



HAL
open science

Evolution expérimentale de bactéries par brassage de génom

Nadia Jezequel

► **To cite this version:**

Nadia Jezequel. Evolution expérimentale de bactéries par brassage de génome. Biologie du développement. Université Pierre et Marie Curie - Paris VI, 2012. Français. NNT : 2012PAO66404 . tel-00831559

HAL Id: tel-00831559

<https://theses.hal.science/tel-00831559>

Submitted on 7 Jun 2013

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.

ECOLE NORMALE SUPÉRIEURE - LABORATOIRE PIERRE AIGRAIN

**THÈSE DE DOCTORAT DE
L'UNIVERSITÉ PIERRE ET MARIE CURIE**

Spécialité

Biologie

École doctorale Interdisciplinaire pour le vivant

Présentée par

Nadia JEZEQUEL

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

**Evolution expérimentale de bactéries par
brassage de génomes**

soutenue le 26 septembre 2012

devant le jury composé de :

| | |
|--------------------------|--------------------|
| Mme Sophie CRIBIER | Présidente du jury |
| M. Olivier TENAILLON | Rapporteur |
| M. Bahram HOUCHMANDZADEH | Rapporteur |
| M. Ivan MATIC | Examineur |
| M. Gilles FISCHER | Examineur |
| M. François HESLOT | Directeur de thèse |

Remerciements

Je souhaite remercier toutes les personnes que j'ai côtoyées au cours de ma thèse et de mes études.

Tout d'abord je tiens à remercier les membres de mon jury de thèse pour leurs regards critiques et leurs échanges de point de vue. Un grand merci à Olivier Tenailon et Bahram Houchmandzadeh qui ont accepté d'être rapporteurs et qui ont émis leur rapport dans des temps records. Je remercie aussi Gilles Fischer, Ivan Matic et Sophie Cribier de m'avoir fait l'honneur d'être examinateurs.

Je tiens aussi à remercier les membres de l'équipe de biophysique du Laboratoire Pierre Aigrain :

- François Heslot pour avoir accepté de me prendre en thèse dans son équipe et pour m'avoir dirigé pendant ces trois ans. Je n'oublierais pas nos longues discussions autour d'une tasse de thé et notamment toutes nos réflexions sur l'amélioration du montage et des protocoles pour le brassage. Merci pour ta grande disponibilité même pendant tes congés !
- Philippe Thomen pour m'avoir accueilli et encadré sur son projet de recherche. Grâce à l'expérience pilote que tu as mis au point et réalisé, tu m'a initié aux expériences d'évolution et surtout à l'analyse passionnante de cette expérience. Je n'oublierai pas l'enthousiasme que m'a procuré l'analyse des mutations apparues au cours de cette expérience. Encore merci pour avoir partagé mon quotidien au labo au cours de ces trois années et d'avoir facilité mon adaptation à un environnement quelque peu hostile pour moi (pas facile pour la biologiste que je suis d'évoluer dans un monde de physiciens!).
- Denis Cote pour son avis éclairé et son soutien notamment lors de la préparation à la soutenance.

Je tiens aussi à remercier Marco Cosentino Lagomarsino ainsi que sa stagiaire Maria Fumagalli pour leur collaboration sur le projet pour la partie théorique. Merci Marco pour avoir voulu collaborer avec nous et d'avoir apporté un nouveau regard sur ce projet.

Je souhaite aussi remercier tous les membres du Laboratoire Pierre Aigrain. Remerciements particuliers aux services de mécanique et d'électronique pour leur aide apportée sur la conception du montage et la mise en place de la mesure de DO : Pascal

Morfin, Jean-Charles Dumont, Jules Silembo, Philippe Pace, David Darson et Anne Denis. Remerciements aussi à Jean-Marc Berroir, directeur du laboratoire Pierre Aigrain, et aux secrétaires Fabienne Renia et Anne Matignon pour leur gentillesse et leur aide administrative.

Je remercie aussi l'école doctorale Interdisciplinaire pour le Vivant pour avoir retenu ma candidature pour ce projet de thèse. Je tiens à féliciter l'ouverture d'esprit de cette école doctorale qui permet à des personnes de formations différentes de se réunir autour de l'étude du vivant, rompant ainsi le cloisonnement souvent réalisé entre les disciplines. Je remercie en particulier Germain Trugnan, Sophie Cribier et Karine Gherdi pour le suivi de ma thèse.

Je tiens aussi à remercier Baya Cherif-Zahar et Giovanni Stevanin pour m'avoir pris en stage en master et de m'avoir ainsi donné la chance de mettre un pied dans le monde de la recherche.

Enfin je souhaite remercier ma famille et mes amis pour leur soutien et leurs encouragements tout au long de mes études.

Résumé

Cette thèse porte sur des expériences d'évolution de bactéries, avec le but d'explorer expérimentalement comment des transferts horizontaux d'ADN au sein d'une population de bactéries peuvent ou non modifier l'évolution de cette population. La méthode utilisée pour permettre des échanges d'ADN à l'intérieur de la population repose sur la propriété de transformation naturelle de la bactérie *Acinetobacter baylyi*. Une expérience pilote sans brassage de génomes a été faite sur 3000 générations et nous avons analysé par séquençage les mutations et la diversité apparues pendant cette expérience. On observe une structure en multiples sous-populations compatible avec un régime d'interférence clonale. Une expérience d'évolution avec échange d'ADN ainsi qu'une expérience contrôle avec une bactérie non compétente ont été réalisées dans les mêmes conditions sur 1300 générations. Des séquençages de génomes ont été effectués sur des populations et des clones à différents temps de chaque expérience d'évolution afin de mettre en évidence les mutations apparues ainsi que leur fréquence dans la population. La diversité dans l'expérience d'évolution avec brassage est nettement plus élevée jusqu'à 970 générations et est due à différents types de mutations alors que dans l'expérience sans brassage, la diversité est essentiellement due à des SNP. Cependant dans l'expérience avec brassage, on observe à 1300 générations une baisse de la diversité qui est liée à la fixation d'une mutation dans un gène essentiel pour la transformation. Ces résultats préliminaires laissent penser que le brassage a un avantage à court terme mais qu'à long terme l'avantage lié à l'échange s'amenuise.

Mots clés : évolution, bactérie, transferts horizontaux d'ADN, transformation naturelle, chemostat, reséquençage, diversité

Abstract

This thesis focuses on experiments in bacterial evolution, with the aim to explore experimentally how horizontal transfer of DNA within a bacterial population, may or may not modify the evolution of this population. The method used to allow exchanges of DNA within the population, is based on the property of natural transformation of the bacterium *Acinetobacter baylyi*. A pilot experiment without genome shuffling was made over 3000 generations, and we analyzed by sequencing the mutations and the diversity that emerged during this experiment. There is a structure in multiple actively competing subpopulations that appears consistent with a system of clonal interference. A experimental evolution with exchange of DNA and a control experiment with a non-competent bacterium were performed under the same conditions over 1300 generations. The sequencing of genomes have been conducted on populations and clones at different times of each experiment to detect the mutations that occurred and their frequency in the population. Diversity in the evolution experiment with genome shuffling is significantly higher up to 970 generations and is due to different types of mutations, whereas in the control experiment diversity is mainly due to SNP. However in the experiment with shuffling, we observe a decrease in the diversity at 1300 generations which is due

to the fixation of a mutation in a gene essential for transformation. These preliminary results suggest that the genome shuffling has a short term benefit while on the long term the benefit from the exchange is dwindling.

Key words : evolution, bacteria, horizontal DNA transfer, natural transformation, chemostat, resequencing, diversity

Table des matières

| | |
|--|-----------|
| Introduction | 9 |
| 1 Introduction à l'Évolution | 13 |
| 1.1 Mécanismes de l'évolution | 13 |
| 1.1.1 Création de diversité par mutations et réarrangements de séquences | 14 |
| 1.1.1.1 Les mutations ponctuelles | 14 |
| 1.1.1.2 Les réarrangements de séquences | 14 |
| 1.1.1.3 Effets des mutations | 16 |
| 1.1.2 Evolution de la fréquence des variations au sein d'une population par sélection ou dérive | 18 |
| 1.1.2.1 La sélection naturelle | 19 |
| 1.1.2.2 La dérive génétique | 20 |
| 1.2 Apports des expériences d'évolution avec des microorganismes | 21 |
| 1.2.1 Avantages et inconvénients | 21 |
| 1.2.2 Principe | 22 |
| 1.2.3 Principales informations apportées par les expériences d'évolu- tion | 23 |
| 1.2.3.1 Estimation du taux de mutation et distribution des ef- fets des mutations bénéfiques | 24 |
| 1.2.3.2 Comparaison des effets de la dérive et de la sélection . | 25 |
| 1.2.3.3 Diversité | 26 |
| 1.2.3.4 Interaction entre espèces | 26 |
| 1.2.3.5 Concept de macroévolution | 26 |
| 1.2.4 Perspectives | 27 |
| 2 Apport évolutif du sexe et des transferts horizontaux | 29 |
| 2.1 Théorie de Fisher-Müller | 31 |
| 2.2 Cliquet de Müller | 32 |
| 2.3 Théorie de la Reine Rouge | 33 |
| 2.4 Autres théories | 33 |

| | | |
|----------|---|-----------|
| 2.4.1 | Hypothèse mutationnelle déterministique | 33 |
| 2.4.2 | "Rubby in the rubbish" | 33 |
| 2.4.3 | "Tangled bank" | 34 |
| 2.4.4 | Modèle de la loterie | 34 |
| 2.5 | Test expérimental de l'apport évolutif du sexe et des transferts horizontaux | 34 |
| 2.5.1 | Expériences In vitro | 34 |
| 2.5.2 | Expérience d'évolution avec des microorganismes | 35 |
| 2.5.2.1 | Virus | 35 |
| 2.5.2.2 | Bactéries | 35 |
| 2.5.2.3 | Eucaryotes unicellulaires | 37 |
| 2.5.3 | But de la thèse | 38 |
| 3 | <i>Acinetobacter baylyi</i>, une bactérie naturellement compétente | 39 |
| 3.1 | Machinerie de transformation naturelle chez <i>Acinetobacter baylyi</i> | 40 |
| 3.1.1 | Effet du type d'ADN exogène sur la transformation naturelle | 40 |
| 3.1.2 | Dégradation de l'ADN exogène lors de la transformation naturelle | 42 |
| 3.1.3 | Gènes impliqués dans la translocation de l'ADN exogène | 43 |
| 3.1.4 | Modèle de mécanisme de translocation de l'ADN exogène | 44 |
| 3.2 | Induction de la compétence chez <i>Acinetobacter baylyi</i> | 45 |
| 3.2.1 | Taux de transformation en fonction de la phase de croissance | 46 |
| 3.2.2 | Effet du nutriment limitant sur l'induction de la compétence | 47 |
| 3.2.3 | Régulation de l'expression des gènes de compétence | 47 |
| 3.2.4 | Effet du pH sur l'induction de la compétence | 47 |
| 3.3 | Recherche des conditions optimales pour le brassage de génomes | 48 |
| 3.3.1 | Paramètres influençant le taux de transformations de notre souche cultivée en milieu discontinu | 49 |
| 3.3.1.1 | Effet de l'ADN transformant | 49 |
| 3.3.1.2 | Effet du temps de transformation | 50 |
| 3.3.1.3 | Effet de la température | 51 |
| 3.3.1.4 | Effet du milieu et du protocole de transformation | 51 |
| 3.3.2 | Taux de transformation d'une culture d' <i>Acinetobacter baylyi</i> en réacteur | 52 |
| 4 | Montages et protocoles pour les expériences d'évolution en chemostat | 55 |
| 4.1 | Principe de fonctionnement d'un chemostat | 55 |
| 4.2 | Montage pour les expériences d'évolution avec ou sans brassage de génomes | 57 |
| 4.2.1 | Réacteurs | 57 |
| 4.2.2 | Connexions du réacteur | 59 |
| 4.2.3 | Arrivée de milieu | 59 |
| 4.2.4 | Arrivée d'air | 60 |
| 4.2.5 | Evacuation du trop plein du réacteur vers une poubelle ou une bouteille de prélèvement | 60 |
| 4.2.6 | Transfert et système de rinçage | 62 |
| 4.2.7 | Régulation en température | 63 |

| | | |
|----------|---|------------|
| 4.2.8 | Mesure de la densité optique | 63 |
| 4.3 | Protocole du brassage de génomes | 64 |
| 4.3.1 | Extraction de l'ADN génomique de la population du réacteur | 64 |
| 4.3.2 | Transformation de la population du réacteur avec l'ADN génomique extrait | 64 |
| 4.3.3 | Réinjection des bactéries transformées dans le réacteur | 65 |
| 5 | Expériences d'évolution avec ou sans brassages de génomes | 67 |
| 5.1 | Expérience pilote | 67 |
| 5.2 | Expériences d'évolution avec ou sans brassage de génomes (sem et lud) | 78 |
| 5.2.1 | Déroulement des expériences d'évolution avec ou sans brassage (sem et lud) | 78 |
| 5.2.2 | Séquençage de génomes de clones et de populations | 79 |
| 5.2.2.1 | Résultats des séquençages de génomes dans l'expérience d'évolution sans brassage de génome (lud) | 80 |
| 5.2.2.2 | Résultats des séquençages de génomes dans l'expérience d'évolution avec brassage de génomes (sem) | 83 |
| 5.2.3 | Comparaison des expériences d'évolution avec ou sans brassage de génomes | 87 |
| 5.2.3.1 | Loci mutés en commun | 87 |
| 5.2.3.2 | Comparaison de la diversité | 89 |
| 5.2.3.3 | Comparaison des types de mutations | 92 |
| 5.2.3.4 | Comparaison des SNP dans les expériences d'évolution avec ou sans brassage de génome (sem et lud) | 93 |
| 5.2.3.5 | Résumé | 94 |
| | Conclusion | 97 |
| | Annexes | 99 |
| | Protocole d'extraction d'ADN génomique | 99 |
| | Amplification de l'ADN génomique par l'ADN polymérase du bactériophage phi29 | 100 |
| | Données supplémentaires de l'article sur l'expérience d'évolution pilote | 101 |
| | Bibliographie | 109 |

Introduction

Objectif de la thèse : Cette thèse porte sur l'évolution darwinienne de bactéries en laboratoire, avec le but d'explorer expérimentalement comment des transferts horizontaux d'ADN au sein d'une population d'individus initialement identiques, peuvent modifier l'évolution de cette population. Cette problématique est à replacer dans le cadre des acquis récents du séquençage de masse, qui ont accrédité la présence de multiples traces de transferts horizontaux d'ADN, c'est-à-dire l'acquisition au cours de l'évolution de fragments d'ADN non-parentaux. Un échange d'ADN horizontal porte en lui certaines caractéristiques du sexe chez les organismes supérieurs, avec potentiellement le bénéfice évolutif de combinaison entre allèles d'individus différents. La question que nous posons est de savoir si en mettant en place des conditions où les échanges d'ADN inter-bactéries sont soit très facilités, soit (en contrôle) essentiellement interdits, on obtient sur des échelles de temps de laboratoire, une évolution darwinienne différente ou pas de la population sous une contrainte sélective donnée. La bactérie modèle choisie est *Acinetobacter baylyi* ADP1, et la méthode choisie pour permettre des échanges faciles d'ADN à l'intérieur de la population repose sur la propriété de transformation naturelle exceptionnelle de cette bactérie. Un réacteur en flux continu est utilisé, permettant le maintien d'une très grande population de bactérie (10^{10}) se divisant continûment dans des conditions de cultures et d'environnements constants.

Introduction : L'échange d'ADN inter-individus peut se faire via le sexe entre individus compatibles, ou encore via une panoplie de mécanismes qui sont courants chez la bactérie tels la conjugaison, la transduction, via des vecteurs viraux ou plasmidiques, etc ... C'est aussi un problème de santé publique (par exemple propagation de résistance aux antibiotiques, apparition de multirésistances) ou encore de sauvegarde de l'environnement (posé par exemple avec la présence/dissémination possible d'OGM). Du point de vue de l'évolution Darwinienne de la bactérie soumise à un traitement où on cherche à l'éliminer, l'avantage obtenu par la bactérie qui réalise l'acquisition d'une résistance est évident. A l'échelle d'un écosystème, un échange d'ADN entre bactéries génétiquement distantes peut permettre aussi l'acquisition/intégration de gènes isolés ou même d'opérons entiers permettant de nouvelles fonctions chez la bactérie receveuse [94]. Les tailles de populations bactériennes sont en général très importantes,

des conditions de promiscuité/confinement souvent réunies, de sorte que même des échanges rares inter-individus peuvent être à la source d'effets importants. A l'opposé, les tailles de population des organismes supérieurs sont souvent limitées, et il a été proposé depuis longtemps [44,99] que le sexe, qui permet la génération rapide de variants via un réassortiment d'allèles, est un moyen puissant pour la survie (via des individus mieux adaptés) face aux assauts des compétiteurs et /ou de l'environnement et de ses changements.

Sexe, "DNA shuffling" et consorts : Un échange d'ADN, qui peut être vu sous l'angle d'un rebattage des cartes ("DNA shuffling") pourrait être un moteur d'évolution important. De façon extrêmement étonnante, les éléments expérimentaux *in vivo* qui pourraient permettre de trancher (déterminer l'avantage du sexe et/ou de l'échange d'ADN), sont parcellaires et peu convaincants [24, 26, 49, 65, 140] ; même la question de l'avantage du sexe par rapport à son coût en apparence exorbitant (deux individus nécessaires pour en faire un) n'est pas résolue. A l'opposé extrême, dans le domaine des biotechnologies et de l'évolution *in vitro* pour produire des variants d'intérêt (par exemple une protéine particulière), l'adaptation du « sexe » sous la forme brute de "DNA shuffling" [128] a permis d'améliorer considérablement la vitesse avec laquelle on arrive à obtenir des variants intéressants. Ce "DNA shuffling" fait intervenir coupure /mélange /réassortiment *in-vitro* de séquences d'ADN choisies parmi celle qui codent par exemple déjà pour des mutants présentant un avantage partiel. Un ou plusieurs cycles de "shuffling" permettent alors de combiner ces mutants, résultant souvent en des mutants encore améliorés. A l'échelle de la séquence d'ADN, cette évolution est la combinaison de mutations (à laquelle il aurait été virtuellement impossible d'arriver, par cumuls successifs (uniquement sériel) de mutations sur une même séquence, à cause de la taille de la "bibliothèque" de séquence qui serait nécessaire). En d'autre terme le "DNA shuffling" *in vitro* permet par des mécanismes non pas d'accumulations sérielles (clonaux) mais d'échanges (parallélisés) de fragments entre classes, d'obtenir une "accélération d'évolution". Enfin, quelques expériences de "genome shuffling" *in vivo* ont été réalisées avec succès, mais elles font intervenir un nombre faible d'individus pour réaliser ce "shuffling" [29, 114, 142].

