) Thin pilus PilV adhesins of plasmid R64 recognize specific structures of the lipopolysaccharide molecules of recipient cells. *J Bacteriol* 185: 5192–5199

Ishiwa A & Komano T (2004) PilV adhesins of plasmid R64 thin pili specifically bind to the lipopolysaccharides of recipient cells. *J Mol Biol* 343: 615–625

Jackson D a, H. R & Berg P (1972) Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of *Escherichia coli*. *Proc Natl Acad Sci* 69: 2904–2909

Jain A & Srivastava P (2013) Broad host range plasmids. *FEMS Microbiol Lett* 348: 87–96

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA & Charpentier E (2012) A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337: 816–821

Johnston CD, Cotton SL, Rittling SR, Starr JR, Borisy GG, Dewhirst FE & Lemon KP (2019) Systematic evasion of the restriction-modification barrier in bacteria. *Proc Natl Acad Sci U S A* 166: 11454–11459

- Juhas M, Van Der Meer JR, Gaillard M, Harding RM, Hood DW & Crook DW (2009) Genomic islands: Tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* 33: 376–393
- Kamruzzaman M, Shoma S, Thomas CM, Partridge SR & Iredell JR (2017) Plasmid interference for curing antibiotic resistance plasmids *in vivo*. *PLoS One* 12: 1–20
- Kellogg EA (1997) Plant evolution: The dominance of maize. *Curr Biol* 7: 411–413
- Kemmer C, Gitzinger M, Daoud-El Baba M, Djonov V, Stelling J & Fussenegger M (2010) Self-sufficient control of urate homeostasis in mice by a synthetic circuit. *Nat Biotechnol* 28: 355–360
- Kim SR & Komano T (1992) Nucleotide sequence of the R721 shufflon. *J Bacteriol* 174: 7053–7058
- Koonin E V. & Makarova KS (2019) Origins and evolution of CRISPR-Cas systems. *Philos Trans B* 374: 1–16
- Kotula JW, Kerns SJ, Shaket LA, Siraj L, Collins JJ, Way JC & Silver PA (2014) Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc Natl Acad Sci U S A* 111: 4838–4843
- Kozyrskyj AL, Ernst P & Becker AB (2007) Increased risk of childhood asthma from antibiotic use in early life. *Chest* 131: 1753–1759
- Kramer BP, Fischer C & Fussenegger M (2004a) BioLogic gates enable logical transcription control in mammalian cells. *Biotechnol Bioeng* 87: 478–484
- Kramer BP, Viretta AU, Baba MD-E, Aubel D, Weber W & Fussenegger M (2004b) An engineered epigenetic transgene switch in mammalian cells. *Nat Biotechnol* 22: 867–870
- Kuhlman TE & Cox EC (2010) Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Res* 38: e92
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, Fitzhugh W, *et al* (2001) Initial sequencing and analysis of the human genome: International Human Genome Sequencing Consortium. *Nature* 412: 565–566
- Langdon A, Crook N & Dantas G (2016) The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome Med* 8: 39–55
- Laukens D, Brinkman BM, Raes J, De Vos M & Vandenabeele P (2015) Heterogeneity of the gut microbiome in mice: Guidelines for optimizing experimental design. *FEMS Microbiol Rev* 40: 117–132

