

Etude génétique de l'adaptation microbienne au moyen de nouvelles méthodes de séquençage

Cécile Gateau

► **To cite this version:**

Cécile Gateau. Etude génétique de l'adaptation microbienne au moyen de nouvelles méthodes de séquençage. Biologie moléculaire. 2012. <hal-01472950>

HAL Id: hal-01472950

<https://hal-ephe.archives-ouvertes.fr/hal-01472950>

Submitted on 21 Feb 2017

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

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

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

ÉCOLE PRATIQUE DES HAUTES ÉTUDES

Sciences de la Vie et de la Terre

MÉMOIRE

présenté

par

Cécile GATEAU

pour l'obtention du diplôme de l'École Pratique des Hautes Études

**Etude génétique de l'adaptation microbienne
au moyen de nouvelles méthodes de
séquençage**

soutenu le 13 décembre 2012

devant le jury suivant :

Frédéric ARNAUD – Président
Olivier TENAILLON – Tuteur scientifique
Michel VEUILLE – Tuteur pédagogique
Michael DUBOW – Rapporteur

Mémoire préparé sous la direction de :

Oliver TENAILLON – Hervé JACQUIER
U722 Ecologie et évolution des micro-organismes
16 rue Henri Huchard 75018 PARIS
olivier.tenaillon@inserm.fr
Directeur : Erick DENAMUR

et de

Michel VEUILLE
UMR 7205 Equipe de génétique des populations
Muséum National d'Histoire naturelle, bat c39
16 rue Buffon 75005 PARIS
mveuille@ephe.sorbonne.fr
Directeur : Michel VEUILLE. EPHE Sciences de la Vie et de la Terre

ÉCOLE PRATIQUE DES HAUTES ÉTUDES
SCIENCES DE LA VIE ET DE LA TERRE

**Etude génétique de l'adaptation microbienne au moyen de nouvelles
méthodes de séquençage**

Cécile GATEAU

Le 13 décembre 2012

Au cours de ces dernières années, l'évolution des micro-organismes est devenue un problème de santé publique majeur. Le changement d'espèce d'hôte (*E. coli* O157H7, SRAS, grippe aviaire, grippe porcine), la réémergence de maladies passées (tuberculose) et l'émergence et la propagation de la résistance aux anti-infectieux (bactéries multirésistantes aux antibiotiques) sont autant d'exemples inquiétants du potentiel adaptatif des micro-organismes. Afin de comprendre leur mode d'adaptation et ses limites, il est nécessaire de développer et de caractériser au mieux la dynamique de ces adaptations et leurs bases moléculaires. Pour avancer dans ce sens, l'évolution expérimentale permet, dans un milieu contrôlé, de faire évoluer des populations bactériennes et de suivre leur adaptation au cours du temps. Grâce aux nouvelles méthodes de séquençage, il est possible de séquencer les génomes et d'identifier les bases moléculaires de ces adaptations. Ainsi au laboratoire, en collaboration avec Richard Lenski (University of Michigan), Dominique Schneider (Université Grenoble) et Claudine Médigue (CEA, Genoscope), nous analysons 6 populations d'*E. coli* B adaptées au laboratoire pendant 40 000 générations. Pour mieux caractériser la dynamique de l'adaptation, les outils de séquençage peuvent être détournés pour estimer des fréquences alléliques. Je me suis donc attachée à développer des outils moléculaires efficaces à grande échelle, permettant d'introduire des marqueurs génétiques au sein du chromosome. Le système d'intégration du bactériophage λ a ainsi été détourné pour permettre l'introduction de tag dans le génome. En suivant la fréquence par séquençage de ces tags, il est possible de déduire la valeur sélective des différents fonds génétiques dans lesquels se sont intégrés les différents tags. Cette méthode permet donc à la fois la réalisation de compétitions avec de multiples génotypes, mais aussi de suivre l'émergence de variant adaptatifs. Cependant, l'extension de méthodes de génie génétique à des échelles ne permettant pas le contrôle individuel de chaque clone s'est révélée être particulièrement laborieuse, et certaines incertitudes demeurent sur le protocole développé.

MOTS-CLÉS: Evolution, Compétition, Diversité, *Escherichia coli*, Séquençage.

INTRODUCTION.....	Erreur ! Signet non défini.
I. Les grands principes de l'évolution.....	8
A) Mutation.....	8
B) Les effets des mutations sur la valeur sélective	9
C) Sélection naturelle	9
D) Dérive génétique.....	10
II. L'évolution des bactéries	11
A) L'exemple de la résistance aux antibiotiques.....	11
B) Choix d'une bactérie modèle de laboratoire	12
a. Résistance et évolution de la virulence chez les bactéries	13
b. <i>E. coli</i> une espèce modèle commensale des vertébrés	13
III. L'évolution Expérimentale	17
A) Quelques exemples de modèles non bactériens	17
B) Les grands avantages de modèles bactériens	18
C) Principes : évolution en chemostat, en tubes, <i>in vivo</i>	19
a. Evolution en chemostat	19
b. Evolution en tubes ou flasques	21
c. Evolution <i>in vivo</i>	21
D) Le succès de Lenski.....	21
a. Une expérience de longue durée.....	21
b. Principaux résultats de Lenski	22
IV. La valeur sélective en évolution expérimentale.....	24
A) Mesure de la valeur sélective par courbe de croissance	24
a. Principe.....	24
b. Avantages/Inconvénients.....	25
B) Mesure de la valeur sélective par compétition et étalement sur boîte de Pétri	25
a. Principe.....	25
b. Avantages/Inconvénients.....	26
C) Mesure de la valeur sélective par compétition au cytomètre de flux	26
a. Principe.....	26
b. Avantages/Inconvénients.....	26
D) Les mesures de compétitions nécessitent des marqueurs	27
a. Le marqueur neutre de sucre	27
b. Les marqueurs croisés	27
c. La fluorescence	27
V. Les mutations avantageuses	28

Liste des abréviations

A: Adénine

ADN: **A**cide **D**ésoxyribo**N**ucléique

Amp: **A**mpicilline

Ara: **A**rabinose

bio: **b**iotine

C : Cytosine

CD1 : Cluster de **D**ifférenciation **1**

CFP: **C**yan **F**luorescent **P**rotein

cm: chlora**m**phenicol

DAP: Acide **D**i**A**mino**P**imélique

ddNTP: **d**idésoxyribonucléotide

DNA: **D**eoxyribo**N**ucleic **A**cid

dNTP: **d**ésoxyribonucléotide

DO: **D**ensité **O**ptique

E. coli: ***Escherichia coli***

ECOR: ***E. coli*** Reference Collection

E. fergusonii: ***Escherichia fergusonii***

G: **G**uanine

gal: **g**alactose

GFP: **G**reen **F**luorescent **P**rotein

g.L⁻¹: **g**ramme par **l**itre

h: **h**eure

HPLC: high performance liquid chromatography

int: **int**égrase

Kan: **K**anamycine

Kb: **K**ilobase

Kv: **K**ilovolt

LB: Luria **B**ertani

mg: **m**illigramme

ml: **m**illilitre

mm: **m**illimètre

MLEE: **M**ulti**L**ocus **E**nzyme **E**lectrophoresis

MLST: **M**ulti**L**ocus **S**equene **T**yping

NaCl: Chlorure de Sodium

nm: **n**anomètre

O/N: **O**ver **N**ight

Ori: **O**rigine

OriT: **O**rigine de **T**ransfert

PCR: **P**olymerase **C**hain **R**eaction

Pfu: *Pyrococcus furiosus*

^R: **R**ésistance

rpm: **r**otation **p**ar **m**inute

S: **S**trong

T: **Thymine**

t: **temps**

TAE: **Tris Acétate EDTA**

Taq: Thermus aquaticus

Tet: **Tetracycline**

tra: **transfert**

U: **Unité**

UFC : **Unité Formant Colonie**

W: **Weak**

μL: **microlitre**

μm: **micromètre**

°C: **Degrés Celsius**

% : **pour cent**

' : **minute**

'' : **seconde**

INTRODUCTION

« Comme il naît beaucoup plus d'individus de chaque espèce qu'il n'en peut survivre, et que, par conséquent, il se produit souvent une lutte pour la vie, il s'ensuit que tout être, s'il varie, même légèrement, d'une manière qui lui est profitable, dans les conditions complexes et quelques fois variables de la vie, aura une meilleure chance pour survivre et ainsi se retrouvera choisi d'une façon naturelle. En raison du principe dominant de l'hérédité, toute variété ainsi choisie aura tendance à se multiplier sous sa forme nouvelle et modifiée » Charles Darwin 1859, « L'origine des espèces ». C'est grâce à l'étude et à l'observation des fruits de sa collecte que Darwin a eu l'idée de sa théorie de l'évolution par sélection naturelle. A la même époque, les expériences de Johann Gregor Mendel ont fondé les bases théoriques de la génétique et de l'hérédité qui sont encore utilisées aujourd'hui.