La compétence naturelle : Parmi les mécanismes d'échange entre bactéries, il en est un qui présente des caractéristiques hors du commun : il s'agit de la compétence naturelle [20], une propriété que possèdent certaines bactéries. Cette compétence naturelle se traduit par la capacité pour une bactérie à capturer de l'ADN si celui-ci est présent dans le milieu de culture, le faire entrer dans son milieu intracellulaire, et ensuite l'intégrer sous certaines conditions dans son génome. A l'issue de ces étapes, la bactérie se retrouve munie dans son génome d'un nouveau segment d'ADN. Cet effet est remarquable, car il n'existe que peu de contraintes sur la source de l'ADN qui est intégré : il n'est pas nécessaire d'avoir une bactérie donneuse spécialisée, il n'est pas nécessaire que l'ADN ait été au préalable intégré dans un vecteur viral ou plasmidique. La simple présence d'un fragment d'ADN linéaire « nu » c'est-à-dire sans protection particulière, et issu par exemple des débris d'une bactérie morte, peut suffire. Cet ADN

peut se recombinaison dans le génome de la bactérie compétente si des homologies partielles existent entre un/des segments de l'ADN entrant, et un/des segment de l'ADN génomique. Une bactérie naturellement compétente se présente donc comme une généraliste de l'échange d'ADN à tout crins. La question se pose évidemment de savoir quel est l'avantage évolutif fourni par une telle capacité, deux options extrêmes peuvent être citées [125] : dans l'une il est proposé que l'ADN sert de nourriture et l'intégration dans le génome est un épiphénomène ; dans l'autre, cette capacité fournit à la bactérie un avantage évolutif (mais alors pourquoi la compétence ne s'est alors pas systématiquement répandue chez les bactéries ?). Si on examine le catalogue des diverses espèces de bactéries identifiées comme étant naturellement compétentes, on trouve des bactéries très diverses, avec *Bacillus subtilis* bactérie courante et non-pathogène, *Deinococcus radiodurans* (bactérie ayant une capacité hors du commun à résister à la dessiccation (et aux radiations), etc ... On trouve aussi des bactéries pathogènes, telles *Streptococcus pneumoniae*, *Helicobacter pylori*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Haemophilus influenzae*, etc..., suggérant éventuellement que la compétence naturelle pourrait aussi être un moyen de générer des variants permettant à la bactérie d'échapper à la surveillance du système immunitaire. On débouche donc encore sur l'idée qu'un rebattage des cartes ("DNA shuffling") entre individus pourrait permettre à la bactérie d'explorer rapidement des combinaisons d'allèles, et "d'accélérer" son évolution sous contrainte sélective forte [69].

Organisation des chapitres de la thèse : Dans un premier chapitre introductif, j'introduis les principaux mécanismes de l'évolution ainsi que l'apport des expériences d'évolution dans la compréhension de ces mécanismes.

Dans le deuxième chapitre, j'aborde la bibliographie des travaux portant sur l'apport évolutif du sexe et des transferts horizontaux.

Dans un troisième chapitre, je présente les propriétés de transformation naturelle d'*Acinetobacter baylyi* ainsi que la mise au point des conditions pour réaliser des brassages de génomes au sein de la population.

Dans un quatrième chapitre, je présente le montage utilisé ainsi que les protocoles utilisés pour les expériences d'évolution avec brassage de génomes. La contrainte sélective utilisée pour l'évolution expérimentale est l'utilisation d'un milieu minimum où la seule source d'azote est du nitrate. Une expérience pilote a d'abord été menée, afin de mettre en place le montage expérimental basé sur un réacteur à flux continu, de déterminer quelle était l'évolution obtenue sur un temps de l'ordre de quelques mois, et également pour déterminer quelles étaient les variations obtenues par évolution. J'ai ensuite mené une expérience d'évolution avec brassage de génome avant de réaliser une expérience de contrôle dans les mêmes conditions, mais avec un mutant d'*Acinetobacter baylyi* rendu non-compétent de façon à interdire les transferts horizontaux.

Dans le cinquième chapitre, je décris les résultats obtenus avec des analyses par réséquençage réalisés tous les 165 générations à la fois sur la population de bactérie, et sur un clone choisi aléatoirement. Ce type d'analyse a été mené sur l'expérience de brassage et son contrôle. Les résultats obtenus lors de l'expérience pilote sont aussi présentés (article soumis).

Dans la conclusion, je résume les résultats saillants, les conclusions qu'on peut en tirer et leurs limitations.

Chapitre 1

Introduction à l'Évolution

Dans ce chapitre introductif, je présenterai dans un premier temps, les principales notions liées aux mécanismes de l'évolution (voir 1.1) avant d'aborder l'apport des expériences d'évolutions avec des microorganismes dans la compréhension des mécanismes de l'évolution (voir 1.2). L'influence du mode de reproduction sur l'évolution sera abordée plus en détail dans le chapitre 2.

1.1 Mécanismes de l'évolution

Plusieurs théories se sont succédées pour tenter d'expliquer la diversité et l'évolution des espèces. La théorie actuellement adoptée est la théorie synthétique de l'évolution ou néodarwinisme. Elle repose sur la théorie darwinienne de la sélection naturelle [30] et a été complétée par les connaissances de la génétique, inconnues à l'époque de Darwin. Les éléments principaux sont les suivants :

- tout d'abord, des variations aléatoires des caractères apparaissent chez les individus de même espèce et certaines de ces variations sont hérissables. Cette diversité phénotypique est créée par des mutations et des réarrangements aléatoires au niveau du génome de l'espèce (voir 1.1.1). Ces mutations peuvent avoir différents effets sur le phénotype (voir section 1.1.1.3).
- la fréquence de ces variations hérissables change selon les forces évolutives qui s'exercent. Les principales forces sont les mutations, la sélection naturelle (voir section 1.1.2.1), la dérive (voir section 1.1.2.2) et la migration.

Le processus d'évolution dépend du mode de reproduction et j'aborderai dans le chapitre 2 la question de l'apport évolutif du sexe.

1.1.1 Création de diversité par mutations et réarrangements de séquences

1.1.1.1 Les mutations ponctuelles

Les mutations ponctuelles correspondent à la substitution d'une base par une autre (SNP pour Single Nucleotide Polymorphism) ou à l'insertion/délétion d'une base (indel). Elles résultent souvent d'erreurs lors de la réplication de l'ADN par l'ADN polymérase ou lors de la réparation de l'ADN. En effet bien que l'ADN polymérase ait une activité de relecture (proofreading) et qu'il existe des systèmes de correction/réparation, des erreurs arrivent à subsister avec pour la bactérie *Escherichia coli*, un taux de mutation ponctuelle par base et par génération estimé entre $1,5 \cdot 10^{-11}$ et $5,4 \cdot 10^{-10}$ [38, 101, 136]. Cependant toutes les substitutions n'ont pas la même probabilité : par exemple, chez les bactéries, il a été montré que les transitions $G \rightarrow A$ et $C \rightarrow T$ sont les mutations ponctuelles plus fréquentes [57, 59, 136]. Les mutations ponctuelles peuvent aussi être induites par des agents chimiques (EMS, agents alkylants, analogue de base ...) ou physiques (UV, radiations ionisantes ...).

1.1.1.2 Les réarrangements de séquences

Parmi les réarrangements de séquences, on distingue :

- **les insertions ou délétions de séquences (indel)**. Elles peuvent aussi résulter d'erreur lors de la réplication notamment lorsqu'il y a des séquences répétées qui peuvent induire des décalages de l'ADN polymérase à cause de la formation de boucles d'ADN à la fois au niveau du brin d'ADN qui sert de matrice et du brin en cours d'élongation [89]. Ainsi lorsqu'une boucle se forme au niveau du brin matrice, il n'y aura pas réplication de la séquence de la boucle sur le brin synthétisé et on aura donc une délétion. En revanche, lorsqu'une boucle se forme au niveau du brin en cours de réplication, l'ADN polymérase peut être amenée à répliquer deux fois la même région (insertion par duplication en "tandem"). Elles peuvent aussi résulter d'événements de transposition (voir plus bas) ou d'événements de recombinaison entre des séquences homologues (recombinaison inégale entre chromosome homologue, recombinaison entre des séquences répétées) (figure 1.1.1.2).
- **les duplications de séquences** qui sont un type particulier d'insertion de séquence car elles correspondent à l'insertion d'une séquence identique à une autre présente dans le génome d'origine. Les séquences dupliquées peuvent être collées (duplication en tandem) ou distantes l'une de l'autre. Elles peuvent résulter d'erreur de réplication (voir précédemment) ou d'événements de recombinaison entre des séquences répétées (dont les éléments transposables)
- **les transpositions d'éléments transposables** qui provoquent des insertions et/ou délétions de séquences selon le mode de transposition. En effet les éléments transposables sont des séquences qui sont capables de se déplacer dans le génome de façon autonome selon deux modes de fonctionnement : "copier/coller" ou "couper/coller". Dans le premier cas on aura une duplication de l'élément transposable

qui sera inséré ailleurs dans le génome alors que dans le second cas, on aura une délétion de l'élément transposable dans sa localisation d'origine pour permettre son insertion dans la séquence cible. Les éléments transposables les plus simples sont les séquences d'insertion (IS) qui ne contiennent que les éléments nécessaires à leur transposition.

- **les inversions de séquences** qui peuvent être provoquées par des événements de recombinaison entre des séquences répétées (par exemple entre des éléments transposables) (figure 1.1.1.2).

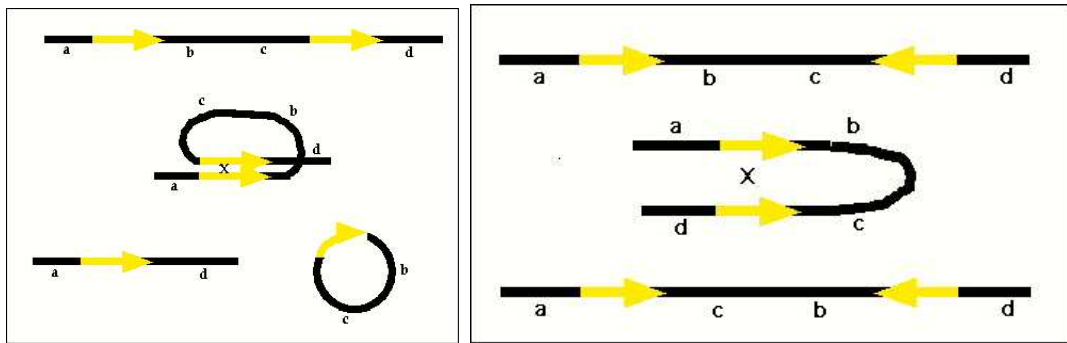


Figure 1.1 – Illustration d'évènements de délétion (figure de gauche) et d'inversion (figure de droite) dus à des réarrangements entre des séquences répétées. Les flèches représentent l'orientation des deux séquences répétées

- **les échanges de séquences** qui ont lieu lors de la reproduction sexuée ou de transfert horizontal d'ADN.

Chez les espèces à reproduction sexuée, les organismes produisent par méiose des cellules de deux types sexuels différents (les gamètes). Chez les organismes diploïdes qui ont deux copies de chaque chromosome, la méiose permet de créer des gamètes haploïdes qui n'ont qu'une copie de chaque chromosome et c'est la fusion de deux gamètes de type sexuel différent (fécondation, syngamie) qui permettra de produire un descendant diploïde qui aura la moitié de l'information génétique provenant d'un parent et la moitié de l'information provenant de l'autre. Lors de la méiose, il peut y avoir des échanges de séquences par recombinaison homologue entre les exemplaires d'un même chromosome (chromosomes homologues). Ce brassage intrachromosomique ou "*crossing-over*" permet de recombinaison les mutations (allèles) présentes sur les deux exemplaires de chaque chromosome. Il y a aussi du brassage interchromosomique car chaque gamète aura une combinaison aléatoire d'exemplaire de chaque chromosome recombinaison. Lors de la fécondation, il y a aussi du brassage génétique car il y a combinaison aléatoire des mutations provenant des deux parents.

Le transfert horizontal d'ADN, par opposition au transfert vertical par descendance, permet le transfert d'ADN entre un donneur et un récepteur. Le transfert horizontal d'ADN est surtout présent chez les procaryotes (archéobactéries et eubactéries) mais quelques cas de transfert de bactéries vers les eucaryotes

(levures [54], plantes [14], d'eucaryotes vers des bactéries (par exemple de plantes vers des bactéries [33, 67]), voir entre eucaryotes ont été observés expérimentalement ou supposés d'après des analyses comparatives de génomes [13, 68, 77, 102]. Le transfert horizontal d'ADN avec des eucaryotes est plus controversé car il est plus difficile à mettre en évidence (souvent uniquement par analyse de génomes) alors que chez les procaryotes, l'existence de transferts horizontaux a été montrée par des expériences et/ou par analyses de génomes [13, 68, 77, 102]. Il existe trois modes principaux de transfert horizontal d'ADN chez les procaryotes [13] :

- la conjugaison qui transfère de l'ADN entre une bactérie donneuse et une bactérie réceptrice grâce à la formation d'un lien physique entre elles (le pilus).
- la transduction qui transfère de l'ADN entre bactéries par l'intermédiaire d'un bactériophage (virus de bactérie)
- la transformation qui permet aux bactéries de récupérer de l'ADN présent dans l'environnement (voir chapitre 3).

Contrairement au sexe qui permet de brasser entièrement deux génomes, seulement une partie du génome du donneur recombine avec le génome du receveur lors d'un transfert horizontal d'ADN. De plus ce transfert est très occasionnel et moins fréquent que le sexe, ce qui fait que l'importance des transferts horizontaux dans l'évolution des eucaryotes (sexués) est encore controversé [68, 77, 102]. Cependant pour les organismes à reproduction asexuée (clonale) comme les procaryotes, on peut retrouver certaines composantes du sexe dans le transfert horizontal [125].

1.1.1.3 Effets des mutations

La plupart des mutations sont délétères ou neutres et seulement une petite fraction des mutations est bénéfique. Ainsi chez la bactérie *Escherichia coli*, le taux de mutation bénéfique a récemment été estimé expérimentalement à environ $5,4 \cdot 10^{-12}$ par base et par génération [116] soit environ 1 à 6% du taux de mutation global estimé expérimentalement par [38, 136].

L'effet d'une mutation dépend :

- **de sa localisation dans le génome.** En effet dans le génome, on distingue trois grands types de séquences : les séquences codantes, les séquences régulatrices de l'expression des gènes et les autres séquences (sans fonction connue). Des mutations dans les deux premières catégories peuvent avoir différents effets : bénéfique, délétère (voire létal) ou neutre alors que des mutations dans la troisième catégorie sont généralement neutres.

Des mutations dans des séquences régulatrices ou des séquences codant pour des protéines régulatrices peuvent modifier l'expression de plusieurs gènes. Par exemple, chez les bactéries, les gènes d'un opéron¹ ont des éléments régulateurs en commun (promoteur, séquences régulatrices et régulateurs transcriptionnels) si

1. un opéron est une unité transcriptionnelle composée de plusieurs gènes sous le contrôle des mêmes séquences régulatrices (dont le promoteur) et des mêmes régulateurs transcriptionnels. Un ARN commun (ARN polycistronique) permet la traduction des protéines correspondant aux gènes de l'opéron

bien que la mutation dans un des ces éléments affectera l'expression de l'ensemble des gènes de l'opéron. De façon plus générale la modification de l'expression ou de l'efficacité d'une protéine régulatrice pourra aussi modifier l'expression des gènes qu'elle régule.

Des mutations dans la séquence codante d'un gène peut dans certains cas modifier l'expression d'autres gènes. C'est le cas des mutations dans les protéines régulatrices comme on vient de le voir mais cela peut aussi être le cas de mutations dans les premiers gènes d'un opéron si elles provoquent un arrêt prématuré de la transcription ou si elles déstabilisent le transcrit commun : on parle d'effet polaire.

- **de l'environnement dans lequel vit l'organisme.** Par exemple si une mutation entraîne la perte de la fonction d'une enzyme nécessaire pour l'utilisation d'un substrat, elle sera délétère si elle utilise ce substrat mais sera neutre voire bénéfique si le substrat est absent dans l'environnement.
- **du nombre de copie de chaque gène :** pour un organisme diploïde, il y a deux exemplaires de chaque chromosome et par conséquent deux copies (allèles) de chaque gène. Une mutation sur un gène peut être récessive si les deux copies du gène doivent être mutées pour avoir un effet ou dominante si une seule mutation suffit. De plus, même quand il n'y a qu'un exemplaire de chaque chromosome, il peut y avoir de la redondance fonctionnelle notamment quand il y a eu des duplications de gènes.
- **des autres variations présentes ailleurs dans le génome.** Contrairement aux interactions entre les différents allèles d'un même gène (dominance/récessivité) qui nécessitent la présence d'au moins deux exemplaires de chaque chromosome, les interactions entre différents gènes ou entre différentes mutations dans une même copie d'un gène peuvent avoir lieu chez les espèces n'ayant qu'un exemplaire de chaque chromosome (organismes haploïdes, par exemple les bactéries). C'est le phénomène d'**épistasie au sens large**² [81, 118].

On parle d'épistasie synergique entre deux mutations quand l'effet des deux mutations est plus important que la somme des effets de chaque mutation et d'épistasie antagoniste quand l'effet est plus faible. Pour une mutation bénéfique, l'épistasie synergique est dite positive car le phénotype est meilleur que prévu (supérieur à la somme des effets de chaque mutation) alors que l'épistasie antagoniste est négative car le phénotype est moins bon que prévu. Pour une mutation délétère, c'est l'inverse. Prenons l'exemple de la biosynthèse d'un produit. S'il n'y a qu'une seule voie conduisant à la production de ce produit alors l'effet d'une mutation délétère sera généralement plus faible s'il y a déjà une autre mutation délétère dans cette voie (épistasie positive antagoniste). Par contre s'il y a plusieurs voies

2. Le terme d'épistasie a été introduit pour la première fois par William Bateson en 1909 et désignait le fait que la mutation d'un gène masquait l'effet de mutations sur d'autres gènes. R.A Fisher en 1918 réutilise ce terme d'épistasie pour faire référence à toute déviation statistique à l'additivité des effets des mutations puis à partir des années 1960 avec la génétique des populations, l'épistasie désigne aussi toute déviation par rapport à un modèle multiplicatif. C'est cette dernière définition utilisée en génétique des populations que j'utiliserai par la suite.

permettant la synthèse de ce produit alors l'effet de mutations délétères tendront à être de plus en plus important (épistasie négative synergique).

L'épistasie est dite signée si elle change "le signe" de l'effet d'une mutation : une mutation peut être délétère ("négative") dans un fond génétique donné mais bénéfique ("positive") dans un autre et vice versa.

- **du type de mutation.** Les mutations ponctuelles de type substitution lorsqu'elles sont localisées dans la séquence codante d'un gène sont classées en trois catégories :
 - **les mutations faux sens ou non synonymes** qui provoquent la substitution d'un acide aminé par un autre dans la protéine correspondant au gène. L'effet de ce changement dépend de l'importance de l'acide aminé modifié dans la structure ou la fonction de la protéine et du nombre de copie du gène (voir précédemment).
 - **les mutations non sens** qui changent un codon en un codon stop et conduit à la production d'une protéine plus courte qui peut être fonctionnelle ou non selon la taille et l'importance de la région délétée. Elles conduisent souvent à une perte de fonction.
 - **les mutations silencieuses ou synonymes** qui n'ont pas de conséquence sur la séquence en acides aminées de la protéine codée par le gène. Elles sont en générale neutres, cependant elles peuvent avoir un effet si elles changent un codon par un autre plus ou moins fréquent. En effet la vitesse de traduction, le taux d'expression de la protéine ainsi que sa conformation peuvent être modifiés selon la fréquence du nouveau codon [72].