- Lawley TD & Walker AW (2012) Intestinal colonization resistance. *Immunology* 138: 1–11
- Lazdins A, Maurya AP, Miller CE, Kamruzzaman M, Liu S, Stephens ER, Lloyd GS, Haratianfar M, Chamberlain M, Haines AS, *et al* (2020) Potentiation of curing by a broad-host-range self-transmissible vector for displacing resistance plasmids to tackle AMR. *PLoS One* 15: 1–23
- Leavitt JM & Alper HS (2015) Advances and current limitations in transcript-level control of gene expression. *Curr Opin Biotechnol* 34: 98–104
- Lee DJ, Bingle LEH, Heurlier K, Pallen MJ, Penn CW, Busby SJW & Hobman JL (2009) Gene doctoring : a method for recombineering in laboratory and pathogenic *Escherichia coli* strains. *BMC Microbiol* 9: 1–14
- Leventhal DS, Sokolovska A, Li N, Plescia C, Kolodziej SA, Gallant CW, Christmas R, Gao JR, James MJ, Abin-Fuentes A, *et al* (2020) Immunotherapy with engineered bacteria by targeting the STING pathway for anti-tumor immunity. *Nat Commun* 11: 1–15
- Levsikaya A, Weiner OD, Lim WA & Voigt CA (2009) Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* 461: 997–1001
- Levy A, Goren MG, Yosef I, Auster O, Manor M, Amitai G, Edgar R, Qimron U & Sorek R (2015) CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* 520: 505–510
- Ley R, Turnbaugh P, Klein S & Gordon J (2006) Microbial ecology: human gut microbes associated with obesity. *Nature* 444: 1022–3
- Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD & Gordon JI (2005) Obesity alters gut microbial ecology. *Proc Natl Acad Sci* 102: 11070–11075
- Li M, Wang R, Zhao D & Xiang H (2014) Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process. *Nucleic Acids Res* 42: 2483–2492
- Liang W, Xie Y, Xiong W, Tang Y, Li G, Jiang X & Lu Y (2017) Anti-Restriction Protein, KlcAHS, Promotes Dissemination of Carbapenem Resistance. *Front Cell Infect Microbiol* 7: 150
- Licht TR, Christensen BB & Krogfelt KA (1999) Plasmid transfer in the animal intestine and other dynamic bacterial populations: the role of community structure and environment. *Microbiology* 145: 2615–2622
- Llosa M, Gomis-Rüth FX, Coll M & De la Cruz F (2002) Bacterial conjugation: A two-step mechanism for DNA transport. *Mol Microbiol* 45: 1–8

- Lloyd-Price J, Abu-Ali G & Huttenhower C (2016) The healthy human microbiome. *Genome Med* 8: 51
- Lloyd CJ, Ebrahim A, Yang L, King ZA, Catoi E, O'Brien EJ, Liu JK & Palsson BO (2018) COBRAME: A computational framework for genome-scale models of metabolism and gene expression. *PLoS Comput Biol* 14: 1–14
- Loessner MJ, Rees CED, Stewart GSAB & Scherer S (1996) Construction of luciferase reporter bacteriophage A511 :: *luxAB* for rapid and sensitive detection of viable *Listeria* cells. *Appl Environ Microbiol* 62: 1133–1140
- López-Igual R, Bernal-Bayard J, Rodríguez-Patón A, Ghigo JM & Mazel D (2019) Engineered toxin–intein antimicrobials can selectively target and kill antibiotic-resistant bacteria in mixed populations. *Nat Biotechnol* 37: 755–760
- Lu TK & Collins JJ (2007) Dispersing biofilms with engineered enzymatic bacteriophage. *Proc Natl Acad Sci* 104: 11197–11202
- Luczynski P, Neufeld KAMV, Oriach CS, Clarke G, Dinan TG & Cryan JF (2016) Growing up in a bubble: Using germ-free animals to assess the influence of the gut microbiota on brain and behavior. *Int J Neuropsychopharmacol* 19: 1–17
- Lundberg R, Toft MF, August B, Hansen AK & Hansen CHF (2016) Antibiotic-treated versus germ-free rodents for microbiota transplantation studies. *Gut Microbes* 7: 68–74
- Machado AMD & Sommer MOA (2014) Human intestinal cells modulate conjugational transfer of multidrug resistance plasmids between clinical *Escherichia coli* isolates. *PLoS One* 9: e100739
- Macinga DR & Rather PN (1999) The chromosomal 2'-N-acetyltransferase of *Providencia stuartii*: physiological functions and genetic regulation. *Front Biosci* 4: 132–140
- Mahowald MA, Rey FE, Seedorf H, Turnbaugh PJ, Fulton RS, Wollam A, Shah N, Wang C, Magrini V, Wilson RK, *et al* (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc Natl Acad Sci* 106: 5859–5864
- Mali P, Esvelt KM & Church GM (2013) Cas9 as a versatile tool for engineering biology. *Nat Methods* 10: 957–963
- Mandell DJ, Lajoie MJ, Mee MT, Takeuchi R, Kuznetsov G, Norville JE, Gregg CJ, Stoddard BL & Church GM (2015) Biocontainment of genetically modified organisms by synthetic protein design. *Nature* 518: 55–60
- Manrique P, Bolduc B, Walk ST, van der Oost J, de Vos WM & Young MJ (2016) Healthy human gut phageome. *Proc Natl Acad Sci* 113: 10400–10405

- Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA & Ferrari R (2018) Genome, transcriptome and proteome: The rise of omics data and their integration in biomedical sciences. *Brief Bioinform* 19: 286–302
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bembem LA, Berka J, Braverman MS, Chen YJ, Chen Z, *et al* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380
- Martin VJJ, Piteral DJ, Withers ST, Newman JD & Keasling JD (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol* 21: 796–802
- Mathur R, Oh H, Zhang D, Park S, Seo J, Hayden MS & Ghosh S (2012) A mouse model of *Salmonella typhi* infection. *Cell* 151: 590–602
- McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, *et al* (2013) The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* 57: 3348–3357
- Medema MH, Blin K, Cimermancic P, De Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E & Breitling R (2011) AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* 39: 339–346
- Meslin JC, Sacquet E & Guenet J (1973) ACTION DE LA FLORE BACTÉRIENNE SUR LA MORPHOLOGIE ET LA SURFACE DE LA MUQUEUSE DE L'INTESTIN GRÈLE DU RAT. *Ann Biol Anim* 13: 203–214
- Modi SR, Lee HH, Spina CS & Collins JJ (2013) Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499: 219–222
- Modrich P & Lehman IR (1973) Deoxyribonucleic Acid Ligase. *J Biol Chem* 248: 7502–7511
- Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin E V. & Van Der Oost J (2016) Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353:aad5147
- Morrow JF, Cohen SN, Chang ACY, Boyer HW, Goodman HM & Helling RB (1974) Replication and Transcription of Eukaryotic DNA in *Escherichia coli*. *Proc Natl Acad Sci* 71: 1743–1747
- Mutalik VK, Guimaraes JC, Cambray G, Lam C, Christoffersen MJ, Mai QA, Tran AB, Paull M, Keasling JD, Arkin AP, *et al* (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat Methods* 10: 354–360

- Neil K, Allard N, Grenier F, Burrus V & Rodrigue S (2020) Highly efficient gene transfer in the mouse gut microbiota is enabled by the Incl2 conjugative plasmid TP114. *Commun Biol* 3: 1–9
- Neumann H & Neumann-Staubitz P (2010) Synthetic biology approaches in drug discovery and pharmaceutical biotechnology. *Appl Microbiol Biotechnol* 87: 75–86
- Nicklas W, Baneux P, Boot R, Decelle T, Deeny AA, Fumanelli M & Illgen-Wilcke B (2002) Recommendations for the health monitoring of rodent and rabbit colonies in breeding and experimental units. *Lab Anim* 36: 20–42
- Nobrega FL, Costa AR, Santos JF, Siliakus MF, Lent JWM Van, Kengen SWM, Azeredo J & Kluskens LD (2016) Genetically manipulated phages with improved pH resistance for oral administration in veterinary medicine. *Sci Rep* 6: 1–12
- Nordgård L, Brusetti L, Raddadi N, Traavik T, Averhoff B & Nielsen KM (2012) An investigation of horizontal transfer of feed introduced DNA to the aerobic microbiota of the gastrointestinal tract of rats. *BMC Res Notes* 5
- Norman A, Hansen LH & Sorensen SJ (2009) Conjugative plasmids: vessels of the communal gene pool. *Philos Trans R Soc L B Biol Sci* 364: 2275–2289
- Núñez B & De la Cruz F (2001) Two atypical mobilization proteins are involved in plasmid CloDF13 relaxation. *Mol Microbiol* 39: 1088–1099
- Nuñez JK, Lee ASY, Engelman A & Doudna JA (2015) Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* 519: 193–198
- Ott LC, Stromberg ZR, Redweik GAJ, Wannemuehler MJ & Mellata M (2020) Mouse Genetic Background Affects Transfer of an Antibiotic Resistance Plasmid in the Gastrointestinal Tract. *mSphere* 5: 1–14
- Pardee K, Green AA, Ferrante T, Cameron DE, Daleykeyser A, Yin P & Collins JJ (2014) Paper-based synthetic gene networks. *Cell* 159: 940–954
- Pei R & Lamas-Samanamud GR (2014) Inhibition of biofilm formation by T7 bacteriophages producing quorum-quenching enzymes. *Appl Environ Microbiol* 80: 5340–5348
- Pei Z, Bini EJ, Yang L, Zhou M, Francois F & Blaser MJ (2004) Bacterial biota in the human distal esophagus. *Proc Natl Acad Sci U S A* 101: 4250–4255
- Pereira VB, Saraiva TDL, Souza BM, Zurita-Turk M, Azevedo MSP, De Castro CP, Mancha-Agresti P, dos Santos JSC, Santos ACG, Faria AMC, *et al* (2014) Development of a new DNA vaccine based on mycobacterial ESAT-6 antigen delivered by recombinant invasive *Lactococcus lactis* FnBPA+. *Appl Microbiol Biotechnol* 99: 1817–1826