C'est grâce aux concepts théoriques de l'évolution et de la génétique que la génétique des populations a émergé. Elle a permis sous une multitude de modèles spécifiques d'observer des grandes caractéristiques du monde vivant : adaptation des populations, diversifications, spéciations, conséquences des modes de reproductions, coévolution entre espèces... Cependant, les observations sont la conséquence des processus s'étant déroulés sur des temps évolutifs souvent très long et la validation des modèles reste indirecte et repose sur des concepts scientifiques relativement élaborés. Une des conséquences de cette complexité est la persistance de détracteurs de la théorie de l'évolution, qui utilisent des arguments populistes et simplistes pour réfuter les principes identifiés par Charles Darwin.

Néanmoins, quelques processus adaptatifs peuvent aussi se passer sur des échelles de temps relativement courts et sont observables sur des échelles de temps humaines ou même au laboratoire. C'est sur ce principe que repose l'évolution expérimentale, qui propose de faire évoluer des organismes dans des milieux contrôlés au laboratoire et d'en suivre les caractéristiques. Mon projet a pour but de développer des outils permettant une meilleure caractérisation de l'adaptation se produisant dans ces expériences.

I. Les grands principes de l'évolution

Dans ce chapitre, je vais présenter une description des grandes forces évolutives régissant l'évolution de populations d'organismes isolés qui nous seront utiles par la suite. Avant d'introduire les mutations et la sélection naturelle, nous devons mentionner ici une précision terminologique qui nous aidera à éviter toute confusion pour le reste du manuscrit, la valeur sélective ou « fitness » représente le succès de reproduction moyen d'un génotype particulier dans un environnement donné. Elle est souvent exprimée de façon relative par rapport à un autre génotype, tel que l'ancêtre dans les expérimentations d'évolution. Pour des bactéries, cette valeur sélective est souvent étudiée sur un cycle complet d'évolution qui contient plusieurs phases : la phase de latence, la phase exponentielle et la phase stationnaire. La valeur sélective représente alors un mélange entre la capacité à initier la croissance rapidement, la vitesse de croissance et la survie.

A) Mutation

La mutation correspond à une erreur dans la transmission du matériel génétique d'un individu à son descendant. Biologiquement, elle correspond à la transmission d'une erreur produite dans la molécule d'ADN, support de l'information génétique, suite à une altération chimique ou physique ou due à une erreur de réplication. Malgré la présence de nombreux mécanismes de réparation, il arrive que certaines mutations ne soient pas détectées et soient ainsi transmises d'une génération à une autre. Plusieurs types de modifications peuvent être observées telles que les substitutions, les délétions courtes ou longues, les insertions ou les réarrangements de chromosome. Aussi, des objets parasites, tels que les transposons (séquence d'ADN capable de se déplacer et de se multiplier de manière autonome dans un génome), provoquent des mutations en intégrant de nouvelles copies dans le génome qui peuvent inactiver des gènes ou en modifier la régulation. Ces effets mutagènes sont à la base de l'utilisation des transposons en génétique.

B) Les effets des mutations sur la valeur sélective

Une fois que la mutation est apparue dans un individu, son destin évolutif au sein de la population est ainsi principalement déterminé par son effet sur la valeur sélective appelé alors l'effet sélectif. La connaissance de l'effet des mutations sur la valeur sélective est l'un des indices fondamentaux des génétiques évolutives. Les mutations peuvent être de trois types : délétères, neutres ou avantageuses.

- Une mutation est délétère lorsqu'elle a un effet néfaste sur la valeur sélective de l'organisme. La plupart des mutations sont délétères, ce fait est attendu puisque, grâce à la sélection naturelle, les organismes sont *a priori* bien adaptés à leur environnement et la probabilité qu'une mutation aléatoire améliore des phénotypes doit être très faible. La distribution des effets néfastes est vaste, allant de l'effet complètement mortel ou létal à l'effet faiblement nuisible.

- Une mutation est neutre lorsqu'elle n'a aucun effet sur la valeur sélective de l'organisme. Une mutation synonyme a peu d'effet et peut alors être perçue comme une mutation neutre.

- Enfin, une mutation est avantageuse lorsqu'elle a un effet bénéfique sur la valeur sélective de l'organisme. Puisqu'elles sont très rares, les mutations avantageuses sont difficiles à détecter et la distribution de leur effet est encore difficile à évaluer (Eyre-Walker et *al.*, 2007). L'évolution adaptative est le procédé par lequel une population améliore sa valeur sélective moyenne, pour des populations bactériennes issues d'un clone. Ce processus requiert l'apparition et la sélection de mutations avantageuses.

C) Sélection naturelle

La sélection naturelle est une force évolutive qui permet aux populations de s'adapter. C'est par l'action de la sélection naturelle que les organismes au sein d'une population ayant la valeur sélective la plus forte laisseront plus de descendants que ceux de valeur sélective moindre. On peut quantifier la différence de valeur sélective de 2 génotypes en suivant l'évolution de leur fréquence au sein d'une population (Fisher, 1930).

Dans le cas d'organismes haploïdes, notons W_1 et W_2 la valeur sélective de 2 génotypes **A** et **a**, de fréquence p et $q=1-p$. Notons W la valeur sélective moyenne de la population soit $W=pW_1+qW_2$

Après une génération, nous avons :

$$p' = \frac{W_1 p}{W} \text{ et } q' = \frac{W_2 q}{W} \text{ et donc } \frac{p'}{q'} = \frac{p}{q} \left(\frac{W_1}{W_2}\right)$$

on peut donc écrire $\frac{p(t)}{q(t)} = \frac{p(0)}{q(0)} \left(\frac{W_1}{W_2}\right)^t$

On voit donc que si $W_1 > W_2$, le génotype **A** envahit la population d'autant plus vite que W_1/W_2 est grand. Inversement si $W_2 > W_1$, l'allèle **A** est éliminé de la population d'autant plus vite que W_1/W_2 est petit.

D) Dérive génétique

Dans une population de taille infinie, comme illustré dans le modèle précédent, il n'existe pas de fluctuations aléatoires dans les changements de fréquences alléliques d'une génération à une autre : un allèle procurant un avantage même minime est donc irrémédiablement sélectionné pour devenir l'allèle majoritaire de la population, un allèle délétère purgé et un allèle neutre garde sa fréquence fixe dans la population. Dans une population de taille finie, les fluctuations aléatoires, qui apparaissent due à l'échantillonnage des allèles fondateurs de la nouvelle génération, modifient ces résultats. Ces fluctuations sont d'autant plus importantes que la taille de la population est faible et correspondent au phénomène de dérive génétique (Kimura et *al.*, 1969).

Ce terme peut être illustré dans le cas de 2 allèles de même valeur sélective. Dans ce cas précis, en population infinie, la fréquence p de l'allèle **a** doit rester constante. Cependant, si la population a une taille finie de N individus, la composition génétique de N individus de la génération suivante correspond à un échantillonnage. Ainsi, en moyenne, la fréquence de **a** ne change pas mais elle va légèrement fluctuer avec une variance $p(1-p)/N$ qui est d'autant plus importante que N est petit. Par conséquent, au lieu d'être stable la fréquence de **a** va dériver et amener irrémédiablement l'allèle **a** à une fréquence de 0 (perte de l'allèle) ou une fréquence de 1 (fixation de l'allèle). L'existence de la dérive génétique permet la fixation d'une mutation neutre. Mais l'action de la dérive affecte aussi la fixation de mutations

délétères ou avantageuses. Si les effets sélectifs sont faibles alors les fluctuations liées au hasard peuvent être plus importantes au point de rendre négligeable l'action de sélection.

Une étude quantitative de ces forces évolutives suggère que l'adaptation des organismes sera efficace et rapide si le temps de génération est court, les pressions de sélection sont fortes et les tailles de population importantes permettant à la fois l'apparition et la sélection de mutants (Sniegowski et *al.*, 1997).