Les **insertions/délétions** dans la séquence codante d'un gène peuvent provoquer en aval un décalage du cadre de lecture si la longueur de la séquence délétée ou insérée n'est pas un multiple de trois. Un décalage du cadre de lecture conduit, en aval de l'insertion/délétion, à la modification de la séquence en acides aminés de la protéine et peut même provoquer un arrêt prématuré de la traduction (apparition d'un codon stop). Tout comme une mutation non sens, un décalage du cadre de lecture entraîne souvent une perte de fonction. Si la séquence insérée ou délétée est grande et/ou modifie les séquences régulatrices du gène, on peut avoir une perte ou une modification de l'expression du gène. Si la longueur de l'insertion/ délétion est un multiple de trois et est relativement petite, des acides aminés sont insérés ou délétés dans la protéine et l'effet de la mutation dépend de l'importance de la séquence modifiée dans la structure et la fonction de la protéine.

1.1.2 Evolution de la fréquence des variations au sein d'une population par sélection ou dérive

Les mutations apparaissent aléatoirement et permettent de générer de la diversité au sein d'une population. Cette diversité tend à être réduite soit par l'action de la sélection naturelle qui est le moteur principal de l'adaptation, soit par l'effet du hasard (dérive génétique).

1.1.2.1 La sélection naturelle

La sélection naturelle agit au niveau du phénotype de l'individu qui dépend à la fois de l'environnement et de l'ensemble des mutations présentes dans son génome (son génotype). L'individu qui aura le meilleur succès reproductif dans un environnement donné laissera plus de descendants permettant ainsi de transmettre ses mutations plus fréquemment que les autres individus. Les mutations bénéfiques qui confèrent un avantage reproductif se propageront dans la population tandis que les mutations délétères seront progressivement éliminées car les individus porteurs de ces mutations laisseront moins de descendants. La sélection naturelle ne peut agir que sur les mutations qui sont transmises et par conséquent pour les organismes pluricellulaires, uniquement sur les mutations présentes dans les cellules donnant naissance aux gamètes (cellules de la lignée germinale). L'adaptation d'un individu peut être quantifiée par sa *fitness* (appelée aussi valeur adaptative ou valeur sélective) qui correspond à la mesure du succès reproductif³ d'un génotype dans un environnement donné. Dans le cas de générations non chevauchantes, la *fitness* absolue W_A d'un génotype A correspond au nombre de descendants ayant ce génotype A (N_{Ag}) divisé par le nombre de parents de ce génotype (N_{Ag-1}) : $W_A = N_{Ag}/N_{Ag-1}$.

En général, on compare toujours la performance d'un génotype A par rapport à un génotype B (qui peut servir de référence) et c'est pourquoi, au lieu d'utiliser la *fitness* absolue d'un génotype A, on calcule sa *fitness* relative w qui correspond au ratio des *fitness* absolues des deux génotypes comparés : $w = W_A/W_B$

Pour mesurer l'avantage d'un génotype A par rapport à un génotype B de référence dont on fixe arbitrairement sa *fitness* à 1, on utilise aussi le coefficient de sélection s qui est défini par $w = 1 + s$. Si A a la même *fitness* que la référence alors $s = 0$ et $w = 1$. Si A a un avantage sur le génotype de référence alors $s > 0$ et $w > 1$ par contre si A a un désavantage $-1 < s < 0$ et $0 < w < 1$.

Dans le cas où les générations sont chevauchantes, si on considère une population asexuée (de bactéries par exemple) possédant deux génotypes A et B, on définit le taux de sélection $r = d \ln N_{At}/dt - d \ln N_{Bt}/dt$ avec t qui est un temps, N_{At} et N_{Bt} , la taille (ou la densité) des sous-populations de génotypes A et B respectivement [21]. Si la sélection est densité indépendante et qu'il n'existe pas d'interactions entre les deux génotypes alors $r = m_A - m_B$ avec m le taux de croissance de chaque génotype (paramètre malthusien) et la *fitness* relative (malthusienne) est $w = m_A/m_B$.

Quand une mutation nouvelle apparaît, elle est susceptible de disparaître par effet stochastique d'échantillonnage (voir 1.1.2.2). Si elle arrive à se maintenir dans ce premier régime stochastique et que la taille correspondante de la sous-population portant cette mutation parvient à une taille critique $1/s$, alors s'il n'y a pas d'interférence avec d'autres mutations, cette sous-population va de façon déterministe envahir toute la population [35, 127]. Dans ce cas de figure, la probabilité de fixation⁴ d'une mutation par sélection est proportionnel à s : plus s est grand, plus la probabilité de fixation par sélection est grande. Le temps de fixation par sélection est quant à lui inversement

3. le succès reproductif dépend à la fois du taux de survie et du taux de fécondité du génotype

4. une mutation est fixée quand toute la population possède la mutation

proportionnel à s : plus s est grand, plus le temps de fixation est court. Cependant dans le cas d'une population asexuée, les mutations bénéfiques apparues chez différents clones sont en compétition pour la fixation (interférence clonale) et la probabilité de fixation n'est pas gouvernée uniquement par s contrairement aux populations sexuées où des mutations bénéfiques apparues chez différents individus peuvent se combiner (pour plus de détail voir chapitre 2). Selon les différents paramètres cités ci-dessus, on définit plusieurs régimes de sélection [35, 127]. Pour une population asexuée, dans un régime de sélection forte ($s \gg 1/N$) avec un faible nombre de mutations apparaissant, les mutations se fixent les unes à la suite des autres alors que dans un régime de sélection forte avec un grand nombre de mutations apparaissant, des mutations bénéfiques peuvent apparaître avant que les précédentes mutations ne soient fixées ou perdues.

La sélection est une force évolutive qui tend à optimiser la *fitness* vers un optimum. Cependant plusieurs chemins adaptatifs sont possibles et plusieurs optima peuvent exister : c'est l'optimum le plus proche qui est atteint (optimum local) mais il n'est pas forcément le meilleur (optimum global). Un paysage adaptatif est une représentation des différents chemins adaptatifs possibles pour atteindre les génotypes optimaux (aucune mutation unique dans ce génotype ne permet d'augmenter la *fitness*) [104, 120]. Ce paysage peut être plus ou moins rugueux selon le nombre d'optima possibles et les interactions épistatiques contribuent à la formation de ce paysage [118, 120]. Le passage d'un optimum local à un optimum global par sélection est difficile dans un régime où les mutations se fixent une par une car il nécessite une baisse transitoire de la *fitness*.

1.1.2.2 La dérive génétique

Bien que la sélection combinée aux mutations soit le principal moteur de l'adaptation des espèces à leur environnement, d'autres forces modulent l'évolution. Parmi ces forces on retrouve la dérive génétique concept fondée par Sewall Wright, Ronald Fisher et John Haldane au début du XX siècle et connu aussi sous le nom de modèle de Wright-Fisher [44, 138]. La dérive correspond à des biais d'échantillonnage des mutations qui peuvent faire varier aléatoirement les fréquences des mutations indépendamment de leur avantage et des autres forces évolutives (mutation, sélection naturelle, migration...). Ces variations aléatoires des fréquences peuvent être dues par exemple au fait que, par hasard, certains individus n'ont pas de descendants et d'autres en ont plus. Ainsi par dérive génétique, certaines mutations bénéfiques peuvent être perdues et inversement certaines mutations délétères peuvent se fixer. L'équilibre entre la sélection et la dérive dépend fortement de la taille de la population : les effets de la dérive sont d'autant plus grand que la population est petite. La probabilité de fixation d'une mutation par dérive est de $1/N$ pour une population haploïde de taille N ($1/2N$ pour des diploïdes) et le temps de fixation par dérive est proportionnel à N .

La dérive peut être particulièrement importante lorsque la taille de la population est petite ou lorsqu'elle est brutalement réduite soit à cause d'une extinction massive (épidémie, catastrophe écologique...) soit à cause d'une migration. Lorsqu'une petite partie de la population migre géographiquement, la fréquence des allèles dans la petite population de migrants peut par des biais d'échantillonnage aléatoire ne pas être re-

présentative de la grande population de laquelle les migrants sont issus. On appelle ce phénomène l'effet fondateur. De façon plus générale, on parle de goulot d'étranglement quand la taille de la population est brutalement réduite.

Contrairement à la théorie Darwinienne où la sélection est le principal moteur de l'évolution, la dérive est au coeur de la théorie neutraliste avancée par Kimura [73]. Kimura fait l'hypothèse que la majorité des polymorphismes observés dans les séquences génétiques sont neutres : la plupart sont synonymes ou s'ils sont non synonymes, le changement d'acide aminé est supposé ne pas entraîner de conséquence fonctionnelle sur la protéine et sur la *fitness* de l'individu. Selon Kimura et son disciple Ohta [103], la plupart des mutations ponctuelles que l'on observe ont un coefficient de sélection si faible ($s \ll 1/N$) qu'elles peuvent être considérées comme des allèles neutres soumis à la dérive. La théorie neutraliste minimise le rôle de la sélection à l'échelle de l'évolution moléculaire sans pour autant nier l'importance de la sélection naturelle sur l'évolution phénotypique (à l'échelle de l'individu).

1.2 Apports des expériences d'évolution avec des microorganismes

La biologie de l'évolution n'a longtemps été qu'une science historique où on essaie d'inférer les événements passés inconnus à partir des organismes vivants actuels et quand c'est possible de fossiles. Les expériences d'évolution sur des microorganismes ont permis de suivre en temps réel l'évolution de populations dans des conditions contrôlées et ont été ainsi un apport important pour la compréhension des mécanismes de l'évolution (pour revue voir [15, 41]).

1.2.1 Avantages et inconvénients

Les expériences d'évolution avec les microorganismes présentent de nombreux avantages. En effet les microorganismes ont un temps de génération court qui permet de les faire évoluer sur des centaines voire des milliers de générations sur des échelles de temps relativement courtes (tableau 1.1)

Leur petite taille permet de maintenir de grandes population dans un espace restreint et d'en avoir plusieurs en parallèle (tableau 1.1).

Un autre avantage des microorganismes, est la possibilité de conserver vivants, au congélateur (à $-80^{\circ}C$), les ancêtres et des échantillons de population à différents temps (générations) de l'expérience. Cela permet de les réétudier plus tard en comparaison avec les individus issus de l'évolution (tableau 1.1).

Les conditions dans lesquelles a lieu l'expérience peuvent être contrôlées, ce qui permet de modifier les contraintes environnementales (nutriments limitants, composition du milieu...) et de reproduire les expériences dans les mêmes conditions. De plus on peut tester l'effet de certains paramètres influençant l'évolution comme la taille de la population, le taux de mutation (création de souches mutatrices en inactivant certain gène lié à la réparation des erreurs par exemple), le taux de recombinaison génétique

| | <i>Homo sapiens</i> | <i>Escherichia coli</i> |
|--|---------------------|----------------------------|
| temps pour 10 000 générations | environ 200 000 ans | 1-2 ans |
| espace nécessaire pour 6 milliards d'individus | la Terre | 200 mL |
| information sur les ancêtres | fossiles | conservés vivants (-80 °C) |

Tableau 1.1 – Avantages des microorganismes pour les expériences d'évolution : comparaison entre l'Homme et la bactérie *Escherichia coli*

(nul dans le cas d'une reproduction asexuée, plus ou moins important selon la fréquence de la reproduction sexuée vs asexuée ou l'efficacité de la machinerie de recombinaison homologue). Cependant certaines critiques ont été émises sur l'application des résultats de ces expériences à l'environnement naturel car les conditions de laboratoire peuvent être considérées comme non naturelles et trop simplifiées par rapport aux conditions environnementales naturelles. Ces critiques sont justifiées mais étudier des systèmes simples permet de mieux comprendre les composantes élémentaires de l'évolution en plus de tester les différentes théories [15, 131]. Les expériences peuvent être complexifiées pour donner une image plus écologique, mais paradoxalement, en augmentant les paramètres, il est plus difficile d'en tirer des principes généraux applicables à d'autres conditions et la reproductibilité des expériences est moins évidente.

1.2.2 Principe

Une population ancestrale (en générale, issue d'un seul clone) est mise en culture dans un environnement donné avec souvent comme pression de sélection un nutriment limitant ou des conditions environnementales très défavorables. Deux méthodes de culture peuvent être utilisées pour propager les microorganismes : la dilution en série ou la technique de culture en continu. Dans le cas de la dilution en série, la population estensemencée dans du milieu de culture liquide. La population va croître jusqu'à ce que le milieu s'épuise (plus assez de substrat limitant). Une fraction de la population est ensuite diluée dans du milieu frais pour permettre à la population de croître à nouveau. Cette dilution est répétée périodiquement, d'où le nom de dilution en série. L'expérience modèle d'évolution par dilution en série est celle menée par R. Lenski depuis 1988 et qui est toujours en cours (> 55 000 générations en juillet 2012) ([83], myxo.css.msu.edu/ecoli/). Il a mis en culture en parallèle 12 populations d'*E. coli* dans du milieu minimum limité en glucose. Chaque jour, 1 % de la population est transféré dans du milieu frais permettant ainsi aux bactéries de croître sur environ 6,6 générations avant d'entrer en phase stationnaire jusqu'à la prochaine dilution qui relancera la croissance et ainsi de suite. La taille de la population efficace est de l'ordre de 10^7 .

Dans le cas des techniques de culture en continu de type chemostat (voir section 4.1), le milieu de culture est renouvelé grâce à l'apport en continu de milieu frais et le trop plein de milieu de culture est éliminé afin de maintenir un volume de culture

constant. Contrairement à la dilution en série où les bactéries passent par toutes les phases de croissance (phase de latence, croissance exponentielle, phase stationnaire), dans un chemostat les bactéries sont maintenues dans un état de croissance donné. Pour une revue d'expériences en chemostat : voir la revue de Ferenci [43].

Les ancêtres et leurs descendants peuvent être comparés à la fois au niveau de leur phénotype et au niveau de leur génotype. Pour les comparaisons phénotypiques, on utilise souvent des expériences de compétition afin de mesurer la *fitness* relative des descendants par rapport à l'ancêtre ou d'un descendant par rapport à un autre (voir 1.1.2.1). Les deux clones à comparer sont d'abord cultivés indépendamment dans les mêmes conditions que l'expérience d'évolution. Avant de les diluer dans du milieu frais pour relancer la croissance, les deux populations sont mélangées (généralement avec un ratio 1 pour 1). Le ratio initial des deux populations est estimé en étalant sur milieu solide juste après dilution et le ratio final après un cycle de dilution en série. Pour pouvoir distinguer les deux sous populations, une des deux doit posséder un marqueur génétique qui est sélectivement neutre (par exemple dans l'expérience de Lenski, une des souches est modifiée génétiquement pour former des colonies rouges). Le changement du ratio des deux sous populations permet d'estimer la *fitness* relative. Dans le cas où une des deux sous populations (ou les deux) décroît au cours de la compétition comme ce peut être le cas en chemostat, la *fitness* relative n'a pas de sens et il faut plutôt utiliser le taux de sélection r qui correspond à la différence des *fitness* des deux sous populations.

Les comparaisons phénotypiques peuvent aussi être très ciblées si on recherche un phénotype particulier (capacité de métaboliser un sucre,...).

Les comparaisons génotypiques sont de plus en plus utilisées dans les expériences d'évolution. Auparavant les séquençages de génome entier étaient relativement coûteux et les analyses génétiques étaient souvent ciblées d'après le phénotype que l'on voulait améliorer. Par exemple si on ciblait une adaptation à l'utilisation d'un substrat, l'analyse était souvent limitée à la recherche de mutations dans cette voie métabolique. Même si cette approche permettait d'identifier des mutations, ces dernières n'expliquaient souvent pas entièrement les changements de *fitness* car des mutations non détectées par cette méthode étaient présentes ailleurs dans le génome. Par la suite les séquençages de génome entier se sont généralisés. Avec la baisse des coûts de séquençage et l'amélioration des outils bioinformatiques, les séquençages de génome entier sont devenus la norme et ne se font plus uniquement sur des clones mais aussi sur des populations [5, 8, 9, 36, 56, 130].

1.2.3 Principales informations apportées par les expériences d'évolution

Les expériences d'évolution ont apporté une multitude d'informations sur le fonctionnement de l'évolution en permettant de tester directement différentes théories [15, 41]. Ici je me focaliserais sur certains aspects clés.

1.2.3.1 Estimation du taux de mutation et distribution des effets des mutations bénéfiques

Plusieurs stratégies ont été adoptées pour déterminer le taux de mutation dans les populations de bactéries. Une première stratégie (test de fluctuation) est d'ensemencer un petit nombre de bactéries sans mutant (environ 500 dans [66]) dans du milieu riche et de laisser croître sur environ 20 à 30 génération (taille de la population finale est de $2,8 \cdot 10^8$ dans [66]). Les cultures sont ensuite étalées sur un milieu solide sélectif (par exemple avec de l'acide nalidixique [9,66]) où seul des mutants ayant acquis une résistance à l'agent sélectif peuvent pousser. D'autres variantes de cette méthode sont aussi utilisées pour déterminer le taux de mutation [38] ainsi que les expériences d'accumulation de mutation par dérive (voir principe de ces expériences plus bas) [37,71,86,141]. Avec l'essor des séquençages de génome, le taux de mutation peut aussi être déterminé à partir du nombre de mutations synonymes présent chez des clones issus d'expériences d'évolution [9,136]. Cependant les mutations synonymes ne sont pas forcément neutres à cause du biais d'usage des codons [39]. Le taux de mutation peut aussi être inféré avec des méthodes de génomiques comparatives [101]. Chez *Escherichia coli*, le taux de mutation ponctuelle par base et par génération est estimé entre $1,5 - 4,5 \cdot 10^{-11}$ et $5,4 \cdot 10^{-10}$ [38,101,136] et les variations entre ces estimations sont probablement dues aux différentes hypothèses faites dans chaque méthode de détermination. Cependant au cours d'expérience d'évolution, le taux de mutation peut devenir 10 à 100 fois plus élevé au cours de l'adaptation car des mutateurs (mutés dans des gènes de réparation) peuvent apparaître et se maintenir [5,9,108].