- Pérez-Cobas AE, Artacho A, Knecht H, Ferrús ML, Friedrichs A, Ott SJ, Moya A, Latorre A & Gosalbes MJ (2013) Differential effects of antibiotic therapy on the structure and function of human gut microbiota. *PLoS One* 8: e80201
- Phillips AW, Newcombs HR, Lachapelle R & Balish E (1962) Rearing of germfree and monocontaminated chicks in rigid plastic isolators. *Appl Microbiol* 10: 224–30
- Pilmis B, Le Monnier A & Zahar JR (2020) Gut microbiota, antibiotic therapy and antimicrobial resistance: A narrative review. *Microorganisms* 8: 1–17
- Pinero-Lambea C, Ruano-Gallego D & Angel Fernandez L (2015) Engineered bacteria as therapeutic agents. *Curr Opin Biotechnol* 35: 94–102
- Pines G, Freed EF, Winkler JD & Gill RT (2015) Bacterial Recombineering: Genome Engineering via Phage-Based Homologous Recombination. *ACS Synth Biol* 4: 1176–1185
- Poolman JT & Wacker M (2016) Extraintestinal pathogenic *Escherichia coli*, a common human pathogen: challenges for vaccine development and progress in the field. *J Infect Dis* 213: 6–13
- Pósfai G, III GP, Fehér T, Frisch D, Keil GM, Umenhoffer K, Kolisnychenko V, Stahl B, Sharma SS, Arruda M de, *et al* (2006) Emergent Properties of Reduced-Genome *Escherichia coli*. *Science* 312: 1044–1046
- Praszkier J & Pittard AJ (2005) Control of replication in I-complex plasmids. *Plasmid* 53: 97–112
- Purnick PEM & Weiss R (2009) The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol* 10: 410–422
- Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, *et al* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464: 59–65
- Rapid amplification of 5' complementary DNA ends (5' RACE) (2005) *Nat Methods* 2: 629–630
- Rawls JF, Mahowald MA, Ley RE & Gordon JI (2006) Reciprocal Gut Microbiota Transplants from Zebrafish and Mice to Germ-free Recipients Reveal Host Habitat Selection. *Cell* 127: 423–433
- Rawls JF, Samuel BS & Gordon JI (2004) From The Cover: Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. *Proc Natl Acad Sci* 101: 4596–4601
- Rawson DM, Willmer AJ & Turner APP (1989) Whole-cell biosensors for environmental monitoring. *Biosensors* 4: 299–311