II. L'évolution des bactéries

Les micro-organismes avec leurs fortes tailles de population et leurs temps de génération courts semblent donc être des organismes modèles pour étudier l'évolution sur des échelles de temps courts (Santiago et *al.*, 2003).

A) L'exemple de la résistance aux antibiotiques

L'exemple type d'une évolution rapide sur un temps court, observable sur le temps de vie d'un Homme est celle de la résistance aux antibiotiques. Le phénomène de résistance aux antibiotiques n'est pas du tout récent, il existait bien avant la découverte des antibiotiques par l'Homme. Certaines espèces bactériennes ont développé cette « résistance naturelle » dans leur propre milieu il y a de cela des millions d'années, soit parce qu'elles étaient elles-mêmes productrices de molécules antibiotiques et devaient donc les tolérer, soit pour contrer l'effet d'antibiotiques produits par d'autres organismes présents dans leur entourage (Fischer et *al.*, 2005). La résistance aux antibiotiques est donc une conséquence de la compétition entre les micro-organismes pour occuper une même niche écologique. Toujours est-il que l'Homme n'a que récemment commencé à diriger et accélérer cette résistance de façon importante, par une utilisation parfois inadéquate et abusive d'antibiotiques en santé humaine et vétérinaire ainsi que dans l'industrie agroalimentaire. L'ampleur du problème est telle que dès la sortie d'un nouvel antibiotique, ses jours sont comptés. Il ne s'agit généralement que d'une question de mois ou d'années avant qu'une résistance considérable ne se développe. (Davies, 1994) On ne se demande même plus si une résistance surviendra, mais plutôt à quel moment surviendra-t-elle (Davies, 1994 ; Levy, 1998). Lorsque la résistance n'est pas naturelle, elle peut être mutationnelle ou acquise par un élément génétique mobile, comme un plasmide ou un

transposon. Dans le cas de la résistance mutationnelle (ou chromosomique), la mutation subie par le gène cible de l'antibiotique du chromosome de la bactérie se traduit par la synthèse d'une protéine modifiée et le gène de résistance est transmis à la descendance de cette bactérie nouvellement résistante. Cet évènement est assez fréquent en raison du très grand nombre de bactéries impliquées dans un cycle d'infection, de leur temps de génération très rapide et du taux intrinsèque de mutation qui est d'environ 1 sur 10^{10} , ce qui peut paraître faible, mais si une seule de ces mutations confère la résistance, la bactérie résistante survivra et deviendra dominante (Walsh,2000). Dans le cas de la résistance reliée à l'acquisition d'un plasmide (ou d'un transposon) portant un gène de résistance, non seulement la bactérie transmettra ce gène à sa descendance, mais il lui sera également possible d'échanger cet élément mobile avec d'autres bactéries présentes dans son environnement par conjugaison ou transduction. Ce type de résistance est le plus fréquent et est de plus en plus répandu. Il concerne tous les types d'antibiotiques et peut même se transmettre entre bactéries d'espèces différentes.

Enfin, il est également possible qu'une espèce bactérienne acquiert plusieurs plasmides portant chacun un gène de résistance à un antibiotique différent. Lorsque ces éléments fusionnent, la totalité des gènes de résistance est transmise à une autre bactérie et des plasmides porteurs de résistances multiples sont ainsi transférés à de nombreuses populations (Encyclopédie Encarta, 2007).

Cette résistance est le fruit de la remarquable adaptabilité des bactéries, qui ont su développer différents mécanismes afin de contrer les effets nocifs des antibiotiques (Walsh,2000 ; Fischer et *al.*, 2005 ; Nikaido,1996).

B) Choix d'une bactérie modèle de laboratoire

L'évolution rapide des bactéries ne se cantonne pas à l'émergence de résistance. Leurs conflits avec le système immunitaire de l'hôte, leurs bactériophages et les environnements fluctuants qu'ils rencontrent sont aussi des sources d'évolution rapide. A ce titre, en tant qu'organisme modèle en génétique et microbiologie, *E. coli* se révèle être un modèle évolutif intéressant.

a. Résistance et évolution de la virulence chez les bactéries

Si la plupart des souches de *E. coli* se comportent comme commensal du tube digestif (Tenailon et *al.*, 2010), certaines peuvent être impliquées dans des infections intestinales ou extra-intestinales. Les bactériémies à *E. coli* font partie des formes les plus sévères des infections extra-intestinales, et sont relativement fréquentes. Par ailleurs, certaines souches d'*E. coli* peuvent entraîner des complications graves voire mortelles, notamment celles responsables de diarrhées hémorragiques. Récemment, plusieurs épidémies dues à *E. coli* sont apparues en France et en Allemagne. Les souches responsables étaient pourtant connues et caractérisées comme non pathogènes dans l'intestin des ruminants. En Allemagne et dans la région de Bordeaux, la souche mise en cause est *E. coli* O104:H4 : souche rare n'ayant jamais été à l'origine d'épidémie ([<http://www.frm.org/actualite/toute-l-actu/maladies-infectieuses-escherichia-coli-une-bacterie-dangereuse.html>]) (Rosko et *al.*, 2011).

Longtemps considérée comme une entérobactérie sensible aux antibiotiques, *E. coli* a supplanté les *Klebsiella sp.* et *Enterobacter sp.* dans la résistance aux céphalosporines de 3ème génération ([<http://www.rivm.nl/earss/database/>]) tandis que sa résistance aux fluoroquinolones atteint également des taux préoccupants, ces bactéries se retrouveraient par portage intestinal et ce à des taux importants dans certains pays (Kader et *al.*, 2007 ; Nurul-Atifah et *al.*, 2005 ; Pallecchi et *al.*, 2007).

La compréhension de cette complexité ne peut relever de la seule approche phénotypique conventionnelle mais doit résulter d'une étude de la structure génétique de l'espèce et de ses rapports avec la virulence et la résistance aux antibiotiques ainsi que son évolution.

Dans ce sens, l'étude de l'évolution d'*E. coli* est importante. Pour mieux comprendre cette évolution, une approche réductionniste, qui consiste à faire évoluer la bactérie au laboratoire a été largement utilisée.

b. *E. coli* une espèce modèle commensale des vertébrés

E. coli est une espèce bactérienne paradoxale. Elle est en effet à la fois le commensal aérobie le plus abondant du tube digestif de nombreux vertébrés à sang chaud, tels que les

mammifères, les oiseaux, mais aussi des vertébrés à sang froid et une bactérie retrouvée dans l'environnement. Par ailleurs, c'est le pathogène le plus fréquemment impliqué au cours d'infections intestinales et extra-intestinales en médecine humaine et vétérinaire (Tenaillon et *al.*, 2010). Cependant, *E. coli* est aussi un pathogène versatile pouvant être un pathogène opportuniste, alors que certains pathovars sont des pathogènes stricts. Il est classiquement admis que *E. coli* peut représenter un simple commensal, mais en réalité la bactérie présente de véritables caractéristiques mutualistes avec l'hôte qui l'héberge (Nowrouzian et *al.*, 2003). Elle permet en effet l'apport de vitamines (Bentley et *al.*, 1982), la stimulation du système immunitaire ou encore constitue une barrière contre l'implantation de souches pathogènes (Schamberger et *al.*, 2004 ; Vollard et *al.*, 1994).

Dans le but d'étudier les raisons de ce polymorphisme écologique et afin de tenter de différencier les souches pathogènes des souches commensales, plusieurs travaux ont été réalisés sur le génome de l'espèce *E. coli*, permettant de mieux comprendre les déterminants de l'expression de la virulence.