Comme l'adaptation dépend de l'apparition de mutations bénéfiques, le taux de mutation bénéfique et la distribution des effets des mutations bénéfiques sont des paramètres clés. Au cours d'une évolution, ce taux n'est pas constant, les mutations d'effet important sont généralement rares et fixées en premier. Les premières estimations du taux de mutation bénéfique à partir d'expérience d'évolution de bactéries (*Escherichia coli* et *Pseudomonas fluorescens*) sont autour de 10^{-8} et 10^{-9} par génome et par génération [48,62,66,126]. Plus récemment, ce taux de mutation bénéfique a été estimé expérimentalement à $2 \cdot 10^{-5}$ par génome et par génération [116] soit un taux plus de mille fois plus élevé que les estimations précédentes. Pourquoi tant de différences entre ces estimations ? Une première raison peut être la spécificité de chaque expérience. Un autre paramètre peut être la taille de la population. Dans les premières estimations, la taille de la population est grande ($>10^7$) alors que dans cette dernière expérience, la taille de population a été choisie plus petite (10^4) afin de minimiser l'interférence clonale qui purge une grande partie des mutations bénéfiques à faibles effets (l'effet de l'interférence clonale a été montré dans [116] en prenant une taille de population de 10^7 pour laquelle on retrouve un taux de $2 \cdot 10^{-8}$). Cependant les effets de la dérive et de l'interférence clonale avaient été pris en compte dans les estimations de [48,126] mais leur modèle avait des hypothèses trop restrictives sur la gamme de valeurs de ce taux de mutation bénéfique. En reprenant leur données avec un nouveau modèle moins restrictif [127], le taux de mutation bénéfique est estimé à $5,7 \cdot 10^{-5}$, valeur en accord

avec [116]. Ces dernières estimations du taux de mutation bénéfique [116, 127] correspondent à environ 1 à 6% du taux de mutation global estimé expérimentalement ($2,5 \cdot 10^{-3}$ par génome et par génération [38] et $4 \cdot 10^{-4}$ par génome et par génération [136]). Ainsi les mutations bénéfiques pourraient représenter jusqu'à environ 10% des mutations et finalement le taux de mutation ne serait pas aussi bas que prédits précédemment pouvant ainsi changer la vision de l'adaptation par un régime de sélection avec de un faible taux de mutation bénéfique. La distribution des effets des mutations bénéfiques semble sur quelques cas ressembler à une distribution exponentielle avec un excès de mutations à faibles effets [48, 62, 66, 116, 126] comme prédit [44, 104].

1.2.3.2 Comparaison des effets de la dérive et de la sélection

Pour tester les effets de la dérive, une façon de faire est de soumettre la population à des séries de goulots d'étranglement extrêmes. Un clone est ensemencé dans du milieu liquide et au lieu de diluer une fraction de la population dans du milieu, la population est étalée sur milieu solide afin de sélectionner un clone qui va être ensemencé dans du milieu pour relancer la croissance et on répète ces actions autant de fois que souhaité. Le choix d'un clone à chaque cycle au lieu d'une fraction de la population permet d'éliminer la diversité à chaque cycle et comme ce choix est aléatoire, la dérive et les mutations aléatoires sont les principales forces gouvernant ce type d'évolution. Ce type d'expérience a été mené avec des virus [17, 32], des bactéries [71, 86] ainsi que des levures [37, 141]. Dans toutes ces études, la *fitness* décroît à cause de la fixation de mutations délétères. En comparant les différentes répliques d'une expérience, le nombre de mutations fixées tend à augmenter avec le nombre de cycle de goulots d'étranglements mais la variance entre les différentes lignées tend elle aussi à augmenter avec le nombre de cycle car par chance certaines lignées fixent plus ou moins de mutations. Ce type d'expérience mettent aussi en évidence le Cliquet de Muller (voir chapitre 2).

Dans les populations de grandes tailles ($> 10^6$), la sélection est la force prédominante. En général, la *fitness* moyenne de la population augmente rapidement au début puis ralenti brutalement et continue à augmenter à un taux nettement plus faible. Ainsi dans l'expérience d'évolution de Lenski, l'augmentation en *fitness* dans les 5000 premières générations est environ dix fois plus grande qu'entre 15 000 et 20 000 générations. L'augmentation rapide de la *fitness* au début est généralement due à l'accumulation de mutations bénéfiques à large effet à cause de l'interférence clonale qui est forte. Lorsque le nombre de mutations bénéfiques disponibles ainsi que leurs effets moyens diminuent, des mutations à plus faibles effets, qui mettent plus de temps à se répandre, peuvent se fixer grâce à la diminution de l'interférence clonale et on observe ainsi une augmentation moins rapide de la *fitness*. La diminution de la vitesse d'adaptation peut aussi être liée à la présence d'épistasie négative entre les mutations bénéfiques [22, 70]. Bien qu'on retrouve souvent des évolutions similaires entre les répliques d'une même expérience d'évolution [9, 25, 27, 28, 56, 105, 130, 132], on observe aussi des évolutions différentes. Ces divergences peuvent parfois être expliquées par la présence de dérive. En effet bien que la taille des populations dans les expériences d'évolution soit généralement grande et que par conséquent on est dans un régime de sélection forte et de

dérive faible, certaines mutations bénéfiques rares peuvent par chance se fixer modifiant ainsi la trajectoire de l'évolution et les phénomènes d'épistasie peuvent renforcer cette divergence. Les mutations rares ou à faible effet sont les plus soumises à ces effets de dérive bien qu'on peut imaginer que si on a deux mutations de même effet (qu'il soit large ou faible), une des deux mutations peut par chance se fixer avant l'autre.

1.2.3.3 Diversité

Les expériences d'évolution se sont aussi intéressées à la création et au maintien de la diversité dans les populations ainsi qu'aux interactions entre les sous populations. Même dans les environnements les plus simples (milieu homogène avec une seule source de chaque nutriment), de la diversité apparaît et se maintient. Par exemple, une sous population spécialisée dans l'utilisation de déchets métaboliques peut se maintenir dans un chemostat : des évolutions répétées d' *Escherichia coli* en chemostat limité en glucose [132] montrent la coexistence d'une sous population majoritaire utilisant plus efficacement le glucose et d'une sous population minoritaire capable d'utiliser comme source secondaire l'acétate produit par la sous population majoritaire [132]. Cette sous population utilisant l'acétate a une *fitness* plus élevée quand elle est minoritaire car elle dépend de la quantité d'acétate produit par la sous population la plus efficace dans l'utilisation du glucose. D'autres expériences en milieu simple ont aussi montré d'autres types de maintien de sous population qui sont aussi dépendant de la fréquence des deux sous populations [80]. Des expériences d'évolution ont aussi montré que le degré de diversité augmente avec la complexification du milieu que ce soit au niveau de l'hétérogénéité spatiale [123] ou du type de ressources.

1.2.3.4 Interaction entre espèces

La dynamique des interactions entre espèces a aussi été étudiée à l'aide d'expérience de coévolution de phage avec leur hôte bactérien [108, 113]. Ces expériences ont montré que lorsque la bactérie *Pseudomonas fluorescens* évolue avec son bactériophage, son taux de mutation est 10 à 100 fois plus élevé que lorsqu'elle évolue séparément [108]. Il y a ainsi une sélection pour l'évolvabilité avec l'apparition de mutateurs qui sont mutés dans des gènes de réparation et qui peuvent ainsi avoir un taux de mutations plus élevé permettant une réponse plus rapide. De même le bactériophage évolue plus rapidement lorsqu'il coévolue avec son hôte que lorsqu'il est mis en présence d'un hôte ayant toujours le même génotype [113].

1.2.3.5 Concept de macroévolution

Enfin certaines expériences d'évolution [11, 40] ont mis en avant des concepts de macroévolution comme la contingence et les équilibres ponctués [51, 52]. La théorie de la contingence historique postule que plusieurs chemins évolutifs sont possibles mais qu'ils dépendent des événements passés qui ont eu lieu. Ainsi dans l'expérience de Lenski [11], une population a acquis une mutation lui permettant de métaboliser le citrate mais seulement après 30 000 générations. L'apparition de cette mutation est

contingente à l'histoire de cette population car sa probabilité d'apparition est plus élevée dans les clones tardifs (après 20 000 générations).

La théorie des équilibres ponctués postule que l'évolution comprend de longues périodes d'équilibre, ou quasi-équilibre, ponctuées de brèves périodes de changements importants. Dans l'expérience de Lenski, le début de l'évolution (les 3000 premières générations) est ponctué par la sélection de rares mutations bénéfiques qui envahissent successivement la population [40].

1.2.4 Perspectives

Avec la baisse des coûts de séquençage de génomes, la plupart des expériences d'évolution séquençent des génomes de clones de la fin (voir des clones intermédiaires) mais peu de populations ont été séquençées. Cependant le développement des séquençage de population [8] à différents temps de l'évolution permettraient de mieux étudier la diversité et la dynamique des sous populations au cours de l'évolution.

Enfin la biologie des systèmes permet potentiellement une approche plus intégrative et complémentaire dans l'étude des expériences d'évolution afin de mieux comprendre l'impact des mutations sur les réseaux de régulations.

Chapitre 2

Apport évolutif du sexe et des transferts horizontaux

Il existe une multitude de stratégies reproductives qui sont classées en deux grandes catégories : reproduction sexuée et reproduction asexuée.

La reproduction sexuée donne des descendants issus de la fusion des génomes de deux gamètes produits par méiose (pour plus de détails voir section 1.1.1.2). Lors de la méiose, le brassage intrachromosomique permet de recombinaison les allèles présents sur les deux exemplaires de chaque chromosome alors que le brassage interchromosomique permet à chaque gamète d'avoir une combinaison aléatoire d'exemplaire de chaque chromosome recombiné ou non. Enfin lors de la fécondation (fusion de deux gamètes), il y a combinaison aléatoire des allèles provenant des deux parents.

La reproduction asexuée comporte différents modes de reproductions. Les bactéries se reproduisent par mitose pour donner des descendants qui sont des clones (reproduction clonale) et bien que les transferts horizontaux permettent occasionnellement des recombinaisons entre allèles, ils ne constituent pas un mode de reproduction sexué au sens strict (voir section 1.1.1.2). Chez les organismes pluricellulaires, la reproduction asexuée peut être clonale ou non clonale. La reproduction est non clonale si les descendants sont issus d'un gamète non fécondé produit par méiose car le brassage génétique pendant la méiose permet de donner des descendants génétiquement différents. Cependant comme il n'y a pas de fusion de gamètes (fécondation), il ne s'agit pas d'une reproduction sexuée.

La reproduction sexuée est très répandue chez les organismes pluricellulaires (seulement 0,1% des animaux et 1% des plantes sont asexués [106]) bien que le sexe ait un certain coût : coût de la méiose, de la recombinaison, de la production des mâles, de la recherche d'un partenaire ... [82, 93, 107]. Il a été argumenté par Maynard-Smith [93] que la production des mâles conduit à un facteur deux en coût du sexe (voir figure 2.1).

Le coût de cette production des mâles (voir figure 2.1) suppose que les femelles

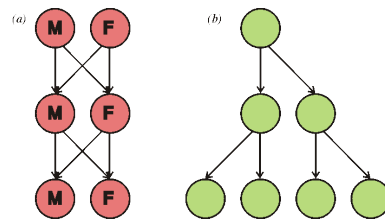


Figure 2.1 – Facteur 2 et coût du sexe. Si on prend l'exemple d'une espèce avec deux femelles, une se reproduisant de manière asexuée (b) et une se reproduisant de manière sexuée avec un mâle (a). Chaque femelle a deux descendants : deux femelles pour la reproduction asexuée, un mâle (M) et une femelle (F) pour la reproduction sexuée. A la génération suivante, dans le cas de la reproduction asexuée, les deux femelles pourront donner chacune deux filles alors que dans le cas de la reproduction sexuée, le mâle et la femelle se reproduiront ensemble pour donner deux descendants. On voit donc que dans le cas de la reproduction asexuée, le nombre de descendants double à chaque génération alors que dans le cas de la reproduction sexuée, il reste constant. Pour que le sexe se maintienne face à l'invasion des individus à reproduction clonale, il faut que l'avantage du sexe soit au moins deux fois plus grand que le coût de production des mâles.

investissent plus dans la production des descendants (gamète femelle plus gros qui fournit les réserves énergétiques, mâles ne s'occupant pas de la progéniture). Enfin dans le cas des eucaryotes unicellulaires où les gamètes sont de même taille (isogamie), les deux partenaires investissent équitablement dans la production de la descendance et le facteur 2 associé dans le coût de production des mâles est moindre (c'est pourquoi les types sexuels ne sont pas appelés mâle et femelle).

La méiose est plus coûteuse que la mitose et nécessite souvent plus de temps : par exemple la levure *Saccharomyces cerevisiae* se divise asexuellement par mitose en 90 min environ alors que l'induction de la méiose peut durer jusqu'à plusieurs jours. La méiose est plus longue car deux divisions sont nécessaires au lieu d'une pour la mitose et des contrôles cellulaires s'assurent du bon appariement des chromosomes pour la recombinaison homologe et la bonne ségrégation des chromosomes. La recherche d'un gamète ou d'un partenaire de sexe opposé est aussi coûteuse en temps et en énergie et un individu sexué prend en plus le risque de se reproduire avec un individu porteur de mutations délétères. La reproduction sexuée peut ainsi produire des descendants ayant une *fitness* plus faible car la recombinaison peut défaire les combinaisons d'allèles bénéfiques et la fécondation peut avoir lieu entre des gamètes porteurs de combinaisons d'allèles plus ou moins bénéfiques.

Il existe des espèces possédant à la fois une reproduction sexuée et asexuée, ce qui suggère que chaque mode a ses avantages mais à différents temps. Quels sont les bénéfices du sexe qui expliqueraient que le sexe se soit maintenu au cours de l'évolution ? Les bénéfices peuvent être directs (avantage pour l'individu) ou indirect (avantage

pour la population avec le danger d'arguments de sélection de groupe). Les bénéfices directs donnent un avantage immédiat à la descendance (avantage soumis directement à la sélection naturelle). Ces avantages directs de la reproduction sexuée sont souvent spécifiques à l'espèce : par exemple soin apportés par les deux parents à la progéniture, oeuf qui est une forme de survie aux conditions environnementales défavorables ... Les bénéfices indirects, qui sont souvent cités pour le sexe, sont liés aux brassages génétiques qui conduisent à une augmentation de la diversité au sein de la population. Plusieurs théories ont été avancées pour expliquer l'avantage du sexe. La plupart avance des avantages indirects. Les trois principales sont la théorie de Fisher-Müller énoncée en 1930 par R.A Fisher [44] et reprise en 1932 par Müller [99], la théorie du cliquet de Müller [99, 100] et la théorie de la Reine Rouge [133].

2.1 Théorie de Fisher-Müller

R.A Fisher en 1930 [44] a avancé l'hypothèse que l'avantage du sexe est de permettre de combiner chez un seul individu des mutations apparues indépendamment chez différents individus de la population et de permettre ainsi une réponse plus rapide à la sélection naturelle. En effet dans une population à reproduction sexuée, il y a à chaque génération des brassages de génomes entre les individus de la population, si bien qu'à chaque génération un individu peut cumuler des mutations bénéfiques provenant de ses deux parents, ce qui permet de construire plus rapidement le génotype le mieux adapté. Müller a repris cette idée en développant un modèle basé sur la compétition entre mutations. Dans une lignée clonale, les mutations apparaissent les unes à la suite des autres. Si deux mutations bénéfiques apparaissent chez deux individus différents de la population, elles vont entrer en compétition pour la fixation et dans le cas d'un régime de sélection forte/faible nombre de mutations, la mutation la plus bénéfique sera sélectionnée tandis que l'autre mutation disparaîtra. Ce phénomène de compétition entre les mutations présentes chez différents individus de la population est appelé *interférence clonale* [99].

L'avantage du sexe avec cette théorie est très dépendant de la taille de la population, du taux de mutation bénéfique et de l'avantage de ces mutations. En effet l'interférence clonale et l'effet du sexe dépendent du nombre de mutations apparaissant à chaque génération et de leur avantage sélectif. Si le nombre de mutations apparaissant est faible (taux de mutation faible ou taille de population faible) ou que les mutations ont un fort avantage sélectif, peu de mutations seront en compétition et/ou elles se fixeront rapidement¹. Dans ce cas, on s'attend à ce que le sexe ait peu d'effet sur l'évolution car la diversité est faible. De la même manière, si la taille de la population ou le taux de mutation sont très grands, plusieurs mutations bénéfiques peuvent apparaître chez un même individu diminuant ainsi l'interférence clonale et l'effet bénéfique du sexe.

L'avantage du sexe avancé par cette théorie est donc présent dans un régime intermédiaire aux deux évoqués précédemment. Cependant cette théorie ne prend en compte

1. le temps de fixation d'une mutation par sélection est proportionnel à $1/s$ et le temps de fixation par dérive est proportionnel à la taille de la population N

que l'interférence clonale entre les mutations bénéfiques et ne considère pas les phénomènes d'épistasie entre les mutations. L'effet de la présence de mutations délétères sur la fixation des mutations et le coût de la dissociation des combinaisons bénéfiques par le sexe ne sont pas considérés dans cette théorie.

2.2 Cliquet de Müller

Le Cliquet de Müller est une théorie qui prend en compte l'effet des mutations délétères alors que la théorie de Fisher-Müller dépend uniquement de l'apparition de mutations bénéfiques. Comme les mutations délétères sont beaucoup plus nombreuses que les mutations bénéfiques [42], cette théorie avance que des mutations délétères finissent inexorablement par apparaître chez les individus les mieux adaptés. Comme chez les populations asexuées, les descendants ont au moins le même nombre de mutations que leur parents (la probabilité d'une réversion est quasiment nulle), les individus ne peuvent qu'accumuler des mutations. Dans les populations sexuées, les mutations délétères peuvent être éliminées par recombinaison avec des individus non porteur de ces mutations et le sexe permet ainsi de reconstituer les meilleurs génotypes. Cependant cette théorie est très dépendante du taux de mutation délétère, de la taille de la population et des effets des mutations. Si le taux de mutations est élevé (en particulier celui des mutations légèrement délétères), une mutation bénéfique peut apparaître chez un individu porteur d'une mutation légèrement délétère. Si l'avantage global des deux mutations est suffisant, la mutation délétère peut ainsi se fixer avec une mutation bénéfique (on parle d'"auto-stop" ou "hitch-hiking"). Dans ce cas de figure, l'avantage du sexe peut être grand et permet de purger les mutations délétères.

Quand la taille de la population est très grande ou que le taux de mutation est faible, le régime de sélection prédomine sur la dérive et la probabilité de fixation d'une mutation délétère est négligeable. Dans ce cas, le sexe a peu d'impact sur l'élimination des mutations délétères.

Le cliquet de Müller dépend aussi du coefficient de sélection s des mutations et des interactions épistatiques entre les mutations. Plus le désavantage d'une mutation délétère est grand, plus vite elle sera éliminée par sélection et sa probabilité de fixation sera d'autant plus faible. De même s'il y a de l'épistasie synergique entre les mutations délétères, le désavantage augmente et la probabilité d'accumuler des mutations délétères devient plus faible.

En conclusion cette théorie s'applique surtout pour des populations de petite taille où la dérive prédomine et présuppose que le taux de mutation délétère soit suffisamment élevé pour que des mutations délétères puissent se fixer par dérive ou ne pas être contre-sélectionnées.

2.3 Théorie de la Reine Rouge

La théorie de la Reine Rouge², proposée par Van Valen en 1973 [133] et reprise par Bell en 1982 [10], rejoint en partie la théorie de Fisher-Müller car toutes les deux avancent que le sexe permet de combiner rapidement des mutations bénéfiques et de s'adapter ainsi plus rapidement. Cependant l'hypothèse de la Reine rouge intègre les facteurs environnementaux et les interactions entre espèces (les co-évolutions).

L'exemple classique pour illustrer la théorie de la Reine Rouge est l'exemple de la co-évolution entre hôtes et parasites. Le sexe permet de recombinaison les allèles et les parents peuvent donner naissance à des génotypes rares que le parasite n'est pas capable d'infecter. Le génotype rare échappant au parasite va être avantageux et va se répandre dans la population. Le parasite n'étant plus adapté à son hôte, il va lui aussi être sélectionné pour s'adapter à son hôte et on va donc avoir ainsi une course poursuite sans fin entre le parasite et l'hôte afin que chacun puisse s'adapter au changement de l'autre. Le sexe apparaît donc comme un moyen de maintenir une certaine *fitness*.