- Reisch CR & Prather KLJ (2015) The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Sci Rep* 5: 15096
- Rettedal EA, Gumpert H & Sommer MOA (2014) Cultivation-based multiplex phenotyping of human gut microbiota allows targeted recovery of previously uncultured bacteria. *Nat Commun* 5: 4714
- Rezzonico E, Mestdagh R, Delley M, Combremont S, Dumas ME, Holmes E, Nicholson J & Bibiloni R (2011) Bacterial adaptation to the gut environment favors successful colonization: microbial and metabonomic characterization of a simplified microbiota mouse model. *Gut Microbes* 2: 307–318
- Rhee, Sang H, Pothoulakis, C and Mayer EA (2009) Principles and clinical implications of the brain-gut-enteric microbiota axis. *Nat Rev Gastroenterol Hepatol* 6: 306–314
- Rice LB (2008) Federal funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE. *J Infect Dis* 197: 1079–1081
- Richaud F, Richaud C, Ratet P & Pattet J (1986) Chromosomal Location and Nucleotide Sequence of the *Escherichia coli* *dapA* Gene. 166: 297–300
- Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, *et al* (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440: 940–943
- Roberts RJ, Vincze T, Posfai J & Macelis D (2015) REBASE-a database for DNA restriction and modification: Enzymes, genes and genomes. *Nucleic Acids Res* 43: D298–D299
- Rojas-Lopez M, Monterio R, Pizza M, Desvaux M & Rosini R (2018) Intestinal pathogenic *Escherichia coli*: Insights for vaccine development. *Front Microbiol* 9: 1–17
- Ronda C, Chen SP, Cabral V, Yaung SJ & Wang HH (2019) Metagenomic engineering of the mammalian gut microbiome *in situ*. *Nat Methods* 16: 167–170
- Rooks MG & Garrett WS (2016) Gut microbiota, metabolites and host immunity. *Nat Rev Immunol* 16: 341–352
- Ross A, Ward S & Hyman P (2016) More is better: Selecting for broad host range bacteriophages. *Front Microbiol* 7: 1–6
- Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, *et al* (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475: 348–352

- Rozwandowicz M, Brouwer MSM, Fischer J, Wagenaar JA, Gonzalez-Zorn B, Guerra B, Mevius DJ & Hordijk J (2018) Plasmids carrying antimicrobial resistance genes in *Enterobacteriaceae*. *J Antimicrob Chemother* 73: 1121–1137
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB & Erlich HA (1987) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487–491
- Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA & Arnheim N (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230: 1350–1354
- Samuel BS & Gordon JI (2006) A humanized gnotobiotic mouse model of host-archaeal-bacterial mutualism. *Proc Natl Acad Sci* 103: 10011–10016
- Sanger F, Nicklen S & Coulson AR (1977) The role and use of progestogens. *Proc Natl Acad Sci* 74: 5463–5467
- Sawasvirojwong S, Srimanote P, Chatsudthipong V & Muanprasat C (2013) An Adult Mouse Model of *Vibrio cholerae*-induced Diarrhea for Studying Pathogenesis and Potential Therapy of Cholera. *PLoS Negl Trop Dis* 7
- Schaedler RW, Dubois R & Costello R (1965) The development of the bacterial flora in the gastrointestinal tract of mice. *J Exp Med* 122: 59–66
- Schaik W Van (2015) The human gut resistome. *Phil Trans R Soc B* 370: 1–9
- Schloss JA (2008) How to get genomes at one ten-thousandth the cost. *Nat Biotechnol* 26: 1113–1115
- Schoen C, Stritzker J, Goebel W & Pilgrim S (2004) Bacteria as DNA vaccine carriers for genetic immunization. *Int J Med Microbiol* 294: 319–335
- Schofield D, Sharp NJ & Westwater C (2012) Phage-based platforms for the clinical detection of human bacterial pathogens. *Bacteriophage* 2: 105–121
- Shabbir MAB, Hao H, Shabbir MZ, Wu Q, Sattar A & Yuan Z (2016) Bacteria vs. Bacteriophages: Parallel Evolution of Immune Arsenals. *Front Microbiol* 7: 1292
- Sharan SK, Thomason LC, Kuznetsov SG & Court DL (2009) Recombineering: A homologous recombination-based method of genetic engineering. *Nat Protoc* 4: 206–223
- Shepherd ES, DeLoache WC, Pruss KM, Whitaker WR & Sonnenburg JL (2018) An exclusive metabolic niche enables strain engraftment in the gut microbiota. *Nature* 557: 434–438