La collection ECOR : collection représentative de l'espèce

Afin de représenter la diversité de l'espèce *E. coli*, plusieurs collections ont été rassemblées. La collection ECOR (« *E. coli* reference collection ») est la plus étudiée. Elle a été établie à partir de 2600 isolats naturels par Ochman et Selander en 1984 et est composée de 72 souches provenant de zones géographiques différentes (USA, Suède, Bali, Indonésie, Nouvelle-Guinée, Iles Tonga). Parmi ces souches, 61 sont issues de la flore fécale de sujets sains (humains ou espèces animales) et 11 souches isolées au cours de cystites ou de pyélonéphrites aiguës chez l'Homme (Ochman et *al.*, 1984). Les souches ont été sélectionnées d'après leur variabilité électrophorétique pour 11 enzymes métaboliques, déterminées par électrophorèse en gel d'amidon lors d'étude de *multilocus enzyme electrophoresis* (MLEE), et pour leur diversité d'hôtes (diversité d'espèces et diversité d'origine géographique). Des travaux ont permis de générer d'importantes données génétiques et phénotypiques et conforter l'idée que cette collection est représentative de la diversité génétique et phénotypique de l'espèce.

b.1. Une espèce génétiquement structurée

Les premières données d'analyse de la phylogénie de l'espèce *E. coli* ont été effectuées par MLEE (Whittam et al., 1983) faisant intervenir 38 enzymes. Dans un second temps, les analyses ont été réalisées sur des données de MLST (Multi Locus Sequence Typing) (Escobar et al., 2004) des 72 souches de la collection ECOR. Ainsi, 4 groupes phylogénétiques principaux ont été déterminés (A, B1, B2 et D) et 2 groupes accessoires (C et E). Par la suite, le groupe D a pu être séparé en 2 groupes, un groupe D et un groupe F, phylogénétiquement proche des B2. Depuis peu, un arbre phylogénétique robuste réalisé sur les données les plus complètes disponibles a été obtenu à partir de 1878 gènes du « core » génome. (Touchon et al., 2009).

Les résultats des analyses phylogénétiques de l'espèce *E. coli* laissent supposer que le groupe le plus ancien, apparu il y a 40 millions d'années serait le groupe B2. Le groupe D se serait ensuite distingué puis seraient apparu les groupes A et B1 (Lecointre et al., 1998). Les souches de *Shigella* et entéro-hémorragiques O157:H7 sont regroupées coupant l'arbre en 2 avec d'une part les groupes B2 et D et d'autre part les groupes A et B1 (Escobar et al., 2004 ; Paramo et al., 2003).

Une étude plus approfondie de la diversité génétique a pu mettre en évidence une grande diversité au sein du groupe B2 que l'on peut diviser en 9 sous-groupes (Le Gall et al., 2007). Parmi les 9 sous-groupes définis au sein du groupe B2 (Le Gall et al., 2007), les souches responsables d'infections urinaires sont principalement retrouvées dans les sous-groupes II et IX (Jaureguy et al., 2008).

On voit donc à travers ces analyses que les différents groupes et sous-groupes phylogénétiques sont associés à différentes pathologies. De façon similaire, on trouve que la prévalence de ces groupes est très variable en fonction des espèces hôtes et des styles de vie des hôtes. Ainsi les B2 sont majoritaires dans les selles des populations humaines occidentalisées, mais représentent de l'ordre de 5% des souches des populations humaines de pays tropicaux (Tenailon et al., 2010).

La notion de « core » génome et de « pan » génome

Au sein de l'espèce *E. coli*, la taille du génome varie entre 4068 gènes et 5379 gènes ce qui représente une variation de plus de 1000 gènes entre 2 souches. Cette différence correspond à environ 500 protéines (Touchon et al., 2009). Lors de la comparaison des 2 premiers génomes séquencés de *E. coli* (*E. coli* K-12 et *E. coli* O157 :H7), il a été montré que seulement 30% des gènes sont communs aux 2 souches (Perna et al., 2001). 18 autres génomes de souche ont été séquencés par la suite et leur analyse comparative a permis de confirmer et de préciser ces données : L'ensemble des génomes de *E. coli* possèdent en moyenne 4721 gènes dont 1976 sont communs chez toutes les souches (Touchon et al., 2009). Ces gènes communs composent le « core » génome (Grimont, 1988 ; Wayne et al., 1987). Les autres gènes sont variables d'une souche à l'autre. L'ensemble des gènes communs et des gènes variables est appelé le « pan » génome (Touchon et al., 2009).

b.2. Quel déterminisme génétique pour les différents groupes et sous-groupes ?

La souche de référence est *E. coli* K-12 et c'est elle qui a servi de modèle pour l'accumulation des connaissances en génétique, physiologie, biologie moléculaire et biochimie. Il s'agit également du premier *E. coli* qui a été complètement séquencé en 1997 (Blattner et al., 1997). Historiquement *E. coli* K-12 a été isolé de fèces d'un patient atteint d'une diphtérie en 1922 à Palo Alto (USA) (Hobman et al., 2007). Cette souche a été conservée plusieurs années à l'université de Stanford avant d'être adoptée comme souche de référence par Charles Clifton puis Edward Tatum dans les années 1940. Cette souche a subi des modifications sous l'effet des ultraviolets et d'une croissance dans différents milieux (gélose au sang ou en présence d'acridine orange). Plusieurs souches dérivent de cette souche de référence dont *E. coli* K-12 MG1655 qui est étudiée dans notre laboratoire.

E. coli est habituellement étudiée en milieu riche, milieu LB (Luria Bertani) ou en milieu minimum glucose. Ces conditions expérimentales sont très éloignées des habitats naturels d'*E. coli* que sont le tube digestif, l'urine ou des environnements hydriques comme nous l'avons vu précédemment. Le milieu de vie étant susceptible d'influencer l'expression

des gènes de virulence, il est nécessaire que les souches soient étudiées dans des conditions mimant au mieux leur « style de vie » (Dougan, 2009).

Aussi au cours des dernières années, de nombreuses études se sont intéressées à *E. coli*, non seulement au travers de sa diversité génétique mais également au travers des environnements plus complexes. De gros progrès ont été faits pour une meilleure compréhension des bases génétiques de la virulence, avec l'identification de facteurs de virulence. Cependant, les facteurs impliqués dans le commensalisme restent bien moins définis. En fait, il semblerait que la plupart des gènes de virulence aient essentiellement un rôle dans la colonisation, et que leur impact dans la virulence (surtout extra-intestinale) agisse comme un dégât collatéral (Le Gall *et al.*, 2007). Pour essayer d'aller plus loin, l'analyse de génome complet est apparue comme prometteuse pour identifier des gènes spécifiques des groupes et sous-groupes et leur répartition en tant que commensale. Cependant, très peu de gènes ont finalement été identifiés et donc les déterminants de l'adaptation différentielle au tube digestif restent très largement méconnus.

III. L'évolution Expérimentale

L'évolution expérimentale est une approche puissante pour comprendre l'évolution des bactéries, cependant elle ne se cantonne pas à celles-ci, d'autres organismes ayant été utilisés dans cette discipline (Allen Orr, 2005).

A) Quelques exemples de modèles non bactériens

Depuis des siècles, l'Homme a choisi, isolé et favorisé pour son usage personnel quelques organismes parce qu'il trouvait en eux des qualités lui convenant (exemple de choix de plantes pour sa production élevée de graines ou d'animaux pour sa quantité de viandes). Désireux d'obtenir d'avantage, il a pratiqué la sélection dont il a découvert empiriquement l'efficacité. La domestication si l'on s'en réfère à la doctrine Darwinienne se présente comme une évolution accélérée et dirigée par l'Homme. Accélérée parce qu'on ne conserve d'emblée que les mutants jugés profitables se trouvant ainsi avantagés par rapport au type sauvage. Ils correspondent aux porteurs du génotype privilégié par la sélection naturelle.

Les premiers généticiens disposaient de variétés naturelles, c'est-à-dire des mutants établis par la nature ou des mutants apparus spontanément dans leur élevage. Puis, au cours du temps, des protocoles de mutagenèse et de sélection de mutants ont été mis au point par les biologistes permettant une amélioration considérable de l'évolution expérimentale. De nombreux organismes ont été utilisés pour les expériences d'évolution expérimentale et certains sont encore utilisés aujourd'hui :

- La drosophile est un petit insecte (eucaryote et pluricellulaire) facilement manipulable. Elle se reproduit facilement et elle est pratique à manipuler génétiquement puisqu'elle possède un petit génome (170 mégabases) (Molly et *al.*, 2009). C'est l'un des premiers organismes vivants à avoir été utilisé pour l'évolution expérimentale (1920).

- La levure est un micro-organisme eucaryote et unicellulaire facilement manipulable au laboratoire. Son génome fait 13 mégabases. C'est l'organisme eucaryote le plus simple qui existe. L'analyse génétique chez la levure présente une particularité car elle maintient 4 cellules issues de la méiose au sein d'une poche.

- Enfin des virus et bactériophages de virus tels que *VSV* (Virus Stomatite Vésiculaire), *Phi6* ou encore *T7*, et *phiX174*, *Phi2* ont été régulièrement utilisés pour comprendre la dynamique d'adaptation des virus.