2.4 Autres théories

2.4.1 Hypothèse mutationnelle déterministique

Cette hypothèse formulée initialement par Kondrashov en 1984 [75, 76] reprend la théorie du cliquet de Müller en ajoutant l'hypothèse d'un effet synergique entre les mutations délétères. Le sexe permettrait de créer des individus porteurs d'un nombre différent de mutations délétères. Comme il existerait des interactions synergiques entre les mutations, les individus porteurs d'un plus grand nombre de mutations délétères disparaîtraient en éliminant avec eux plusieurs mutations délétères, réduisant ainsi le nombre de mutations délétères dans la population. Cependant cette hypothèse repose sur des conditions très restrictives et pas forcément réalistes : en plus de supposer que toutes les mutations délétères agissent de façon synergique, le taux de mutation délétère par génome et par génération doit être supérieur à 1, ce qui semble très supérieur aux taux de mutations observés (de l'ordre de 10^{-3} par génome et par génération).

2.4.2 "Rubby in the rubbish"

Cette théorie, proposée par Peck en 1994 [115], reprend la théorie de Fisher-Müller et du Cliquet de Müller pour prendre en compte à la fois les mutations bénéfiques et les mutations délétères. Cette théorie suppose comme dans le Cliquet de Müller que le nombre de mutations délétères dans les populations asexuées est élevé mais ajoute que les individus ayant une *fitness* plus faible que la moyenne sont génétiquement condamnés à mourir. Cependant si une mutation bénéfique arrive chez ces individus et que son avantage est suffisant pour dépasser le coût des mutations délétères alors cette

2. la Reine Rouge fait référence au roman de Lewis Carroll (*A travers le miroir*) et plus particulièrement au moment où Alice, qui est au pays des merveilles, doit, sur les conseils de la Reine Rouge, courir sans cesse afin de rester à la même place.

mutation bénéfique peut se fixer entraînant la fixation des mutations délétères présentes chez cet individu asexué. Dans une population sexuée, quelque soit le fond génétique dans lequel est apparue la mutation bénéfique, elle peut se fixer indépendamment des autres mutations évitant ainsi la fixation de mutations délétères par "auto-stop".

2.4.3 "Tangled bank"

Cette théorie repose sur un avantage du sexe dans un environnement spatialement hétérogène [74]. Le sexe permettrait de créer une plus grande diversité parmi les descendants et laisseraient ainsi plus de chance pour ses descendants d'explorer de nouvelles "niches" écologiques, diminuant ainsi la compétition entre les descendants.

2.4.4 Modèle de la loterie

Ce modèle [92,137] propose de comparer à la loterie l'adaptation à un environnement instable. Les individus asexués produisent des clones identiques et sont comparés à des acheteurs de plusieurs billets de loterie identiques. Les individus sexués peuvent donner des descendants ayant des génotypes différents et sont donc comparés à des acheteurs de plusieurs billets différents. Ainsi les individus sexués ont plus de chance de produire un descendant ayant la combinaison gagnante pour le nouvel environnement. Cependant dans ce modèle, le sexe est avantageux s'il existe une certaine hétérogénéité spatiale.

2.5 Test expérimental de l'apport évolutif du sexe et des transferts horizontaux

Les différentes théories présentées précédemment avancent des idées intéressantes mais les conditions d'applications de ces modèles sont souvent restrictives et des données expérimentales sont nécessaires pour comprendre comment ces modèles interagissent et s'appliquent à l'évolution du sexe et par extrapolation aux transferts horizontaux entre individus de même espèce.

2.5.1 Expériences In vitro

Une des composantes du sexe est le brassage intrachromosomique qui, lors de la méiose, permet de recombinaison les mutations présentes sur les chromosomes d'une même paire. L'étude in vitro de Stemmer [128] a permis de montrer que la recombinaison entre des variants de séquences d'ADN ("DNA shuffling") permet d'accélérer l'évolution in vitro de protéines. Dans cette expérience, le gène codant pour la protéine est amplifié par PCR avec une ADN polymérase commettant des erreurs aléatoires dans la séquence d'ADN. Les protéines exprimées à partir de ces gènes mutés sont sélectionnées afin de ne retenir que les gènes qui ont des mutations bénéfiques et un nouveau cycle d'amplification avec erreur est réalisé et ainsi de suite. Pour tester l'effet de la recombinaison sur ces cycles de mutagenèse aléatoires/sélection, une étape de recombinaison est rajoutée

après l'étape de mutagenèse : les différentes copies mutées du gènes sont fragmentées et les différents fragments sont réassemblés par une pseudo-PCR qui par recombinaison homologue permet de reconstituer la séquence entière. Les mutations présentes sur différentes copies du gène peuvent ainsi à chaque cycle se combiner.

A l'échelle de la séquence d'ADN, cette évolution est la combinaison de mutations à laquelle il aurait été virtuellement impossible d'arriver, par cumuls successifs (uniquement sériel) de mutations sur une même séquence, à cause de la taille de la "bibliothèque" de séquences qui serait nécessaire. En d'autre terme le "DNA shuffling" in-vitro permet par des mécanismes non pas d'accumulation sérielle (clonale) mais d'échanges (parallélisés) de fragments entre classes, d'obtenir une "accélération d'évolution".

Cette technique du "DNA-shuffling" est utilisée dans le domaine des biotechnologies et de l'évolution in-vitro pour produire des variants d'intérêt (par exemple une protéine particulière) [12, 98, 128] et a permis d'améliorer considérablement la vitesse avec laquelle on arrive à obtenir des variants intéressants.

2.5.2 Expérience d'évolution avec des microorganismes

Différentes expériences d'évolution avec des microorganismes ont aussi cherché à tester l'apport évolutif du sexe ou des transferts horizontaux intraspécifiques. Cet apport a pu être testé chez différents microorganismes (virus/bactériophages, bactéries, levures, algues unicellulaires...).

2.5.2.1 Virus

Chez les virus et les bactériophages, lorsqu'il y a de multiples infections, il peut y avoir recombinaison entre les génomes des virus/bactériophages qui coinfectent [18]. Plusieurs expériences ont montré que les multiples infections accélèrent les gains en *fitness* et semblent aller dans le sens de la théorie de Fisher-Müller [90] ou du cliquet de Müller dans le cas où les virus utilisés avaient accumulés des mutations délétères par des séries de goulots d'étranglement [17, 19].

2.5.2.2 Bactéries

Quelques expériences d'évolution de bactéries ont testé l'apport du sexe et des transferts horizontaux d'ADN sur l'évolution des bactéries soit en utilisant la conjugaison [26], la transformation naturelle [4, 6, 117] ou la fusion de protoplastes qui est une technique permettant de brasser les génomes de certaines bactéries [29, 114, 142].

Dans une expérience avec des cycles de mutagenèse aléatoire suivie de sélection et de fusion de protoplastes [142], les bactéries *Streptomyces fradiae* sont sélectionnées pour augmenter la production d'un antibiotique et seulement deux cycles sont nécessaires pour observer une nette augmentation alors que vingt cycles classiques de mutagenèse aléatoire/sélection avaient été nécessaires pour obtenir le même résultat. Cependant à cause des contraintes techniques, ce type d'évolution dirigée ne permet un brassage de génomes ("genome shuffling") qu'entre un petit nombre d'individus [29, 114, 142].

Dans le but de tester la théorie de Fisher-Müller, une expérience d'évolution de bactéries avec ou sans conjugaison a été réalisée avec des taux de mutations élevés ou bas. Il a été montré que la souche pouvant recombiner par conjugaison (rec+) s'adaptait plus rapidement que la souche ne pouvant pas recombiner (rec-) [26]. Cet avantage de la souche pouvant recombiner est d'autant plus grand que le taux de mutation est élevé. La dynamique de fixation d'une mutation dans le gène *spoT* a aussi été suivie dans des populations rec- et rec+ avec un taux de mutation élevé. Les mutations dans *spoT* se fixent plus rapidement dans les populations rec+ suggérant que l'interférence clonale est plus élevée dans la population rec-. Comme l'adaptation est plus rapide quand le taux de mutation est plus élevé, ces résultats semblent en accord avec les prédictions de la théorie de Fisher-Müller car l'interférence clonale est plus élevée quand le nombre de mutations en compétition augmentent rendant ainsi la recombinaison plus avantageuse.

Dans une expérience d'évolution avec *Acinetobacter baylyi* [117], une bactérie naturellement compétente pour la transformation, l'effet de la recombinaison et de la diversité sur l'évolution de la résistance à deux antibiotiques a été testé. Pour chaque souche résistante à un antibiotique (la rifampicine ou la gentamicine), une souche déficiente pour la recombinaison (Rec-) est aussi créée en mutant le gène *recA*. L'évolution des souches résistantes à la rifampicine ou à la gentamicine est réalisée dans un milieu possédant alternativement de la rifampicine et de la gentamicine. L'évolution est initiée soit à partir de monoculture d'une souche résistante à un antibiotique (pas de diversité initiale) soit à partir d'une polyculture constituée d'une souche résistante à la rifampicine et d'une souche résistante à la gentamicine (diversité initiale). Pour les souches Rec-, l'obtention de double résistants est inhibé si on initie l'évolution avec un mélange de souches résistantes à différent antibiotique mais pour les souches pouvant recombiner, la fréquence des doubles résistants augmente plus rapidement dans les polycultures. Les transferts horizontaux d'ADN permettent donc une acquisition plus rapide des résistances multiples aux antibiotiques lorsqu'il existe de la diversité dans le type de résistance car l'interférence clonale entre les simples résistants est levée grâce à la combinaison des résistances présentes chez les différentes sous populations.

La comparaison de l'évolution d'une souche compétente et non compétente pour la transformation naturelle montre parfois que l'adaptation est plus rapide dans les lignées compétentes [6, 117] mais parfois aucune différence n'est observée et une baisse de la compétence peut se produire [4]. Une explication possible à la baisse de compétence est que les bactéries sont transformées avec l'ADN qui proviendrait de bactéries mortes, composées surtout de mutants avec de faibles *fitness* (sous-représentation de mutations bénéfiques et sur-représentation de mutations délétères). Une autre explication possible est le transfert irréversible d'allèles causant une perte de fonction de gènes nécessaires pour la transformation. Cependant il est parfois difficile de déterminer si les différences observées sont réellement dues à des transferts horizontaux d'ADN par transformation car la quantité d'ADN présent dans le milieu n'est pas connue [6].

2.5.2.3 Eucaryotes unicellulaires

Les eucaryotes unicellulaires ont aussi été utilisés pour tester l'apport évolutif du sexe car la plupart de ces organismes peuvent se reproduire asexuellement ou sexuellement. Les deux organismes modèles les plus utilisés sont l'algue unicellulaire *Chlamydomonas reinhardtii* et la levure *Saccharomyces cerevisiae*.

Chlamydomonas reinhardtii est une algue unicellulaire haploïde qui est flagellée et qui existe sous deux types sexuels mt(+) et mt(-). Lorsque le milieu est pauvre en azote, deux gamètes haploïdes de types sexuels opposés fusionnent pour donner un zygote diploïde qui est une forme de dormance et qui ne possède pas de flagelle. Quand le milieu redevient riche en azote et qu'il y a de la lumière (conditions nécessaire pour faire de la photosynthèse), le zygote diploïde peut redonner par méiose quatre gamètes haploïdes qui vont se reproduire asexuellement tant que les conditions sont favorables. La possibilité de contrôler le mode de reproduction selon les conditions de cultures rend cet organisme modèle pour l'étude de l'apport évolutif du sexe. L'effet du nombre de cycles de reproduction sexuée a été étudié dans des expériences d'évolution avec *Chlamydomonas reinhardtii* [65] dans un environnement plus ou moins complexe (avec une à trois sources de carbone). Les résultats ne sont pas très évidents à interpréter car ils ne présentent pas une tendance bien nette. Les lignées avec des cycles de reproduction sexuée ont généralement une augmentation en *fitness* plus rapide que les lignées asexuées et particulièrement dans le milieu le plus complexe. L'effet du sexe sur l'adaptation augmente généralement avec le nombre de cycles de sexe et s'accompagne d'une augmentation de la variance de la *fitness*. Cependant bien que le sexe ait généralement un avantage à long terme, le sexe produit généralement un coût à court terme car après chaque cycle de sexe la *fitness* diminue sur quelques génération avant de réaugmenter.

Une autre expérience d'évolution avec cet organisme a été menée dans le but de déterminer quel est l'effet de la taille de la population sur le bénéfice apporté par le sexe [24]. L'évolution est réalisée par dilution en série en faisant varier, entre les séries d'expériences, la taille du goulot d'étranglement à chaque dilution. Dans la population évoluant asexuellement, le taux d'adaptation augmente avec la taille de la population mais reste plus faible que les lignées avec sexe comme le prédisait la théorie de Fisher-Müller. Cependant la différence en *fitness* entre les lignées sexuées et asexuées augmente aussi avec la taille de la population alors que d'après le modèle d'interférence clonale, lorsque les populations sont grandes, plusieurs mutations peuvent apparaître en même temps dans une même lignée diminuant ainsi l'effet du sexe. De plus la théorie du Cluet de Müller prédit plutôt que l'effet du sexe est plus grand sur les petites population car elles sont plus soumises à l'effet de la dérive qui peut mener à l'accumulation de mutations délétères.

La levure *Saccharomyces cerevisiae* est un eucaryote unicellulaire possédant deux types sexuels différents MATa et MAT α . Les formes diploïdes sont MATa/MAT α et se reproduisent de manière asexuée quand le milieu contient suffisamment de nutriments. Quand les conditions sont défavorables, la forme diploïde sporule pour donner par méiose quatre spores haploïdes qui peuvent se reproduire de façon asexuée par mitose

ou de façon sexuée grâce à la fusion de deux spores haploïdes de types sexuels différents. Tout comme *Chlamydomonas reinhardtii*, la reproduction sexuée est induite par un appauvrissement du milieu. Ce genre d'induction conduit à des conditions expérimentales différentes entre les évolutions sexuées et asexuées car l'évolution sexuée correspond en réalité à un remplacement de quelques cycles de reproduction asexuée par des cycles de reproduction sexuée induits par le passage dans un environnement pauvre. Pour contrer ce problème, une délétion de deux gènes de recombinaison (*SPO11* et *SPO13*) conduit la levure diploïde à se reproduire uniquement par mitose (reproduction uniquement asexuée) car elle ne peut pas produire de spores haploïdes par méiose. Ainsi lors du passage dans un milieu pauvre, la souche possédant les gènes *SPO11* et *SPO13* sporule et se reproduit de manière sexuée alors que la souche déficiente pour ces deux gènes se reproduit de manière asexuée. L'effet du sexe a été étudié en utilisant cette méthode dans plusieurs expériences [49, 53]. Dans un milieu peu favorable, l'évolution de la souche sexuée est plus rapide que la souche asexuée et s'accompagne aussi d'une plus grande variance dans la *fitness* comme vu précédemment avec *Chlamydomonas reinhardtii* [65]. L'effet du taux de mutations a aussi été étudié avec ce système d'étude du sexe [53]. Des souches mutatrices sexuées et asexuées ont été créées en délétant le gène de réparation *MSH2* afin d'augmenter le taux de mutation. Dans des conditions de cultures peu favorables, les souches mutatrices asexuées semble évoluer moins rapidement que les souches asexuées non mutatrices alors qu'aucune différence n'est observée entre les souches sexuées mutatrices et non mutatrices dans des conditions favorables. Ces résultats laissent penser que le sexe permet d'éliminer plus efficacement les mutations délétères en dissociant celle-ci des mutations bénéfiques comme le suggère la théorie du Cliquet de Müller. Cependant ces résultats sont peu convaincants à cause des fluctuations importantes de la mesure de *fitness* au sein de chaque condition.

2.5.3 But de la thèse

Dans beaucoup d'étude testant l'apport évolutif du sexe et des transferts horizontaux entre individus de même espèce, les comparaisons entre évolution avec ou sans sexe (ou transfert) sont souvent uniquement phénotypiques (comparaison de *fitness*) et peu d'études ont regardé les différences d'adaptation au niveau génomique. De plus les conditions expérimentales sont souvent différentes entre l'évolution avec ou sans sexe car le choix entre reproduction sexuée et asexuée est déterminé par les conditions environnementales.

Le but de ma thèse est de déterminer quel est l'impact des échanges d'ADN entre bactéries sur l'évolution d'une population issue d'un ancêtre commun. Mon travail de thèse s'est porté plus particulièrement sur les différences d'adaptation au niveau génomique entre des populations évoluant avec ou sans échanges d'ADN dans des conditions identiques afin de voir l'impact du brassage génétique sur l'évolution.

Chapitre 3

Acinetobacter baylyi, une bactérie naturellement compétente

Les bactéries se reproduisent de manière asexuée, c'est à dire qu'une bactérie se divise pour donner deux clones identiques aux mutations près. Cependant il y a occasionnellement des échanges d'ADN entre différents individus de même espèce ou non. Ces échanges d'ADN entre individus sont appelés transferts horizontaux d'ADN par opposition aux transferts verticaux qui se font par descendance. Ils permettent, comme le sexe, de combiner des mutations apparues chez différents individus et constituent ainsi un bénéfice évolutif. De nouveaux gènes (par exemple des gènes de résistance à des antibiotiques) peuvent aussi être transférés horizontalement entre différentes bactéries (différentes souches ou espèces).

Il existe 3 modes de transfert horizontal d'ADN [13] : **la transduction** qui se fait par l'intermédiaire d'un bactériophage (virus de bactéries), **la conjugaison** qui permet le transfert d'ADN entre une bactérie donneuse et une bactérie réceptrice grâce à la formation d'un lien physique entre elles et enfin **la transformation naturelle** qui permet à certaines bactéries de récupérer l'ADN présent dans son environnement et de l'intégrer dans son génome par recombinaison homologue.

Les bactéries capables de transformation naturelles sont dites naturellement compétentes car contrairement à la transformation induite artificiellement en laboratoire, on ne les rend pas compétentes à l'aide de différents traitements (électroporation, traitement au CaCl_2 ...). Cependant cette compétence naturelle dépend de l'état physiologique de la bactérie et elle est induite dans des conditions différentes selon les espèces. Ainsi pour pouvoir déterminer si une bactérie est naturellement compétente, il faut trouver les conditions dans lesquelles elle est transformable et ceci implique donc qu'on minore le nombre d'espèces de bactéries naturellement compétentes. Cependant on estime qu'environ 1 % du nombre total d'espèces de bactéries connues sont compétents et les 70 espèces répertoriées sont distribuées sur l'ensemble des phyla [64, 88]. Les bactéries compétentes les plus étudiées sont *Bacillus subtilis* (gram⁺, non pathogène), *Streptococcus pneumoniae* (gram⁺, pathogène), *Acinetobacter baylyi* (gram⁻, non pa-

thogène), *Neisseria gonorrhoeae*, *Haemophilus influenzae* et *Helicobacter pylori* (gram⁻, pathogènes).