- Shintani M, Sanchez ZK & Kimbara K (2015) Genomics of microbial plasmids: Classification and identification based on replication and transfer systems and host taxonomy. *Front Microbiol* 6: 1–16
- Sidjabat HE, Chin JJ, Chapman T, Wu K, Ulett GC, Ong CY, Schembri MA, Johnson JR & Trott DJ (2009) Colonisation dynamics and virulence of two clonal groups of multidrug-resistant *Escherichia coli* isolated from dogs. *Microbes Infect* 11: 100–107
- Silley P (2009) Human flora-associated rodents - Does the data support the assumptions? *Microb Biotechnol* 2: 6–14
- Slomovic S, Pardee K & Collins JJ (2015) Synthetic biology devices for *in vitro* and *in vivo* diagnostics. *Proc Natl Acad Sci U S A* 112: 14429–35
- Smillie C, Garcillan-Barcia MP, Francia M V., Rocha EPC & de la Cruz F (2010) Mobility of Plasmids. *Microbiol Mol Biol Rev* 74: 434–452
- Smith K, McCoy KD & Macpherson AJ (2007) Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. *Semin Immunol* 19: 59–69
- Sommer MOA & Dantas G (2011) Antibiotics and the resistant microbiome. *Curr Opin Microbiol* 14: 556–563
- Sonnenborn U & Schulze J (2009) The non-pathogenic *Escherichia coli* strain Nissle 1917 – features of a versatile probiotic. *Microb Ecol Health Dis* 21: 122–158
- Soucy SM, Huang J & Gogarten JP (2015) Horizontal gene transfer: building the web of life. *Nat Rev Genet* 16: 472–482
- St-Pierre F, Cui L, Priest DG, Endy D, Dodd IB & Shearwin KE (2013) One-step cloning and chromosomal integration of DNA. *ACS Synth Biol* 2: 537–541
- Stecher B, Denzler R, Maier L, Bernet F, Sanders MJ, Pickard DJ, Barthel M, Westendorf AM, Krogfelt KA, Walker AW, *et al* (2012) Gut inflammation can boost horizontal gene transfer between pathogenic and commensal *Enterobacteriaceae*. *Proc Natl Acad Sci* 109: 1269–1274
- Steidler L, Neiryneck S, Huyghebaert N, Snoeck V, Vermeire A, Goddeeris B, Cox E, Remon JP & Remaut E (2003) Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10. *Nat Biotechnol* 21: 785–789
- Sun J, Li X, Yang R, Fang L, Huo W, Li S, Jiang P, Liao X & Liu Y (2016) Complete Nucleotide Sequence of an IncI2 Plasmid Coharboring bla CTX-M-55 and mcr-1. *Antimicrob Agents Chemother* 60: 5014–5017
- Suzuki H, Yano H, Brown CJ & Top EM (2010) Predicting plasmid promiscuity based on genomic signature. *J Bacteriol* 192: 6045–6055

- Takeshita T, Kageyama S, Furuta M, Tsuboi H, Takeuchi K, Shibata Y, Shimazaki Y, Akifusa S, Ninomiya T, Kiyohara Y, *et al* (2016) Bacterial diversity in saliva and oral health-related conditions: the Hisayama Study. *Sci Rep* 6: 22164–22175
- Tamminen M, Virta M, Fani R & Fondi M (2012) Large-scale analysis of plasmid relationships through gene-sharing networks. *Mol Biol Evol* 29: 1225–1240
- Tatum EL & Lederberg J (1947) Gene Recombination in the Bacterium *Escherichia coli*. *J Bacteriol* 53: 673–684
- Tolla DA & Savageau MA (2010) Regulation of aerobic-to-anaerobic transitions by the FNR cycle in *Escherichia coli*. *J Mol Biol* 397: 893–905
- Trosset JY & Carbonell P (2015) Synthetic biology for pharmaceutical drug discovery. *Drug Des Devel Ther* 9: 6285–6302
- Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R & Gordon JI (2007) The Human Microbiome Project. *Nature* 449: 804–810
- Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER & Gordon JI (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444: 1027–131
- Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R & Gordon JI (2009) The Effect of Diet on the Human Gut Microbiome: A Metagenomic Analysis in Humanized Gnotobiotic Mice. *Sci Transl Med* 1: 6ra14
- Underhill DM & Iliev ID (2014) The mycobiota: interactions between commensal fungi and the host immune system. *Nat Rev Immunol* 14: 405–416
- Vasu K & Nagaraja V (2013) Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. *Microbiol Mol Biol Rev* 77: 53–72
- Vaudaux P, François P, Haefliger J-A, Entenza J-M, Vaudaux P & Moreillon P (2001) Reassessing the Role of *Staphylococcus aureus* Clumping Factor and Fibronectin-Binding Protein by Expression in *Lactococcus lactis*. *Infect Immun* 69: 6296–6302
- Venter J, Adams M, Myers E, Li P, Mural R, Sutton G, Smith H, Yandell M, Evans C, Holt R, *et al* (2001) The Sequence of the Human Genome. *Science* 291: 1304–1355
- Ventola CL (2015) Metal oxides and hydroxides as rechargeable materials for photocatalysts with oxidative energy storage abilities. *Pharm Ther* 40: 277–283
- Vincent J-L (2003) Nosocomial infections in adult intensive-care units. *Lancet* 361: 2068–2077