B) Les grands avantages de modèles bactériens

Un des organismes les plus faciles à manipuler en laboratoire est la bactérie. Les bactéries sont des organismes procaryotes. Leurs génomes font de 1 à 5 mégabases. Ce sont les organismes les plus simples à maintenir et à utiliser. Les premières utilisations de micro-organismes en évolution expérimentale ont été réalisées dans les années 1940 et 1950, et ont connu un essor avec les travaux de Richard Lenski. (Santiago et *al.*, 2003 ; Barrick et *al.*, 2009).

L'évolution expérimentale basée sur les micro-organismes a de nombreux avantages et a permis de tester et d'élaborer des théories en évolution. Ces études, encore d'actualité, se sont améliorées au fil du temps et la maîtrise des micro-organismes a permis de simplifier les expérimentations.

Les micro-organismes sont adaptés pour les expériences d'évolution pour beaucoup de raisons pratiques (Barrick et *al.*, 2009):

- Ils sont faciles à développer et à énumérer.
- Ils se reproduisent rapidement, ce qui permet de faire des expériences avec de nombreuses générations.
- Ils permettent d'avoir une large population dans un petit espace, ce qui facilite les expériences de réplifications.
- Ils peuvent être stockés dans une suspension à -80°C et sont ultérieurement « revivifiables », ce qui permet une comparaison directe entre l'ancêtre et les formes évoluées.
- Ils se reproduisent de façon asexuée et le clonage résultant améliore la précision de l'expérience de réplification.
- L'asexualité conserve également une liaison entre le marqueur génétique et génomique, ce qui facilite les mesures de valeur sélective.
- On peut faire varier facilement leur environnement.
- Les nombreuses connaissances accumulées sur *E. coli* au niveau moléculaire permettent à l'issue du séquençage des souches évoluées, d'identifier les déterminants moléculaires de l'adaptation.

C) Principes : évolution en chemostat, en tubes, *in vivo*

a. Evolution en chemostat

L'évolution en chemostat est un dispositif qui permet une croissance continue d'une population bactérienne avec un taux de croissance limité par les ressources. Les ressources arrivent avec un taux constant dans le chemostat. La population s'accroît et consomme les ressources jusqu'à ce que les bactéries atteignent un équilibre de densité où le taux de croissance est égal au taux d'écoulement de ressources à travers le récipient de chemostat. Dans les cultures en chemostat, les bactéries font face à une perpétuelle limitation de ressource. Les systèmes de transport pour les ressources limitées sont les cibles probables de la sélection, et les gènes qui encodent ces systèmes sont ainsi des locus désignés.

Cette méthode n'est pas la plus facile qui existe en évolution expérimentale car elle demande beaucoup d'entretien et les risques de contaminations sont élevés. De plus, les

replicats sont limités car l'installation est assez conséquente et la durée de croissance est courte car les bactéries adhèrent sur les parois. Cependant, elle reproduit quelques points du milieu naturel de la bactérie : la perpétuelle limitation de ressource.

b. Evolution en tubes ou flasques

L'évolution en tube se réalise par transfert en série. Un régime de culture dans lequel une proportion de la population est périodiquement diluée dans du milieu frais dans lequel la population croît jusqu'à ce qu'elle épuise les ressources limitées jusqu'au prochain transfert du cycle. C'est une méthode très facile à mettre en œuvre, de nombreux réplicats sont possibles mais ne reproduisent pas du tout le milieu naturel de *E. coli*. Elle permet d'obtenir des données globales (utilisation du milieu de culture...) mais pas sur le comportement de la bactérie dans un organisme.

c. Evolution *in vivo*

L'évolution *in vivo* permet de suivre une population bactérienne dans un organisme. On peut, par exemple, implanter une population bactérienne dans un tube digestif d'un animal. Au préalable, toutes les bactéries d'espèces similaires présentes dans le tube digestif pouvant fausser les résultats sont éliminées par ajout d'un antibiotique. L'animal est ensuite gavé par une population bactérienne, puis les fécès sont prélevés régulièrement afin de suivre l'évolution de la population. Cette méthode est plus difficile à mettre en œuvre car elle nécessite l'utilisation d'organismes vivants. Cependant, elle reproduit presque parfaitement le milieu naturel de *E. coli* et les données générées sont plus proches de la réalité.

D) Le succès de Lenski

De très nombreuses études d'évolution expérimentale ont utilisé *E. coli* mais la plus célèbre reste celle initiée par Lenski (Lenski et *al.*, 1994).

a. Une expérience de longue durée

L'expérience a débuté en 1988 (Lenski et *al.*, 1991), Richard Lenski a sélectionné un clone d'*E. coli* et aensemencé 12 flasques remplies d'un milieu nutritif de même

composition, très pauvre en glucose (étant aussi le facteur limitant) par des colonies presque identiques. Les 12 populations diffèrent par leur activité sur l'arabinose. On peut distinguer les Ara+ (qui peuvent utiliser l'arabinose) des Ara-. Avec un mélange d'arabinose et d'un colorant chimique, le tétrazolium, les bactéries Ara+ se voient sous forme de colonies blanches alors que les bactéries Ara- se voient sous forme de colonies rouges. Au préalable, Lenski a vérifié que ce marqueur n'avait aucun effet sur la valeur sélective des bactéries. Ces populations de bactéries ont été repiquées de façon à ce que 6 possèdent l'allèle Ara+ et les 6 autres Ara-. Chaque jour, avec chacune des 12 populations, un nouveau tube de milieu était ensemencé avec la culture de la veille puis un échantillon de celle-ci est prélevé et congelé, permettant une analyse ultérieure de l'échantillon. L'expérience dure depuis près de 25 ans et si l'on considère une moyenne de 6 à 7 générations de bactéries par jour, cela donne environ 55 000 générations de bactéries à l'heure actuelle. Pendant le même laps de temps, l'espèce *Homo Sapiens* n'en est qu'à sa 7500^{ème} génération ! La question intéressante de cette expérience serait de savoir si les nouvelles générations de bactéries resteraient identiques à leurs ancêtres, ou bien si elles évolueraient... et si elles évolueraient, si les 12 colonies évolueraient de la même façon ou selon des schémas différents.

b. Principaux résultats de Lenski

b.1. Dynamique de l'adaptation

Au cours des 20 dernières années, l'expérience initiée par Richard Lenski a servi de référence pour étudier la dynamique de l'adaptation. Trois grandes tendances ont pu émerger de cette expérience :

- Une vitesse d'adaptation décroissante. Alors que les premières 2000 générations d'adaptation ont été caractérisées par un fort gain de valeur sélective, ce gain décroît par la suite. Cette décélération de l'adaptation a été observée dans toutes les expériences d'évolution expérimentale (Lenski et *al.*, 1994).

- Un fort parallélisme dans les gains en valeur sélective. Les 12 réplicats évoluent de façon globalement similaire et semblent évoluer vers une même valeur sélective finale: on ne trouve pas une population bloquée à une valeur significativement différente des autres (Lenski et *al.*, 1994) .

- Au niveau moléculaire, des régulateurs globaux semblent être impliqués dans les premiers pas de l'adaptation. Ainsi, il semble que ce soit au travers de la régulation, en *trans* plus qu'en *cis*, que la modulation de la régulation se produit (Hindré et *al.*, 2012).

b.2. Plusieurs mutations

En 2008, Lenski et ses collaborateurs ont fait état d'une adaptation particulièrement importante qui s'est produite dans l'une des 12 populations (Blount et *al.*, 2008). Ils ont remarqué qu'à la génération 33127, une des populations a pu considérablement augmenter la taille de sa population par rapport aux autres. La population a acquis l'aptitude d'absorber le citrate aussi bien que le glucose. La quantité de nourriture disponible dans chaque tube a donc considérablement augmenté et lui a permis de se développer d'avantage. Lenski a émis l'hypothèse que ceci était dû à au moins 2 mutations (30 000 générations était suffisantes pour que chaque gène ait muté au moins une fois dans chacune des 12 lignées, et donc que si cela était dû à une seule mutation, il est probable que d'autres populations auraient acquis la même capacité) (Woods et *al.*, 2001). A un moment de l'évolution, une bactérie de cette population aurait subi une mutation n'ayant aucun effet, c'est donc l'accumulation de cette mutation avec une nouvelle ayant apparue qu'une bactérie a acquis la capacité à utiliser le citrate lui permettant d'envahir la population.