3.1 Machinerie de transformation naturelle chez *Acinetobacter baylyi*

On peut distinguer plusieurs étapes dans la transformation naturelle : l'induction de la compétence, la liaison de l'ADN exogène à la surface de la bactérie, son entrée et sa prise en charge jusqu'au cytoplasme et enfin son intégration dans le génome par recombinaison homologue.

Chez les bactéries gram⁻, la translocation de l'ADN du milieu extérieur vers le cytoplasme nécessite de traverser la membrane externe puis le peptidoglycane et périplasme et enfin la membrane interne. Les machineries de translocation de l'ADN identifiées chez les bactéries compétentes font souvent intervenir des protéines similaires à celles retrouvées dans les pseudopili de sécrétion de type II et les pili de type IV [2,20]. Ces derniers interviennent dans les interactions cellulaires, la formation de biofilm et la "twitching motility", une forme de locomotion des bactéries qui se fait grâce à l'extension et la rétraction des pili. Par similarité avec ces 2 systèmes existants, il a été proposé un modèle de structure en pseudopilus de compétence permettant le transport de l'ADN exogène jusqu'au cytoplasme [2,3,20] (figure 3.2). De plus, chez *Acinetobacter baylyi*, il a été montré que cette translocation de l'ADN nécessite de l'énergie [111].

3.1.1 Effet du type d'ADN exogène sur la transformation naturelle

Chez *Acinetobacter baylyi*, la translocation de l'ADN exogène ne dépend pas de sa séquence. En effet comme le montrent des expériences de compétition d'ADN [87,111], il n'y a pas de distinction entre ADN homologue (provenant d'*Acinetobacter baylyi*) et ADN hétérologue (provenant d'autres espèces de bactéries) contrairement à ce qu'on observe chez *Neisseria gonorrhoeae* et *Haemophilus influenzae* où il y a reconnaissance d'une séquence spécifique présente chez ces espèces (séquence "DUS" pour DNA Uptake Sequence) [2,20,88].

L'ADN chromosomique et les plasmides utilisent la même voie d'entrée dans la bactérie comme le montrent aussi ces expériences de compétition d'ADN [111]. Cependant la transformation est sensible à la topologie de l'ADN. Ainsi un plasmide circulaire simple brin ou un plasmide digéré par une seule enzyme de restriction transforment beaucoup moins bien *Acinetobacter baylyi* que l'ADN chromosomique et les plasmides natifs (double brin circulaire) [111]. Ainsi des plasmides linéarisés par une seule enzyme de restriction ont des taux de transformation 1 000 à 10 000 fois plus faibles qu'un plasmide non digéré. Cependant si on transforme avec un mélange de plasmides digérés par différentes enzymes de restriction, on restaure un taux environ 10 fois plus faible qu'un plasmide natif [111]. L'explication serait que les copies de plasmides linéarisés par une enzyme de restriction aideraient à la recircularisation des

plasmides linéarisés par l'autre enzyme et vice versa. De plus il a été montré que cette recircularisation des plasmides est RecA indépendante et ne fait donc pas intervenir la recombinaison homologue [111].

Bien que la translocation de l'ADN ne soit pas séquence spécifique, l'efficacité de transformation peut dépendre en partie de la séquence car l'étape d'intégration de l'ADN par recombinaison est facilitée s'il y a présence d'homologies de part et d'autre de la séquence étrangère. Notamment il a été montré chez *Acinetobacter baylyi* que le taux de transformation dépend de la divergence avec l'espèce donneuse et du locus où s'intègre l'ADN exogène [124].

Les plasmides ne se répliquant pas chez *Acinetobacter baylyi* et l'ADN chromosomique doivent s'intégrer dans le génome pour conduire à une transformation stable. Deux modes principaux d'intégration sont possibles :

1. intégration par deux événements de recombinaison homologue qui se font grâce aux séquences homologues entourant le marqueur qui va s'intégrer (pour les ADN double brin linéaires et les plasmides)
2. intégration par un seul événement de recombinaison conduisant à l'intégration entière du plasmide (uniquement pour les plasmides double brin circulaires).

Il existe d'autres modes d'intégration beaucoup plus rares [13, 60, 134] (figure 3.1).

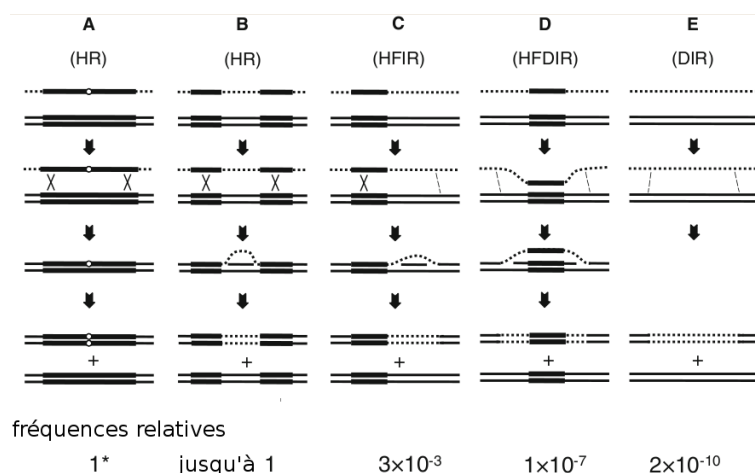


Figure 3.1 – Modes d'intégration d'ADN par recombinaison chez *Acinetobacter baylyi* [13]. Il existe 5 modes d'intégrations : HR, recombinaison homologue, HFIR, recombinaison illégitime facilitée par l'homologie, HFDIR, double recombinaison illégitime facilitée par l'homologie et DIR, double recombinaison illégitime. En pointillé sont représentées les séquences non homologues de l'ADN exogène et en gras sont représentées les séquences homologues. Les croix indiquent des événements de recombinaison homologue et les lignes pointillées obliques, des événements de recombinaison illégitime.

On peut tester la fréquence des deux premiers modes d'intégration en utilisant un plasmide possédant un insert d'ADN chromosomique contenant un gène de résistance à la kanamycine en plus d'un gène de résistance à l'ampicilline propre au plasmide. En effet la résistance à la kanamycine peut s'acquérir par recombinaison homologue alors

que la résistance à l'ampicilline ne peut se faire que par intégration complète du plasmide car il n'est pas entouré de séquences homologues. Ainsi en comparant le nombre de transformants résistants à la kanamycine et le nombre de transformants résistants à l'ampicilline + kanamycine (intégration complète du plasmide), on observe 1000 fois plus d'intégration par double recombinaison homologue [111]. Des résultats similaires sont obtenus lorsqu'on utilise des plasmides non réplcatifs pouvant transformer uniquement par double recombinaison homologue ou uniquement par intégration complète (taux transformation par double recombinaison homologue 100 fois plus élevé) [111]. De plus il a été montré que les plasmides non réplcatifs transforment beaucoup mieux que les plasmides réplcatif ne recombinant pas avec le génome [111]. Ceci suggère que la recombinaison homologue est plus efficace que la recircularisation des plasmides.

3.1.2 Dégradation de l'ADN exogène lors de la transformation naturelle

Au cours de sa translocation, l'ADN exogène double brin est converti en ADN simple brin. Pour démontrer ceci, l'expérience suivante a été réalisée [111]. Une souche d'*Acinetobacter baylyi* ne pouvant pas faire de recombinaison homologue (recA déficiente) a été transformée avec un plasmide. L'ADN cellulaire a été purifié et un Southern blot a été réalisé en condition non dénaturante afin de détecter spécifiquement le plasmide et de déterminer sa forme grâce aux contrôles de plasmides double brin et simple brin. Une bactérie non transformée et une souche non compétente ayant subi le même protocole de transformation ont été utilisées comme contrôle négatif. En comparant les profils de détection, on observe la présence du plasmide uniquement dans la souche transformée et le plasmide est sous sa forme simple brin. Des expériences réalisées chez *Neisseria gonorrhoeae*, laissent penser que l'ADN serait converti en ADN simple brin lors de son passage à travers la membrane cytoplasmique [2, 20, 88] mais aucune démonstration expérimentale n'a été faite chez *Acinetobacter baylyi*.

Il a été montré que l'étape de translocation de l'ADN nécessitait la présence de cations divalent contrairement à la liaison de l'ADN à la surface des bactéries [87, 111]. Ceci suggère l'implication d'une nucléase dans la dégradation de l'un des brin de l'ADN au cours de sa translocation.

L'efficacité de transformation dépend de la longueur de l'ADN exogène [110]. Le taux de transformation augmente avec la taille de manière biphasique. Les fragments en dessous de 600 pb transforment peu efficacement (taux inférieur à 10^{-6}) mais à partir de 1kb environ, le taux de transformation devient élevé ($> 10^{-3}$) et son augmentation avec la taille ralentit. [110]. Il a été aussi montré que, dans une culture d'*Acinetobacter baylyi*, il faut statistiquement 750 fragments internalisés pour obtenir un transformant [110]. Ceci suggère que l'intégration de l'ADN entre fortement en compétition avec sa dégradation. Une bactérie *Acinetobacter baylyi* est cependant capable d'internaliser beaucoup d'ADN (plus de 65 fragments de 2,4 kb par cellule [110]).

| nom du gène | protéine | fonctions |
|-------------|----------|---|
| ACIAD0209 | smf,dprA | import de l'ADN, liaison à l'ADN et recombinaison (fonction potentielle) |
| ACIAD0242 | comM | magnesium chelatase, protéine de compétence potentielle |
| ACIAD0360 | pilD | export et maturation des pseudopilines |
| ACIAD0361 | pilC | stabilisation de l'assemblage du pilus de type IV |
| ACIAD0362 | pilB | biogénèse du pilus de type IV |
| ACIAD0911 | pilU | twitching motility |
| ACIAD0912 | pilT | NTPase, twitching motility |
| ACIAD2639 | comA | transport de l'ADN à travers la membrane interne |
| ACIAD2898 | comL | protéine de compétence potentielle |
| ACIAD3064 | comEA | protéine potentielle de liaison à l'ADN |
| ACIAD3226 | comF | protéine de compétence potentielle |
| ACIAD3314 | comF | pseudopiline, transport de l'ADN à travers le périplasma |
| ACIAD3315 | comE | pseudopiline, transport de l'ADN à travers le périplasma |
| ACIAD3316 | comC | liaison à l'ADN et entrée de l'ADN |
| ACIAD3318 | comB | pseudopiline, transport de l'ADN à travers le périplasma |
| ACIAD3338 | comP | pseudopiline, transport de l'ADN à travers le périplasma |
| ACIAD3355 | comQ | sécrétine, pseudopiline, translocation de l'ADN à travers la membrane externe |
| ACIAD3357 | comO | protéine potentielle de la membrane cytoplasmique |
| ACIAD3359 | comN | protéine potentielle de la membrane cytoplasmique |
| ACIAD3360 | comM | protéine potentielle de la membrane cytoplasmique |

Tableau 3.1 – *liste des gènes de compétence chez Acinetobacter baylyi*. Dans ce tableau est indiquée la liste des gènes liés à la compétence hormis les gènes impliqués dans l'étape d'intégration de l'ADN par recombinaison. Les fonctions de ces gènes ont été déduites par la similarité de leurs protéines avec des homologues chez d'autres espèces naturellement compétentes. Pour les gènes ACIAD2639 (*comA*), ACIAD3314 (*comF*), ACIAD3315 (*comE*), ACIAD3316 (*comC*), ACIAD3318 (*comB*) et ACIAD3338 (*comP*), les fonctions de ces gènes ont été montrées expérimentalement chez *Acinetobacter baylyi* (anciennement nommée *Acinetobacter sp. BD413*) [16, 46, 58, 85, 121, 122].

3.1.3 Gènes impliqués dans la translocation de l'ADN exogène

On a pu identifier 20 gènes de compétence potentiellement impliqués dans la transformation naturelle chez *Acinetobacter baylyi* (étapes de translocation de l'ADN du milieu extérieur jusqu'au cytoplasme) [3, 7] (tableau 3.1). Les fonctions de ces gènes sont pour la plupart déduites par similarité des protéines avec des homologues nécessaires à la transformation naturelle chez d'autres espèces naturellement compétentes.

Pour les gènes ACIAD2639 (*comA*), ACIAD3314 (*comF*), ACIAD3315 (*comE*), ACIAD3316 (*comC*), ACIAD3318 (*comB*) et ACIAD3338 (*comP*), les fonctions de ces

gènes ont été démontrées expérimentalement chez *Acinetobacter baylyi* (anciennement nommée *Acinetobacter* sp. BD413) [3,16,46,58,85,121,122]. Il a été ainsi mis en évidence que comA, comB, comC et comP sont essentiels (mutants non compétents) alors que les mutants comE et comF (ACIAD3314) ont un taux de transformation réduit de 10 et 1000 fois respectivement. De plus comB, comC, comE, comF (ACIAD3314) et comP ne sont pas nécessaires à la formation des pili et la sécrétion de type II ne dépend pas de l'expression de comE et comF (ACIAD3314). Inversement des mutants sans pili ont un taux de transformation sauvage [3, 50]. Ceci suggère que, chez *Acinetobacter baylyi*, la machinerie de translocation de l'ADN est une structure fonctionnellement indépendante.

Pour comP, des études plus poussées ont été faites [122]. Trois formes de comP ont été révélées : une forme de 15 kDa correspondant à la taille déduite de sa séquence, une forme de 20 kDa glycosylée et une forme de 23 kDa qui a subi plus de modifications. La forme glycosylée de 20kDa est la forme majoritaire. Elle est présente surtout dans la membrane externe et la membrane interne mais on en détecte aussi dans le périplasme. La forme de 23 kDa n'est localisée que dans la membrane externe tandis que la forme non modifiée de 15 kDa n'est présente que dans la membrane interne. La localisation de ces différentes formes de comP est la même dans des mutants comE, comF ou comC, ce qui implique que ces trois gènes ne sont pas impliqués dans l'export, la localisation cellulaire et la glycosylation de comP. De plus il a été montré que la glycosylation de comP n'est pas essentielle pour la compétence bien que comP soit indispensable pour la transformation. En plus de ne pas être nécessaire à la formation des pili (voir plus haut), comP n'a pas pu être localisée dans les pili. Cependant sa localisation à la fois dans la membrane externe, le périplasme et la membrane interne reste cohérente avec une organisation en pseudopilus. De plus comme les pseudopilines comP et comB sont absolument essentielles pour la compétence alors que les pseudopilines comE et comF ne le sont pas, on peut émettre l'hypothèse que comP et comB peuvent remplacer comE et comF mais pas l'inverse.

Pour les autres gènes de compétences, il n'y a pas eu de résultats publiés mais certaines informations sont citées dans [3]. Ainsi on sait que comEA, comQ, pilC et pilD sont essentiels pour la transformation naturelle alors que les mutants dprA n'ont qu'un taux de transformation dix fois plus faible.

3.1.4 Modèle de mécanisme de translocation de l'ADN exogène

D'après la fonction des gènes de compétence identifiés chez *Acinetobacter baylyi*, il a été proposé le modèle suivant (figure 3.2). L'ADN exogène se lie à la membrane externe au niveau de comC ou des protéines comQ. Ces dernières forment une structure polymérique en anneau permettant à l'ADN de traverser la membrane externe. L'entrée de l'ADN ne dépend pas de sa séquence (pas de distinction entre ADN homologue et hétérologue [87, 111]) car contrairement à *Neisseria gonorrhoeae* et *Haemophilus influenzae*, il n'y a pas de système de reconnaissance d'une séquence spécifique (séquence "DUS" pour DNA Uptake Sequence) [2, 20, 88]. Ensuite l'ADN traverse le peptidoglycane et le périplasme grâce à la liaison de comEA et à la rétractation du pseudopilus composé de

comB, comE, comF et comP. Ce transport consomme de l'énergie car la rétraction du pseudopilus nécessite l'ATPase PilT et de façon plus globale il a été montré que la transformation naturelle nécessite de l'énergie [111]. Enfin l'ADN double brin traverse la membrane cytoplasmique via le canal comA où il est converti en ADN simple brin par hydrolyse d'un brin.

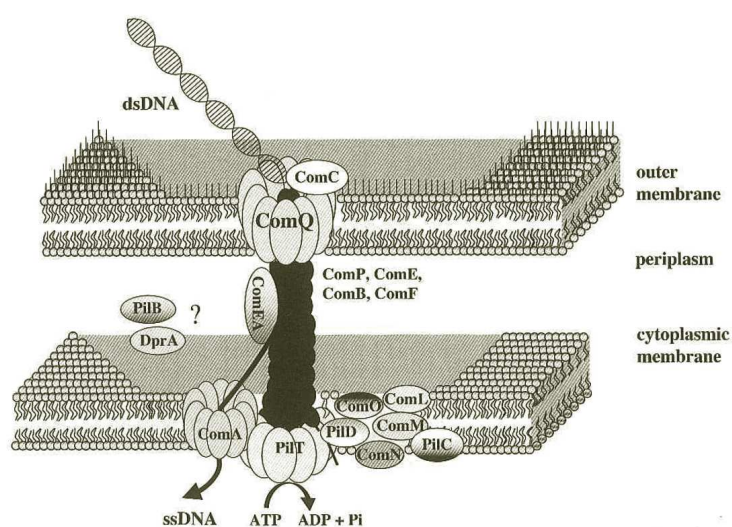


Figure 3.2 – Modèle d'entrée de l'ADN exogène lors de la transformation naturelle chez *Acinetobacter baylyi* (d'après [3]). La liaison de l'ADN exogène sur la membrane externe se fait au niveau de comC ou des protéines comQ. Ces dernières forment une structure polymérique en anneau permettant à l'ADN de traverser la membrane externe. Le transport de l'ADN à travers le peptidoglycane et le périplasm est réalisé grâce à la liaison de comEA et à la rétraction du pseudopilus composé de comB, comE, comF et comP. Ce transport consomme de l'énergie car la rétraction du pseudopilus nécessite l'ATPase PilT. Enfin l'ADN double brin traverse la membrane cytoplasmique via le canal comA où il est converti en ADN simple brin par hydrolyse d'un brin.

3.2 Induction de la compétence chez *Acinetobacter baylyi*

La compétence des bactéries naturellement transformables est un état physiologique induit par différents paramètres physico-chimiques (pH, température, nutriment limitant...) et physiologiques (phase de croissance). Les conditions d'induction sont variables d'une espèce à une autre [64,88].

3.2.1 Taux de transformation en fonction de la phase de croissance

Le taux de transformation maximal est obtenu lorsqu'on dilue dans du milieu frais une culture d'*Acinetobacter baylyi* à saturation. Ensuite ce taux diminue au cours de la phase de croissance exponentielle pour devenir minimal en phase stationnaire [3, 46, 109, 111, 122].