- Voyvodic PL, Pandi A, Koch M, Conejero I, Valjent E, Courtet P, Renard E, Faulon JL & Bonnet J (2019) Plug-and-play metabolic transducers expand the chemical detection space of cell-free biosensors. *Nat Commun* 10: 1–8
- Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR & Church GM (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460: 894–898
- Watson J & Crick F (1953) Molecular Structure of Nucleic Acids. *Nature* 171: 737–738
- Weinstock GM (2012) Genomic approaches to studying the human microbiota. *Nature* 489: 250–256
- Wilkins BM (2002) Plasmid promiscuity: Meeting the challenge of DNA immigration control. *Environ Microbiol* 4: 495–500
- Williams SCP (2014) Gnotobiotics. *Proc Natl Acad Sci* 111: 1661–1661
- Wohlleben W, Mast Y, Muth G, Röttgen M, Stegmann E & Weber T (2012) Synthetic Biology of secondary metabolite biosynthesis in actinomycetes: Engineering precursor supply as a way to optimize antibiotic production. *FEBS Lett* 586: 2171–2176
- Wostmann B & Bruckner-Kardoss E (1959) Development of cecal distention in germ-free baby rats. *Am J Physiol* 197: 1345–1346
- Wostmann BS (1981) The germ-free animal in Nutritional Studies. *Annu Rev Nutr* 1: 257–79
- Xie Z, Wroblewska L, Prochazka L, Weiss R & Benenson Y (2011) Multi-Input RNAi-Based Logic Circuit. *Science* 333: 1307–1312
- Yamamoto M & Pop-vicas AE (2014) Treatment for infections with carbapenem-resistant *Enterobacteriaceae*: what options do we still have? *Crit Care* 18: 229–237
- Yaung SJ, Deng L, Li N, Braff JL, Church GM, Bry L, Wang HH & Gerber GK (2015) Improving microbial fitness in the mammalian gut by *in vivo* temporal functional metagenomics. *Mol Syst Biol* 11: 788–804
- Yin ML, De Groof AJC, Bhattacharjee MK, Figurski DH & Schon EA (2008) Bacterial conjugation in the cytoplasm of mouse cells. *Infect Immun* 76: 5110–5119
- Yosef I, Manor M, Kiro R & Qimron U (2015) Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. *Proc Natl Acad Sci* 112: 7267–7272
- Youming Z, Frank B, Joep M & Francis S (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat Genet* 20: 123–128

Yu L, Batara J & Lu B (2016) Application of Genome Editing Technology to MicroRNA Research in Mammals. *Mod Tools Genet Eng*: 163–185

Yu YA, Zhang Q & Szalay AA (2008) Establishment and Characterization of Conditions Required for Tumor Colonization by Intravenously Delivered Bacteria. *Biotechnol Bioeng* 100: 567–578

Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, Schoen C, Vogel J & Sontheimer EJ (2013) Processing-Independent CRISPR RNAs Limit Natural Transformation in *Neisseria meningitidis*. *Mol Cell* 50: 488–503

Zhou S, Gravekamp C, Bermudes D & Liu K (2018) Tumour-targeting bacteria engineered to fight cancer. *Nat Rev Cancer* 18: 727–743