Ce travail effectué par Lenski et son équipe montre un grand nombre des composants essentiels de l'évolution: les mutations aléatoires soumises à une sélection naturelle, l'adaptation à un environnement par des voies différentes, la façon dont des mutations successives peuvent s'additionner pour opérer une modification évolutive, ou encore que certains gènes ont besoin d'autres gènes pour s'exprimer.

Au cœur de toutes ces expériences, la valeur sélective est le point central. Il est donc important de savoir mesurer une valeur sélective en évolution expérimentale. Cependant comme nous allons le voir, les différentes approches existantes sont laborieuses et possèdent des limitations. L'objectif de mon travail a été de développer une méthode alternative efficace.

IV. La valeur sélective en évolution expérimentale

La notion de croissance bactérienne recouvre 2 aspects : la croissance de la cellule bactérienne (taille, masse, volume) et le phénomène de multiplication cellulaire (population). Pour simplifier, on assimile souvent la croissance à la multiplication cellulaire. Le plus simple est de considérer la croissance comme un ensemble de réactions du métabolisme conduisant à la synthèse de biomasse bactérienne. La croissance est alors définie par l'augmentation de biomasse sèche.

Il existe plusieurs méthodes pour mesurer la valeur sélective. Chacune a ses avantages et inconvénients. Mon projet a pour but de développer une nouvelle approche pour mesurer la valeur sélective.

A) Mesure de la valeur sélective par courbe de croissance

a. Principe

La façon la plus simple d'étudier la croissance bactérienne est de mesurer le nombre de bactéries à intervalles de temps réguliers dans une culture. Pour cela, un prélèvement du milieu de culture contenant les bactéries est réalisé, puis la densité optique est mesurée. Cette technique permet d'évaluer la concentration cellulaire en utilisant un spectrophotomètre à 650 nm. Plus il y a de micro-organismes, plus la lumière est réfléchi, plus l'intensité du faisceau restant est faible et plus la valeur d'absorbance est grande. Les données de densité optique (DO) de la culture bactérienne en fonction du temps (t) permettent de réaliser une courbe de croissance : $DO=f(t)$. Afin de permettre une exploitation plus précise de la courbe de croissance, la courbe $\log DO=f(t)$ peut être tracée car celle-ci est plus linéaire. Cette courbe distingue 6 phases:

- La phase I est la phase de latence, elle correspond au temps d'adaptation des bactéries à leur milieu de culture.
- La phase II est la phase d'accélération, c'est le début de la multiplication bactérienne.
- La phase III est la phase exponentielle, elle correspond à la multiplication des bactéries à leurs vitesses maximales.

- La phase IV est la phase de décélération, le rythme de multiplication des bactéries diminue.
- La phase V est la phase stationnaire où il y a autant de bactéries qui meurent que de bactéries qui se multiplient. Le milieu s'appauvrit en nutriments.
- La phase VI est la phase de déclin, il n'y a plus de nutriments dans le milieu de culture, les cellules meurent.

b. Avantages/Inconvénients

Cette méthode a cependant plusieurs inconvénients:

- Il est difficile d'en mesurer de faibles variations du fait de fortes fluctuations conférant une sensibilité réduite à ces techniques.
- Elles ne mesurent qu'une des composantes de la valeur sélective (biomasse) et ne fournissent qu'une information partielle.

Cette méthode nécessite de nombreux réplicats car elle n'est pas précise. De plus, en plaque, l'oxygénation des cultures n'est pas homogène pour tous les puits et un « effet de puits » est observé, il faut donc le prendre en compte lors de l'analyse des résultats. Cependant, c'est la méthode la plus simple et a l'avantage de pouvoir être réalisée en plaque 96 puits.

B) Mesure de la valeur sélective par compétition et étalement sur boîte de Pétri

a. Principe

Pour parer à ces limitations, la valeur sélective est classiquement mesurée par compétition. Deux souches bactériennes sont placées dans les conditions physiologiques de l'expérience avant d'être mélangées et de subir un cycle d'évolution. Les proportions du mélange avant et après ce cycle sont mesurées par étalement sur un milieu permettant de séparer les 2 génotypes, et à partir du changement de fréquence une valeur sélective est déterminée.

b. Avantages/Inconvénients

Si cette méthode mesure une valeur sélective réelle, elle n'est cependant pas sans problème. Tout d'abord les compétitions sont extrêmement fastidieuses : chaque compétition doit être répétée plusieurs fois et nécessite plusieurs dilutions et étalements sur milieux sélectifs permettant le comptage de chaque population. Ainsi, il est possible de travailler sur des dizaines de souches mais travailler sur des centaines représente un très gros effort. Enfin, idéalement, les compétitions devraient être réalisées sur plusieurs jours afin de s'affranchir des conditions initiales, propres à chacun des génotypes étant initialement en culture isolée.

C) Mesure de la valeur sélective par compétition au cytomètre de flux

a. Principe

Cette méthode consiste à réaliser des cycles de compétition entre une bactérie dont on veut évaluer la valeur sélective et une bactérie « compétitrice » marquée par un marqueur fluorescent. La proportion des 2 populations peut alors être distinguée par passage en cytométrie de flux. Toute la difficulté de la méthode repose dans le choix du marqueur de fluorescence : son intensité, son support (plasmide ou chromosome), son coût physiologique et sa stabilité et enfin la discrimination qu'il permet.

Lorsque 2 bactéries sont en compétition, la discrimination entre les cellules fluorescentes et non fluorescentes n'est souvent pas assez puissante, entraînant un chevauchement des pics au cytomètre où la discrimination est suffisante mais nécessite un tel coût physiologique pour les cellules l'exprimant que la construction génétique n'est pas stable.

b. Avantages/Inconvénients

Cette méthode requiert la présence d'un compétiteur fortement modifié et exprimant une quantité massive de protéines fluorescentes pour permettre son identification, ce qui peut avoir un coût pour la cellule. Cependant lors des compétitions, toutes les cellules possèdent ce marqueur ce qui n'influence pas les résultats de compétitions, étant toutes dans les mêmes

conditions. Elle se limite aussi à la compétition de 2 souches uniquement. Cette méthode permet un résultat rapide et direct.

D) Les mesures de compétitions nécessitent des marqueurs

a. Le marqueur neutre de sucre

C'est le marqueur qu'utilise Lenski dans ses expériences de longue durée. Les populations diffèrent par leur activité sur l'arabinose, on peut distinguer les Ara+ (pouvant utiliser l'arabinose) des Ara-. Avec un mélange d'arabinose et d'un colorant chimique, le tétrazolium, les bactéries Ara+ se voient sous forme de colonies blanches alors que les bactéries Ara- se caractérisent par des colonies rouges. Afin d'utiliser cette méthode, il faut vérifier qu'elle n'ait pas d'effet sur la valeur sélective. Ceci a été confirmé par Lenski de nombreuses fois dans ces conditions mais ne s'avère pas être le cas dans d'autres conditions (Lenski et *al.*, 1991 ; Lenski et *al.*, 1994).

b. Les marqueurs croisés

Si les marqueurs ne sont pas neutres alors il est possible d'utiliser des marqueurs croisés. Les marqueurs croisés sont en général des résistances aux antibiotiques. Dans ce cas, un génotype **a** est marqué avec un antibiotique **A** et un génotype **b** avec un antibiotique **B**. Afin de mesurer les coûts possibles pour les cellules, les antibiotiques sont ensuite inversés. Le génotype **a** est alors marqué avec l'antibiotique **B** et le génotype **b** avec l'antibiotique **A**. Les résultats des 2 compétitions sont ensuite comparés et il est donc possible de connaître le coût qu'a eu l'antibiotique sur la cellule, et d'en déduire le coût associé au fond génétique (Giraud et *al.*, 2001).

L'utilisation de marqueurs croisés n'est pas la méthode la plus simple car elle nécessite de réaliser toutes les manipulations en double.

c. La fluorescence

L'utilisation de marqueur fluorescent tel que la GFP (Green Fluorescent Protein) et de la CFP (Cyan Fluorescent Protein) permet de distinguer les cellules grâce à la couleur différente d'émission : verte pour la GFP et bleue pour la CFP. Cependant, comme nous l'avons vu précédemment, elle nécessite l'utilisation d'un marqueur pouvant avoir un coût pour la cellule. Il suffit donc d'insérer ce marqueur à toutes les cellules présentes lors de la compétition pour qu'elles se retrouvent toutes dans les mêmes conditions (Gallet *et al.*, 2011).