Pour tester s'il est nécessaire d'être passé par une étape de saturation pour induire la compétence, on peut regarder quel est le taux de transformation de bactéries maintenues constamment en phase exponentielle. La première méthode pour maintenir des bactéries en croissance exponentielle est de les diluer dans du milieu frais pendant qu'elles sont en phase exponentielle (à une DO donnée où toutes les bactéries sont en phase exponentielle). L'autre méthode est d'utiliser des techniques de culture en continu (turbidostat ou chemostat) où les bactéries sont constamment alimentées en milieu frais et le volume est maintenu constant en éliminant le trop plein. Dans la première méthode, le taux de transformation décroît avec le nombre de jours que la culture a été maintenue en phase exponentielle. Ainsi après 4 jours de culture, on peut passer d'un taux de transformation initialement de $5 \cdot 10^{-2}$ à un taux de $7 \cdot 10^{-5}$ soit une baisse d'un facteur 1 000 environ [109]. Ceci confirme que pendant la phase exponentielle, il y a une décroissance progressive de la compétence. En turbidostat, la densité optique de la culture est maintenue constante en régulant le débit d'arrivée de milieu (et donc le taux de dilution $D = \text{débit d'arrivée de milieu} / \text{volume de la culture}$). Après 3 jours de culture où les bactéries sont maintenues à plus de 95% de leur taux de croissance maximal μ_{max} ¹, il y a une diminution d'un facteur 100 de la transformabilité des bactéries [109]. Ceci confirme aussi qu'il y a une décroissance de la compétence au cours de la phase exponentielle.

La seconde technique de culture en continu est le chemostat où le taux de dilution D est maintenu constant et impose ainsi un taux de croissance $\mu = D$ (cf section 4.1). Pour essayer de comprendre si c'est la saturation des bactéries et/ou l'apport de milieu riche qui induit la compétence, des expériences de transformation de culture en chemostat après l'ajout ("pulse") de substrat limitant ont été réalisées [109]. Le "pulse" en substrat limitant reproduit l'effet de l'apport de milieu riche qui relance la croissance des bactéries à leur μ_{max} . Pour tester si la saturation des bactéries avant la dilution dans du milieu riche est importante pour l'induction de la compétence, l'alimentation du chemostat a été coupée en même temps ou avant le "pulse" en substrat limitant (ici c'est le potassium). Un "pulse" en potassium immédiatement après l'arrêt de l'arrivée de milieu induit une augmentation du taux de transformation d'un facteur 3-4 alors qu'après un arrêt de 30 min, on a une augmentation d'un facteur 30. Dans les 2 cas, les taux transformations obtenus sont beaucoup plus faibles que ceux obtenus en diluant une culture à saturation (culture sur la nuit) dans du milieu riche où dans ce cas on observe une augmentation du taux de transformation d'un facteur 10 000 ou plus juste après dilution.

1. voir notion de taux de croissance dans la section 4.1

3.2.2 Effet du nutriment limitant sur l'induction de la compétence

Dans un chemostat, les bactéries sont alimentées en continu avec du milieu frais avec un taux de dilution constant $D = \mu$. Pour des cultures en chemostat limitées en azote, en carbone ou en phosphate, le taux de transformation augmente exponentiellement avec le D : il est autour de 10^{-7} - 10^{-6} pour un D faible et augmente jusqu'à 10^{-4} pour un D proche de μ_{max} [109]. Cependant cet effet dépend du nutriment limitant car pour une culture en chemostat limitée en potassium, le taux transformation se maintient autour de $2,2 \cdot 10^{-4}$ [109] .

Il avait été émis l'idée que la transformation naturelle serait un moyen d'apporter des nutriments car l'ADN peut être une source de carbone, d'azote et de phosphate. Les résultats obtenus avec les cultures en chemostat laissent plutôt penser que l'ADN n'est pas utilisé comme source de nutriment car le taux de transformation diminue quand le D diminue, c'est à dire quand l'effet de la limitation en nutriment augmente.

Il a été aussi montré que chez *Acinetobacter baylyi*, la source de carbone et sa concentration initiale n'a pas d'effet significatif sur le taux de transformation. De plus *Acinetobacter baylyi* transforme aussi bien en milieu minimum qu'en milieu riche (LB) [109].

3.2.3 Régulation de l'expression des gènes de compétence

Le taux d'expression de comP et comB varie au cours de la phase de croissance [3, 58, 122]. L'expression de ces protéines est maximale en phase stationnaire et minimale en milieu de phase exponentielle contrairement au taux de transformation qui est minimal en phase stationnaire et maximal en début de phase exponentielle (figure 3.3). L'expression de comP et comB n'augmente pas avec le taux de transformation, ce qui laisse penser que ce n'est pas l'expression de ces gènes qui induit la compétence. Le fait que le taux de transformation soit maximal juste après la dilution d'une culture à saturation dans du milieu riche va dans ce sens car l'induction de la compétence est trop rapide pour impliquer l'induction d'une synthèse de protéines. Ceci est confirmé par le fait que le chloramphénicol (inhibiteur de synthèse protéique) n'inhibe pas la transformation. Cependant le fait que comP et comB soient fortement exprimés en phase stationnaire explique en partie pourquoi on obtient des taux de transformation plus élevés en transformant des bactéries qui viennent de sortir d'une phase stationnaire prolongée. On en déduit que l'induction de la compétence n'est pas due à une synthèse de la machinerie de translocation de l'ADN (qui se ferait plutôt en phase stationnaire) mais à son activation lors du passage dans un milieu riche.

3.2.4 Effet du pH sur l'induction de la compétence

Acinetobacter baylyi est capable de pousser dans une gamme de pH allant de 5 à 8. Son taux de transformation dépend du pH [109]. Si les bactéries sont transformées à un pH en dessous de 6,7, le taux de transformation diminue d'un facteur 20 à pH 6

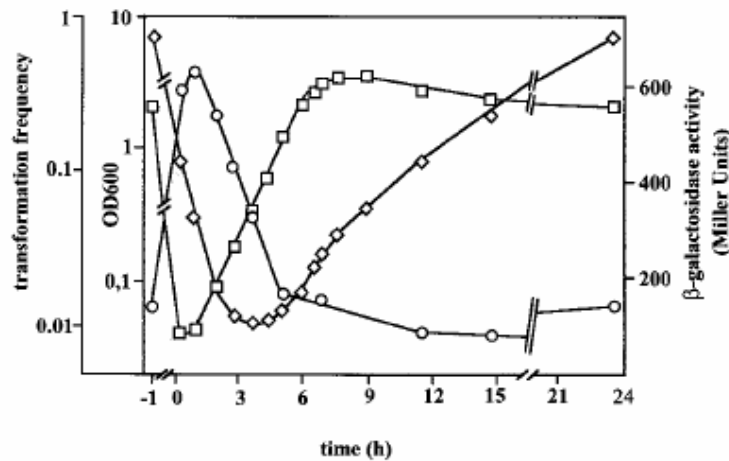


Figure 3.3 – *Expression de comP et taux de transformation en fonction de la phase de croissance chez Acinetobacter baylyi (d’après [3, 122]).* L’expression de *comP* est maximale en phase stationnaire et minimale en milieu de phase exponentielle contrairement au taux de transformation qui est minimal en phase stationnaire et maximal en début de phase exponentielle. □ densité optique des bactéries (OD600), ○ taux de transformation, ◇ expression de *comP* mesurée grâce à l’activité β -galactosidase (fusion en phase de *lacZ* à la séquence contenant le promoteur, le site initiation de la traduction et les 210 premières bases de *comP*)

et d’un facteur 2000 à pH 5,4. Seul le pH pendant la transformation a une influence sur le taux de transformation. En effet si on transforme dans un milieu à pH neutre, quelque soit le pH de la culture avant dilution dans du milieu à pH neutre, le taux de transformation est comparable. Par contre si on dilue une culture à pH neutre dans du milieu à pH acide, le taux de transformation est réduit. Il a été conclu que l’induction de la compétence ne dépend pas du pH mais que c’est plutôt la translocation de l’ADN (ou son intégration) qui est sensible au pH.

3.3 Recherche des conditions optimales pour le brassage de génomes

Afin de choisir dans quelles conditions réaliser le brassage de génomes lors de l’expérience d’évolution en réacteur, il faut tout d’abord déterminer quel est le taux de transformation optimal de notre souche d’*Acinetobacter baylyi* et quels sont les paramètres influençant ce taux de transformation. D’après ce qu’on a vu dans la section précédente, chez *Acinetobacter baylyi*, il est maximal lorsqu’on dilue une culture à saturation dans du milieu riche. C’est pourquoi on a réalisé une première série de test de transformation en diluant une culture à saturation en LB dans du LB (milieu riche).

Une fois déterminés les paramètres influençant le taux de transformation dans ces conditions optimales, on a adapté la méthode de transformation pour des conditions dans lesquelles on réalise le brassage de génomes.

3.3.1 Paramètres influençant le taux de transformations de notre souche cultivée en milieu discontinu

La méthode "classique" de transformation d'*Acinetobacter baylyi* utilisée par [95, 111] est la suivante. Une culture à saturation sur la nuit en LB (ou en milieu minimum) est diluée dans du LB (ou du milieu minimum). Après 2h30 d'incubation à 30°C avec agitation, l'ADN est ajouté et la transformation est réalisée pendant 2h.

Pour les tests suivants, on a choisi de diluer 20 μL d'une culture à saturation en LB dans 300 μL de LB frais (sauf dans le protocole "pulse Nitrate" voir plus bas). On a décidé de mettre l'ADN dès la dilution et d'incuber 4h30 à 30°C (sauf mention contraire) car on n'a pas vu de différence entre les deux méthodes (résultats non montrés).

Pour calculer le taux de transformation, on sélectionne les transformants sur milieu solide LB + antibiotique (kanamycine si l'ADN exogène comporte une cassette de résistance à la kanamycine ou streptomycine s'il possède une mutation ponctuelle conférant une résistance à la streptomycine). Après 24h d'incubation à 30°C, on calcule le taux de transformation qui est le nombre de transformants (nombre de colonies) divisé par le nombre de bactéries étalées (estimé à partir de la densité optique mesurée avant étalement).

3.3.1.1 Effet de l'ADN transformant

Pour les tests de transformation, on a utilisé de l'ADN génomique de la souche d'*Acinetobacter baylyi* qui a été modifiée pour avoir soit une cassette de résistance à la Kanamycine (Kan^R) soit une mutation ponctuelle conférant une résistance à la streptomycine ($Strep^R$). L'ADN génomique a été obtenu de deux façon : soit par extraction (ADNg extrait) soit par amplification par l'ADN polymérase du phage phi29 (ADNg phi29) (voir annexes).

Comme aucun test de transformation d'*Acinetobacter baylyi* n'avait été décrit avec de l'ADNg phi29, un premier test a donc été de comparer les taux de transformation obtenus avec de l'ADNg phi29 Kan^R avec ceux obtenus avec de l'ADNg extrait Kan^R . On a utilisé la méthode "classique" de transformation en LB en mettant l'ADN (à une concentration finale de 1 $\mu\text{g}/\text{mL}$) au moment de la dilution de la culture dans du LB (voir plus haut). Bien qu'on observe une grande variabilité lorsqu'on répète indépendamment les mêmes tests en utilisant des précultures différentes (comme signalé dans [111]), on s'est rendu compte que l'ADNg phi29 Kan^R transformait environ 10 fois moins bien que de l'ADNg extrait Kan^R (tableau 3.2).

Ensuite on a voulu voir si une mutation ponctuelle $Strep^R$ s'intégrait aussi bien dans le génome qu'une cassette Kan^R . Pour cela on a transformé comme précédemment mais en utilisant soit de l'ADNg extrait $Strep^R$ soit de l'ADNg extrait Kan^R . On n'observe pas de différence significative entre les 2 marqueurs $Strep^R$ et Kan^R (tableau 3.2).

| ADN | ADNg phi29 Kan^R | ADNg extrait Kan^R | ADNg extrait $Strep^R$ |
|------------------------|---------------------------------|---------------------------------|---------------------------------|
| taux de transformation | $6, 8.10^{-5} \pm 3, 7.10^{-5}$ | $8, 9.10^{-4} \pm 5, 5.10^{-4}$ | $9, 2.10^{-4} \pm 6, 4.10^{-4}$ |

Tableau 3.2 – Taux de transformation en fonction du type d'ADNg. Une culture à saturation sur la nuit en LB à 30°C a été diluée dans du LB. Au moment de la dilution, on a ajouté de l'ADNg (1 $\mu\text{g}/\text{mL}$) et on a transformé pendant 4h30 à 30°C.

La variabilité des mesures de compétence nous a gêné. Cette variabilité semble d'abord provenir des précultures. Ainsi pour une préculture donnée, la variabilité mesurée entre les répliques d'un même test est beaucoup plus faible que si on part de précultures différentes [111].

Enfin on a voulu déterminer quelle est la concentration optimale d'ADNg extrait pour transformer notre souche d'*Acinetobacter baylyi*. Pour cela on a procédé comme précédemment sauf qu'on a fait varier la concentration d'ADNg extrait Kan^R de 33 ng/mL à 3 $\mu\text{g}/\text{mL}$ (tableau 3.3).

| concentration ADNg extrait | 33 ng/mL | 100 ng/mL | 330 ng/mL | 1 $\mu\text{g}/\text{mL}$ | 3 $\mu\text{g}/\text{mL}$ |
|----------------------------|----------------|----------------|----------------|---------------------------|---------------------------|
| taux de transformation | $1, 4.10^{-4}$ | $3, 7.10^{-4}$ | $5, 5.10^{-4}$ | 1.10^{-3} | $1, 3.10^{-3}$ |

Tableau 3.3 – Taux de transformation en fonction de la quantité d'ADNg extrait. Une culture à saturation sur la nuit en LB à 30°C a été diluée dans du LB. Au moment de la dilution, on a ajouté de l'ADNg extrait Kan^R (1 $\mu\text{g}/\text{mL}$) et on a laissé transformer pendant 4h30 à 30°C. Ces tests ont été réalisés avec une préculture unique afin de limiter la variabilité.

On observe une saturation du taux de transformation à partir de 1 $\mu\text{g}/\text{mL}$ d'ADNg conformément à ce qui a été rapporté dans la littérature [111].

En conclusion, il faut mieux transformer avec de l'ADNg extrait que de l'ADNg phi29 pour faire le brassage de génome et la concentration de 1 $\mu\text{g}/\text{mL}$ a été retenue.

3.3.1.2 Effet du temps de transformation

Il a été montré dans [111] que le taux de transformation augmente avec le temps d'incubation avec l'ADN mais qu'à partir de 2h, ce taux sature. Pour vérifier que notre souche ne sature pas avant 4h30, on a testé différents temps d'incubation avec l'ADN : 30 min, 1h30, 3h et 4h30. On a transformé notre souche avec de l'ADNg phi29 Kan^R ajouté au moment de la dilution de la culture à saturation dans du LB.

On observe une augmentation du taux transformation avec le temps d'incubation avec l'ADN et ce taux ne sature pas avant 4h30 (voir tableau 3.4). On choisit donc de

ne pas réduire le temps de transformation pour les prochains tests et de garder une durée d'incubation de 4h30.

| | | | | |
|-------------------------|-------------|-------------|---------------|-------------|
| temps de transformation | 30 min | 1h30 | 3h | 4h30 |
| taux de transformation | 7.10^{-6} | 2.10^{-5} | $3,8.10^{-5}$ | 1.10^{-4} |

Tableau 3.4 – Taux de transformation en fonction de la durée de transformation. Une culture à saturation sur la nuit en LB à 30°C a été diluée dans du LB frais. Au moment de la dilution, on a ajouté de l'ADNg phi29 Kan^R (1 µg/mL) et on a laissé incubé à 30°C entre 30 min et 4H30.

3.3.1.3 Effet de la température

Comme une première expérience d'évolution sur 3000 générations (expérience pilote AB) a été réalisée à 35°C, on a voulu vérifier que notre souche d'*Acinetobacter baylyi* transforme aussi bien à 35°C qu'à 30°C. Pour cela on a transformé comme précédemment avec de l'ADNg extrait Kan^R ou de l'ADNg phi29 Kan^R (1 µg/mL) sauf qu'on a incubé à 35°C au lieu de 30°C. On retrouve dans les deux cas des taux de transformation situés dans la gamme de ceux obtenus à 30°C (tableau 3.5). Il n'y a donc pas d'effet significatif de la température sur la transformation.

| ADN transformant | ADNg phi29 Kan ^R | ADNg extrait Kan ^R |
|-------------------------------|-------------------------------|-------------------------------|
| taux de transformation à 30°C | $6,8.10^{-5} \pm 3,7.10^{-5}$ | $8,9.10^{-4} \pm 5,5.10^{-4}$ |
| taux de transformation à 35°C | 1.10^{-4} | $3,8.10^{-4}$ |

Tableau 3.5 – Taux de transformation en fonction du type d'ADNg et de la température. Une culture à saturation sur la nuit en LB à 30°C ou 35°C a été diluée dans du LB. Au moment de la dilution, on a ajouté de l'ADNg et on a laissé incubé 4h30 à 30°C ou à 35°C

3.3.1.4 Effet du milieu et du protocole de transformation

Deux méthodes existent pour relancer la croissance d'une culture à saturation : soit diluer la culture dans du milieu frais, soit ajouter du substrat limitant (faire un "pulse"). Pour nos expérience d'évolution, on cultive en continu les bactéries dans un réacteur alimenté en milieu minimum limité en nitrate²⁾ et l'idéal serait donc de faire la transformation directement dans le réacteur. La méthode la plus simple serait de faire un "pulse" de nitrate après l'arrêt de l'alimentation pour relancer la croissance

2. 50 mM de phosphate de potassium, 1 mM de nitrate de potassium, 3 mM de citrate de Potassium, 9 mM d'acide citrique, 1 mM de sulfate de magnésium, des oligoéléments (3 µM FeCl₃, 1 µM ZnSO₄, 1 µM CoCl₂, 0,6 µM CuSO₄, 1 µM MnSO₄, 1 µM Molybdate de sodium, 1 µM MnCl₂ et 1 µM CaSO₄), et du polypropylène glycol 2000 (PPG 2000) à une dilution 1 :20000 comme antimousse.

pendant la transformation (comme ce qui a été fait pour les expériences de transformation en chemostat dans ([109] et voir 3.2.1). Pour tester dans un premier temps si un "pulse" de nitrate peut être envisageable, on a réalisé l'expérience suivante. Une culture à saturation dans le milieu minimum limité en nitrate est transformée en effectuant un "pulse" de 20 mM de nitrate au moment d'ajouter l'ADNg phi29 Kan^R ($1 \mu g/mL$). On laisse incubé 4h30 à 30°C avant d'étaler sur boîtes LB + kanamycine. On obtient un taux de transformation de $6,95.10^{-5} \pm 1,9.10^{-5}$ soit un taux comparable à celui obtenu en transformant avec de l'ADNg phi29 Kan^R après dilution d'une culture à saturation en LB ($6,8.10^{-5} \pm 3,7.10^{-5}$).

On peut donc envisager qu'un pulse de nitrate sera aussi efficace qu'une dilution et qu'on aura des taux de transformation en milieu minimum similaire à ceux en LB. Cependant bien qu'il ait été montré qu'*Acinetobacter baylyi* transforme aussi bien en LB qu'en milieu minimum, il a aussi été observé une baisse de compétence lorsqu'on transforme des cultures qui ont été maintenues en phase exponentielle pendant plusieurs jours (voir 3.2.1 et [109]). D'où la nécessité de faire des tests de transformation en utilisant des cultures en réacteur et ce sera l'objet de la section suivante. Cependant, d'après les expériences qu'on vient de faire, on peut déjà envisager de réaliser des transformations pendant 4h30 à 35°C avec de l'ADNg extrait et en faisant un "pulse de nitrate" plutôt qu'une dilution dans du milieu.