Comme nous pouvons le constater, les méthodes directes sont peu précises et nécessitent donc de très nombreux réplicats et les méthodes par compétition sont soit très laborieuses, soit nécessitent des modifications conséquentes des génomes étudiés qui peuvent en limiter la pertinence.

Nous voulons développer une méthode qui permettrait de modifier efficacement le génome bactérien sans affecter la physiologie de la bactérie transformée et d'utiliser une méthode de séquençage pour mesurer les changements de fréquences alléliques.

V. Les mutations avantageuses

Nous avons vu que la valeur sélective est au cœur des analyses en évolution expérimentale, mais au-delà de l'augmentation de valeur sélective, un des objectifs est d'identifier les mutations avantageuses qui génèrent l'adaptation. Malheureusement, comme les mutations avantageuses sont rares et parfois de faibles effets, il est difficile de les étudier directement. Les premières approches pour les étudier ont donc consisté à essayer d'identifier les mutations qui se sont fixées dans des populations s'étant adaptées. En séquençant les génomes par exemple, on peut identifier les mutations apparues et au moyen d'outils génétiques identifier leur contribution à l'adaptation. Cette méthode est intéressante mais a l'inconvénient de nous renseigner que sur des réalisations *had hoc* du processus d'adaptation et pas de l'ensemble des possibilités. Nous n'avons de l'information que sur le sous-échantillon des mutations qui ont survécu à la dérive et réussi à se fixer au terme d'une potentielle compétition entre souches. Pour aller plus loin, il est possible d'analyser de nombreux réplicats, et l'on voit ainsi en étudiant la convergence entre lignées, à quel point l'adaptation est contrainte et doit utiliser un nombre restreint de mutations. Même avec des

améliorations, ces méthodes ne nous renseignent que peu sur la distribution des effets des mutations avantageuses et sur leur taux (Gallet et *al.*, 2012).

Pour essayer de mieux caractériser la distribution des effets des mutations, d'autres approches ont été développées, basées sur des évolutions à moyen terme en grand nombre de réplicats. L'idée est de mélanger 2 clones portant un marqueur neutre (comme vu précédemment). La fréquence de chacune des 2 sous-populations est alors suivie dans le temps, avec une méthode dépendante du marqueur utilisé. L'apparition d'une mutation avantageuse dans une des 2 sous-populations va entraîner une augmentation en fréquence de celle-ci et à terme son invasion dans la population. La vitesse d'invasion sera directement liée à l'avantage sélectif de la mutation, s et le temps auquel cette invasion sera détectable sera lié à s et au taux de mutation μ . Grâce à la distribution de ces 2 paramètres estimés sur de nombreux réplicats, il est possible d'en estimer une valeur moyenne en utilisant des simulations informatiques reproduisant le processus évolutif. Avec cette méthode, différentes estimations du taux de mutation et de l'effet sélectif ont pu être réalisées.

Bien que cette méthode fournisse des estimations intéressantes et soit capable de percevoir des différences entre différents clones, elle reste néanmoins limitée.

- Ces méthodes ne fournissent qu'une estimation de la moyenne de s et de la moyenne de μ le taux de mutation. Avoir une estimation de la variance serait un avantage non négligeable.
- Il semble qu'il y ait un couplage entre les 2 paramètres, s et μ car la valeur de μ trouvée est toujours la même. En fait, il est possible que le paramètre estimé soit un paramètre composite.
- Le modèle peut être sensible aux interférences clonales qui peuvent rendre la dynamique bien plus complexe que la courbe d'invasion ne permette de le penser. En fait, il est possible que les 2 sous-populations fixent les mêmes mutations avantageuses indépendamment et que l'invasion d'une des 2 sous-populations ne corresponde pas à l'invasion d'une mutation simple mais en fait à l'invasion d'une multitude de mutations.

En conclusion, nos connaissances sur la distribution des effets des mutations avantageuses sont encore très restreintes, et nous ne savons pas encore bien si la décroissance du taux d'adaptation observée en évolution expérimentale correspond avant tout à un

changement d'effet des mutations ou à un changement de taux de mutations avantageuses. Il nous faut alors réfléchir à de nouveaux outils pour progresser, notamment en multipliant le nombre de marqueurs génétiques à utiliser comme je le propose dans mon projet.

Bibliographie

ALLEN ORR H.2005.

The genetic theory of adaptation: a brief history.
Nature 6:119-127

BABIC A., GUEROUT AM., MAZEL D.2008.

Construction of an improved RP4 (RK2)-based conjugative system.
[Res Microbiol](#).159(7-8):545-9.

BARRICK J. E., D. S. YU., S. H. YOON., H JEONG T. KOH., D SCHNEIDER, R. E. LENSKI., and J. F. KIM. 2009.

Genome evolution and adaptation in a long-term experiment with *Escherichia coli*.
Nature 461:1243-7.

BENTLEY R., and R. MEGANATHAN. 1982.

Biosynthesis of vitamin K (menaquinone) in bacteria.
Microbiol Rev 46:241-80.

BLATTNER F.R., G.PLUNKETT,3rd., C.A. BLOCH., N.T. PERNA., V.BURLAND., M.RILEY., J.COLLADO-VIDES., J.D.GLASNER., C.K. RODE., G.F. MAYHEW., J.GREGOR., N.W. DAVIS., H.A. KIRKPATRICK., M.A. GOEDEN., D.J. ROSE., B.MAU., and Y.SHAO., 1997

The complete genome sequence of *Escherichia coli* K-12. Science 277:1453-74

BLOUNT Z D., BORLAND C Z., LENSKI R E.2008.

Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*.
PNAS 105:23 7899-7906

CROZAT E., HINDRE T., KUHN L., GARIN., LENSKI R E, SCHNEIDER D.2011.

Altered regulation of the OmpF porin by Fis in *Escherichia coli* during an evolution experiment and between B and K-12 strains.
[J Bacteriol](#) 193(2):429-40.

DARWINN C. (1859). The origin of Species by Means of Natural Selection, London.

DAVIES J.1994.

Inactivation of antibiotics and the dissemination of resistance genes.
SCIENCE, 264, 375-382.

DAVIES J. 1996.

Bacteria on the rampage.

NATURE, 383, 219-220 .

DEMARRE G., A M GUEROUT., C MATSUMOTO-MASHIMO.,D A ROWE-MAGNUS.,
P MARLIERE.,D MAZEL.2004.

A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncP₋) conjugative machineries and their cognate *Escherichia coli* host strains.

Research in Microbiology 156 : 245–255

DOUGAN G.2009.

Does the Trojan horse have an Achilles' heel?

N Engl J Med 360:83-4

ESCOBAR-PARAMO P., O CLERMONT., A. B BLANC-POTARD H., BUI C., LE
BOURGUENEC.,E DENAMUR. 2004.

A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*.

Mol Biol Evol 21:1085-94.

EYRE-WALKER A., P. D. KEIGHTLEY.2007.

The distribution of fitness effects of new mutations.

Nat Rev Genet 8, 610

FISCHER J.F., MEROUEH S.O., MOBASHERY S. 2005.

Bacterial resistance to beta-lactam antibiotics: compelling opportunism, compelling opportunity.

Chem Rev, 105, 395-424.

FISHER R. A. (1930). The genetical theory of natural selection. Oxford University Press, Oxford.

GALLET R., T F COOPER., SANTIAGO F. E.,T LENORMAND.2011.

Measuring Selection Coefficients Below 10⁻³: Method, Questions, and Prospects.

Genetics 190:175-186

GIRAUD A., MATIC I., TENAILLON O., CLARA A., RADMAN M., FONS M., TADDEI F.2001.

Costs and Benefits of High Mutation Rates: Adaptive Evolution of Bacteria in the Mouse Gut
Science 291 5513 :2606-2608

GRIMONT P. A. 1988.

Use of DNA reassociation in bacterial classification.

Can J Microbiol 34:541-6.

HINDRE O., KNIBBE C, BESLON G., SCHNEIDER D.2012.

New insights into bacterial adaptation through in vivo and in silico experimental evolution.

Nature Reviews Microbiology 10, 352-365

HOBMAN JL., W PENN., MJ PALLEN.2007.

Laboratory strains of *E. coli*: model citizens or deceitful delinquents growing old disgracefully?
Mol Microbiol 64:881-5

JAUREGUY F., L LANDRAUD V., V PASSET., L DIANCOURT., E FRAPY., G GUIGON., E CARBONELLE., O LORTHOLARY., O CLERMONT., E DENAMUR., B PICARD., X NASSIF., S BRISSE.2008.
Phylogenetic and genomic diversity of human bacteremic *E. coli* strains.
BMC Genomics 9:560

KADER A A., A KUMAR., K A KAMATH. 2007.
Fecal carriage of extended-spectrum beta-lactamase-producing *Escherichia coli* and *Klebsiella pneumoniae* in patients and asymptomatic healthy individuals.
Infect Control Hosp Epidemiol 28:1114-6.

KIMURA M, OHTA T. 1969.
The Average Number of Generations until Fixation of a Mutant Gene in a Finite Population.
Genetics 61(3):763–771

KIRILL A., DATSENKO AND BARRY L., WANNER. 2000.
One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.
PNAS 12: 6640-6645.

KUHN J., CAMPBELL A.2001.
The Bacteriophage lambda Attachment Site in Wild Strains of *Escherichia coli*.
J Mol Evol 53:607–614

LANDY A., WEISBERG R A. 1983.
Site-specific recombination in phage lambda.
Lambda II Cold Spring Harbor Laboratory. P212

LECOINTRE G., L RACHDI., P DARLU., E DENAMUR.1998. *E. coli* molecular phylogeny using the incongruence length difference test.
Mol Biol Evol 15:1685-95

LE GALL T., O CLERMONT., S GOURIOU., B PICARD., X NASSIF., E DENAMUR., O TENAILLON.2007.
Extraintestinal virulence is a coincidental by product of commensalisms in B2 phylogenetic group *E. coli* strains.
Mol Biol Evol 24:2373-84

LENSKI R E., ROSE M R., SIMPSON S C., TADLER S C.1991.
Long term experimental evolution in *Escherichia coli*. Adaptation and divergence during 2 000 generations.

The American naturalist 138 (6):1315-1341

LENSKI R E., M TRAVISANO.1994.

Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations.

Proc. Nati. Acad. Sci. USA Vol. 91, pp. 6808-6814, July 1994

LEVY S.B.1998.

The challenge of antibiotic resistance.

Sci Am, 278, 46-53.

MOLLY K., BURKE., M R ROSE.2009.

Experimental evolution with *Drosophila*.

Am J Physiol Regul Integr Comp Physiol 296: R1847–R1854

NIKAIDO H. 1996.

Multidrug efflux pumps of gram-negative bacteria.

J Bacteriol, 178, 5853-5859.

NOWROUZIAN F., B. HESSELMAR., R. SAALMAN., I. L. STRANNEGARD., N. ABERG., A. E. WOLD., and I. ADLERBERTH. 2003.

Escherichia coli in infants' intestinal microflora: colonization rate, strain turnover, and virulence gene carriage.

Pediatr Res 54:8-14.

NURUL ATIFAH M A., H K LOO., G SUBRAMNIAM., E H WONG., P SELVI., S E HO., A KAMARULZAMAN., N PARASAKTHI. 2005.

faecal prevalence of extended spectrum Beta-lactamase (ESBL)-producing coliforms in a geriatric population and among haematology patients.

Malays J Pathol 27:75-81.

OCHMAN H., and R. K. SELANDER. 1984.

Standard reference strains of *Escherichia coli* from natural populations.

J Bacteriol 157:690-3.

PALLECCHI L., A BARTOLONI C., FIORELLI A., MANTELA., T DI MAGGIO., H. GAMBOA., E GOTUZZO., G KRONVALL., F. PARADISI., G M ROSSOLONI. 2007.

Rapid dissemination and diversity of CTX-M extended-spectrum beta-lactamase genes in commensal *Escherichia coli* isolates from healthy children from low-resource settings in Latin America.

Antimicrob Agents Chemother 51:2720-5.

PERNA, N.T., G.PLUNKETT.,^{3rd}, V.BURLAND., B.MAU., J.D. GLASNER., D.J. ROSE., G.F. MAYHEW., P.S. EVANS., J.GREGOR., H.A.KIRKPATRICK., G.POSFAI., J.HACKETT., S.KLINK., A.BOUTIN., Y.SHAO., L.MILLER., E.J. GROTBECK., N.W.DAVIS., A.LIM., E.T. DIMALANTA., K.D. POTAMOUSIS., J.APODACA., T.S.

ANANTHARAMAN., J.LIN., G.YEN., D.C. SCHWATZ., R.A. WELCH., and F.R. BLATTNER. 2001
Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 409:529-33

PUPO GM., R LAN., P R REEVES.2000.
Multiple independent origins of Shigella clones of *E. coli* and convergent evolution of many of their characteristics.
Procs Natl Acad Sci USA 97:105467-72

ROSKO DA., WEBSTER DR, SAHL JW, BASHIR A, BOISEN N, SCHEUTZ F, PAXINOS EE, SEBRA R, CHIN CS, ILIOPOULOS D, KLAMMER A, PELUSO P, LEE L, KISLYUK AO, BULLARD J, KASARSKIS A, WANG S, EID J, RANK D, REDMAN JC, STEYERT SR, FRIMODT-MOLLER J, STRUVE C, PETERSEN AM, KROGFELT KA, NATARO JP, SCHADT EE, WALDOR MK.2011.
[Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany.](#)
N Engl J Med. 8:709-17.

SANTIAGO F. EL., R E. LENSKI.2003.
Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation.
GENETICS 4:457-469

SCHAMBERGER G. P., R. L. PHILLIPS., J. L. JACOBS., and F DIEZ-GONZALEZ. 2004.
Reduction of *Escherichia coli* O157:H7 populations in cattle by addition of colicin E7-producing *E. coli* to feed.
Appl Environ Microbiol 70:6053-60

SNIEGOWSKI P D., PHILIP J. GERRISH., R E. LENSKI.1997.
Evolution of high mutation rates in experimental populations of *E. coli*
NATURE 387:703-705

SOUZA A., MAGALHAES S., GORDO I.
Cost of antibiotic resistance and the geometry of adaptation.
Mol Biol Evol may 2012 1417-28.

TENAILLON O., SKURNIK D., PICARD B., DENAMUR E.2010.
[The population genetics of commensal *Escherichia coli*.](#)
Nat Rev Microbiol 3:207-17

TOUCHON M., C HOEDE., O TENAILLON., V BARBE., S BAERISWYL., P BIDET., E BINGEN., S BONACORSI., C BOUCHIER., O BOUVET., A CALTEAU., H CHIAPELLO., O CLERMONT., S CRUVEILLER., A DANCHIN., M DIARD., C DOSSAT., M E KARAOUI., E FRAPPY., L GARRY., JM GHIGO., AM GILLES., J JOHNSON., C LE BOUGUENEC., M LESCAT., S MANGENOT., V MARTINEZ-JEHANNE., I MATIC., X NASSIF., S OZTAS., MA PETIT., C PICHON., Z ROUY., CS RUF., D SCHNEIDER., J TOURET., B VACHERIE., D VALLENET., C MEDIGUE., EP ROCHA., E DENAMUR.2009.
Organised genome dynamics in the *E. coli* species results in highly diverse adaptive paths.
PloS Genet 5:e1000344

VOLLAARD E. J., and H. A. CLASENER. 1994.

Colonization resistance.

Antimicrob Agents Chemother 38:409-14.

WALSH C. 2000.

Molecular mechanisms that confer antibacterial drug resistance.

NATURE, 406, 775-781.

WAYNE L. B., DJ; COLWELL RR; GRIMONT, PAD; KANDLER, O; KRICHEVSKY, MI;
MOORE, LH; MOORE, WEC; MURRAY, RGE; STACKEBRANDT, E; STARR, MP;
TRUPER, HG. 1987.

Report ad hoc committee on reconciliation of approaches to bacterial systematics.

International Journal of Systematic Bacteriology Oct:463-464.

WHITTAM T. S., H. OCHMAN ., R. K. SELANDER. 1983.

Multilocus genetic structure in natural populations of *Escherichia coli*.

Proc Natl Acad Sci U S A 80:1751-5.

WOODS R .,J E Barrick.,T F. COOPER., U SHRESTHA.,M R KAUTH.,R E LENSKI.2001.

Second-order selection for evolvability in a large *Escherichia coli* population.

Science 331(6023): 1433–1436.

Encarta: Antibiotiques. 2007.

In: *Encyclopédie Microsoft Encarta en ligne 2007*.

Microsoft Corporation; 2007.