3.3.2 Taux de transformation d'une culture d'*Acinetobacter baylyi* en réacteur

Une fois qu'on a déterminé les paramètres influençant le taux de transformation d'une culture d'*Acinetobacter baylyi* en milieu discontinu, on a essayé de transformer une culture d'*Acinetobacter baylyi* en réacteur de type chemostat. Dans un premier temps on s'est inspiré des protocoles utilisés dans des tests de transformation en chemostat (voir section 3.2.1 et [109]). Pour cela, on a cultivé à 35°C notre souche d'*Acinetobacter baylyi* dans un réacteur alimenté en milieu minimum limité en nitrate (MS121N). Pour effectuer la transformation, on a stoppé l'arrivée de milieu avant d'ajouter 20 mM de nitrate ("pulse" de nitrate) et de mettre de l'ADNg extrait Kan^R ($1 \mu g/mL$). Au bout de 4h30, l'alimentation en milieu est remise et un prélèvement est effectué afin de sélectionner les transformants Kan^R sur boîte LB kanamycine. Malheureusement on n'a jamais réussi à transformer dans le réacteur (taux inférieur à 1.10^{-9}) même en modifiant certains paramètres comme le temps de transformation (entre 1h et 5h) et la durée entre l'arrêt de l'alimentation et le "pulse" de nitrate (on a fait varier entre 0 et 90 min).

Pour essayer de comprendre pourquoi on n'arrive pas à transformer dans le réacteur, on a entrepris de faire des tests en dehors du réacteur en utilisant des prélèvements faits en sortie de réacteur. L'avantage de procéder ainsi est de pouvoir réutiliser la même culture en réacteur pour plusieurs tests. De plus, on peut les faire en petit volume (tube eppendorf de 2mL), ce qui permet de les multiplier aisément. Dans un premier temps, on a tenté de refaire les mêmes tests que ceux faits en réacteur pour vérifier

si on n'arrivait toujours pas à transformer des bactéries cultivées en réacteur avec les protocoles utilisés. Pour le vérifier, on a prélevé des bactéries du réacteur et on a ajouté 20 mM de nitrate et 1 $\mu\text{g}/\text{mL}$ d'ADNg extrait *Kan^R*. On a laissé transformer pendant 4h30 à 35°C avec agitation avant de sélectionner les transformants *Kan^R* sur boîtes LB kanamycine. Étonnement, on a réussi à transformer les bactéries avec un taux de $1,5 \cdot 10^{-6} \pm 8 \cdot 10^{-7}$ soit un taux environ 600 fois plus faible que ceux obtenus avec une culture en milieu discontinu (voir section précédente 3.3.1.4). Cependant cette baisse de compétence est du même ordre que celle obtenue pour des cultures en chemostat (voir section 3.2.1 et [109]). Pour essayer de voir si on peut augmenter ce taux en recréant une saturation des bactéries avant la transformation, on a laissé le prélèvement entre 1h et 5h à 35°C avec agitation avant d'ajouter 20 mM de nitrate et de l'ADNg extrait (1 $\mu\text{g}/\text{mL}$). On n'observe pas de différence significative dans les taux de transformation.

Pourquoi arrive-t-on à transformer en dehors du réacteur mais pas en réacteur ? Ce qui distingue la transformation en réacteur de la transformation en tube, c'est la méthode d'oxygénation (par bullage en réacteur ou par agitation en tube) et la surface du récipient (verre ou plastique). Pour cerner le problème, on a fait l'expérience suivante. On a transformé un prélèvement de réacteur en faisant comme précédemment mais en utilisant différentes matières de tube (verre ou plastique) et différentes méthodes d'oxygénation (par agitation ou par bullage d'air via une bougie filtrante). On s'est rendu compte qu'on n'arrivait pas à transformer lorsqu'on faisait buller de l'air pendant la transformation. Ceci explique pourquoi on n'arrive pas à transformer en réacteur vu qu'on oxygène et mélange notre culture en réacteur en faisant buller de l'air par une bougie filtrante.

Vu qu'on ne peut pas transformer en réacteur à cause du bullage, on a essayé de voir comment on pouvait transformer en dehors du réacteur. Comme on veut transformer l'ensemble de la population afin de ne pas créer de goulot d'étranglement ("bottle-neck"), on doit prélever de grand volume de bactéries du réacteur, ce qui rend les protocoles impliquant des dilutions de bactéries peu envisageables. On a essayé de voir si en transformant les bactéries dans un milieu plus riche (en LB) que celui où elles sont cultivées, on arriverait à augmenter le taux de transformation. Pour tester cela, après avoir prélevé des bactéries du réacteur, on les centrifuge et les resuspend dans du LB avant de les transformer pendant 4h30 avec de l'ADNg extrait. On obtient cette fois-ci un taux de transformation de $2,3 \cdot 10^{-5} \pm 5,2 \cdot 10^{-6}$ soit un taux 10 fois plus élevé qu'avec un pulse de nitrate. On a essayé dans un premier temps d'utiliser ce protocole pour les expériences d'évolution avec brassage de génomes mais on s'est ravisé car il y avait trop de risque de contamination pendant les manipulations en dehors du réacteur. En plus les conditions de transformation sont très éloignées des conditions de culture des bactéries en réacteur.

Finalement, pour l'expérience d'évolution avec brassage de génomes, on a opté pour transformer un grand prélèvement du réacteur (plus de 70 % de la population) en faisant une injection de 20mM de nitrate au moment d'ajouter l'ADNg extrait. On laisse transformer 4h30 à 35°C avec agitation avant de remettre les bactéries dans le réacteur.

Chapitre 4

Montages et protocoles pour les expériences d'évolution en chemostat

Pour les expériences d'évolution réalisées au laboratoire, les bactéries ont été cultivées dans des réacteurs de type chemostat.

Dans une première partie, je présente le principe de fonctionnement d'un chemostat avant de détailler dans une deuxième partie le montage qui a été utilisé pour l'expérience d'évolution avec brassage de génomes et son contrôle avec une souche non compétente. Enfin dans une troisième partie, j'aborderai les protocoles utilisés pour faire du brassage de génomes.

4.1 Principe de fonctionnement d'un chemostat

Un chemostat est un (bio)réacteur où on cultive en continu des microorganismes (ou des cellules) en apportant du milieu nutritif à un débit donné Q et en conservant un volume constant de culture (V) par élimination du trop plein avec ce même débit Q ([97], [119]).

Un paramètre caractérisant une culture de microorganismes est le taux de croissance μ . Dans le cas d'une culture en discontinu (appelée aussi "en batch"), si x est le nombre de bactéries dans la culture alors l'accroissement dx de la population pendant l'intervalle de temps dt est $dx = \mu \cdot x \cdot dt$ soit $dx/dt = \mu \cdot x$. Si μ est constant alors

$$x_t = x_0 \cdot e^{\mu \cdot t} \quad (4.1)$$

En "batch", d'après Monod [97], μ dépend de la concentration en substrat limitant S^1 :

$$\mu = \mu_m \times \frac{S}{(k_s + S)} \quad (4.2)$$

1. en "batch", un substrat est limitant si le nombre final de bactéries augmente avec la concentration initiale de ce substrat. Lorsque ce n'est plus le cas, un autre substrat (ou facteur) est limitant

où μ_m est le taux de croissance maximal, S la concentration en substrat limitant et k_s , la concentration en substrat limitant pour laquelle on a $\mu = \mu_m/2$.

Quand $S \gg k_s$ alors $\mu = \mu_m$. C'est en général le cas dans la plupart des milieux de culture et on a donc $\mu = \mu_m$ pendant la phase exponentielle de croissance où toutes les bactéries se divisent à leur vitesse maximale.

Pour une culture en continu de type chemostat, l'accroissement dx de la population du chemostat dépend du nombre de bactéries formées par la croissance des bactéries ($\mu \cdot x \cdot dt$) mais aussi du nombre de bactéries éliminées par le trop plein qui est le produit de la concentration des bactéries dans le réacteur (x/V) par le volume de culture éliminé pendant l'intervalle de temps dt ($Q \cdot dt$). Le nombre de bactéries éliminées pendant l'intervalle de temps dt est donc $Q \cdot x/V \cdot dt = D \cdot x \cdot dt$ où $D = Q/V$ est le taux de dilution. Pour résumer, l'accroissement de la population dans un chemostat est :

$$dx = \mu \cdot x \cdot dt - D \cdot x \cdot dt \quad (4.3)$$

soit

$$dx/dt = (\mu - D) \cdot x \quad (4.4)$$

A l'équilibre dans un chemostat, on a autant de bactéries formées par croissance que de bactéries éliminées par le trop plein ($dx = 0$). D'après l'équation 4.3, on a

$$\mu \cdot x \cdot dt = D \cdot x \cdot dt$$

soit

$$\mu = D$$

Ainsi contrairement au culture en discontinu ("en batch") où les bactéries passent par toutes les phases de croissance², les cultures en continu de type chemostat maintiennent les bactéries dans un état de croissance donné dépendant du taux de dilution D imposé ($\mu = D$).

Cependant si on impose un $D > \mu_m$, l'élimination des bactéries par le trop plein sera plus rapide que leur production par croissance (d'après l'équation 4.4, $dx/dt < 0$ quand $\mu < D$). D'après l'équation 4.4, si on suppose que $\mu = \mu_m$ lorsque le réacteur est en phase de vidage alors

$$x_t = x_0 \cdot e^{(\mu_m - D) \cdot t} \quad (4.5)$$

Si on prend le logarithme, on aura

$$\ln(x_t) = \ln x_0 + (\mu_m - D) \cdot t \quad (4.6)$$

En traçant $\ln(x_t)$ en fonction de t , on obtient une pente qui vaut $(\mu_m - D)$ et comme on impose D , on peut en déduire le μ_m . Ainsi pour déterminer le μ_m d'une population en chemostat, on peut vider partiellement le réacteur en mettant transitoirement un

2. la phase de latence ($\mu = 0$) où les bactéries se préparent à la division, la phase d'accélération ($0 < \mu < \mu_m$), la phase exponentielle où toutes les bactéries se divisent à leur vitesse maximum ($\mu = \mu_m$), la phase de ralentissement ($0 < \mu < \mu_m$) où le substrat limitant commence à s'épuiser, la phase stationnaire dite de saturation ($\mu = 0$) et enfin éventuellement la phase de déclin ($\mu < 0$)

D très élevé ($D \gg \mu_m$).

Un autre paramètre lié au taux de croissance est le temps de génération ou temps de doublement T qui correspond au temps nécessaire pour que le nombre de bactéries double ($x_t/x_0 = 2$).

Si μ est constant ($\mu = D$ en chemostat ou $\mu = \mu_m$ en "batch") alors d'après l'équation 4.5, $x_T/x_0 = e^{\mu \cdot T} = 2$. Si on prend le logarithme, on obtient $\mu \cdot T = \ln 2$ soit

$$T = \ln 2 / \mu \quad (4.7)$$

Pour les expériences d'évolution en chemostat, on a aussi déterminé N , le nombre de générations qui se sont écoulées pendant l'expérience. Si on suppose qu'on est à l'équilibre dans le chemostat pendant la durée dt , alors $N = dt/T$ correspond au nombre de générations qu'il y a eu pendant la durée dt .

4.2 Montage pour les expériences d'évolution avec ou sans brassage de génomes

Le montage utilisé pour les expériences d'évolution avec ou sans brassage comportent deux réacteurs interconnectés permettant de transférer la population d'un réacteur dans l'autre si besoin (par exemple en cas de formation de biofilm). Ces réacteurs peuvent être déconnectés stérilement, ce qui permet d'autoclaver un réacteur si besoin. La figure 4.1 présente un schéma du montage ainsi que le fonctionnement du réacteur 1 en mode chemostat (hors phase de prélèvement/réinjection, de transfert et de rinçage).

4.2.1 Réacteurs

Chaque réacteur est en verre et a été fabriqué par l'atelier de soufflage de verre. Il comprend deux parties :

- un tube en verre de section carrée (largeur = 2,5 cm, hauteur = 25 cm environ) dont le fond a été soudé. Cette section basse contient la culture et a été choisie carrée pour permettre de mettre en place un système de mesure de densité optique.
- une partie haute approximativement sphérique, soudée en haut du tube de verre à section carrée, a des tubes en verre plongeant vers le bas du réacteur et traversant la paroi du haut du réacteur.

Ces tubes se terminent à l'extérieur du réacteur par un pas de vis permettant de visser un bouchon qui peut être connecté à un tuyau. La taille de ces tubes varie.

Les plus courts, en forme de "bec", sont environ 3 cm au dessus du niveau de culture. Un de ces "bec" est utilisé pour alimenter le réacteur en milieu et un autre est branché à un filtre à air ouvrable si besoin (par exemple pour mettre sous pression le réacteur ou pour évacuer la pression dans le réacteur, voir section 4.3.3).

Le tube le plus long va jusqu'à environ 2-3 cm au dessus du fond du réacteur et se termine par une bougie filtrante permettant de faire buller de l'air dans le réacteur afin

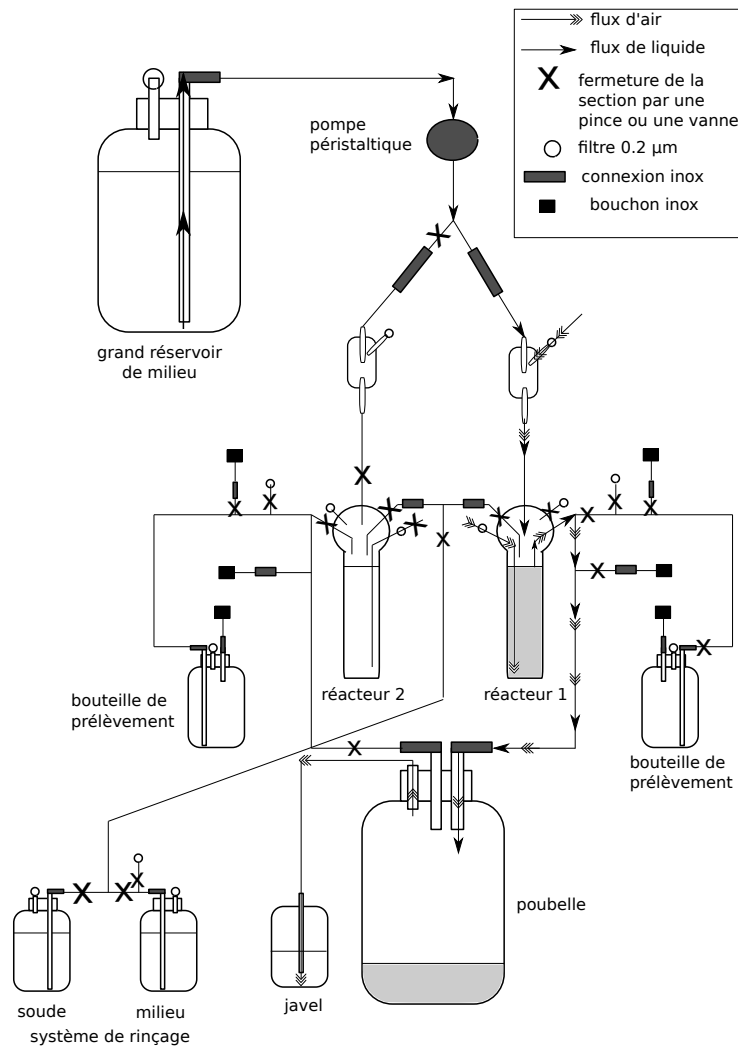


Figure 4.1 – Montage utilisé pour les expériences d'évolution avec ou sans brassage de génome. Les flux d'air et de liquide ainsi que les connexions fermées sont représentés dans le cas du fonctionnement du réacteur 1 en mode chemostat (hors phase de prélèvement/réinjection, de transfert et de rinçage).

d'oxygéner et mélanger la culture. Ce tube longe une paroi du réacteur, permettant ainsi de ne pas gêner le système de mesure de DO (voir section 4.2.8).

Le niveau du réacteur est délimité par la position des tubes qui arrivent à peu près au niveau de la jonction de la section carrée avec la partie haute sphérique :

- un de ces tubes sert à transférer les bactéries vers un autre réacteur et à rincer le réacteur avec du milieu et de la soude (voir section 4.2.6).
- l'autre, légèrement plus long (moins d'un cm plus long), sert à évacuer le trop plein (liquide + air) vers une poubelle et à maintenir un volume constant V de

culture dans le réacteur pour un débit d'air donné.

En effet en plus de dépendre de la position du tube d'évacuation du trop plein (qui impose le niveau maximum de la culture), le volume de culture que peut contenir chaque réacteur dépend du débit d'air qu'on fait buller dans la culture. En effet le volume qu'occupent les bulles d'air dans la culture est d'autant plus grand que le débit d'air est élevé et par conséquent plus le volume de la culture sera faible car on impose en plus un niveau maximum pour la culture.

Pour déterminer le volume effectif du réacteur (V), on fait fonctionner le réacteur en mode chemostat sansensemencer de bactérie et en imposant le débit d'air à 100 mL/min (débit d'air utilisé pour les expériences). Une fois que le niveau de milieu dans le réacteur est à l'équilibre, on arrête l'alimentation en milieu et en air et on déconnecte le réacteur. Ensuite on pèse pour déterminer le volume de milieu contenu dans le réacteur. Sachant que le milieu a une densité de 1 g/mL, on détermine le volume V de culture pour les 2 réacteurs utilisés. V vaut environ 140 mL : 140 mL \pm 0.5 mL pour le réacteur 1 et 142 mL \pm 0.5 mL pour le réacteur 2.

4.2.2 Connexions du réacteur

Chaque tube en verre du réacteur se termine à l'extérieur par un pas de vis permettant de visser un bouchon. Si le tube est utilisé, il est vissé à un embout permettant de connecter de façon étanche des tubes de silicone/tygon (diamètre intérieur 1.6 cm, diamètre extérieur 4.8 cm). Les flux d'air et de liquide entre le réacteur et les différents éléments du montage se font principalement via ce type de tubes.

Pour créer des points de déconnexion/connexion stériles dans le montage en cours de fonctionnement, on utilise des raccords constitués de deux tubes cylindriques en inox qui s'emboîtent l'un dans l'autre. Ces connexions stérilisables permettent notamment de changer le réservoir de milieu, de déconnecter un réacteur pour l'autoclaver ou de brancher/débrancher une bouteille de prélèvement (figure 4.1). Ces connexions/déconnexions se font stérilement à côté d'une flamme et les embouts en inox sont passés sous la flamme.

4.2.3 Arrivée de milieu

Le milieu servant à approvisionner le réacteur est préparé dans de grandes bouteilles de 10 L en verre. Ces bouteilles possèdent un bouchon avec trois ports contrôlés par des vannes manuelles :

- un port est branché à un filtre à air afin que le réservoir reste à pression atmosphérique, tout en préservant la stérilité
- une sortie est traversée par un tube plongeant au fond du réservoir et connecté à l'extérieur via une connexion inox à une pompe péristaltique qui va pomper le milieu à un débit donné.
- le troisième port n'est pas utilisé (vanne fermée)

Après la pompe péristaltique, le tube d'alimentation bifurque en deux grâce à un T afin d'alimenter un des deux réacteurs selon l'ouverture d'une des deux bifurcations

(fermeture à l'aide d'une pince). Juste après cette bifurcation, une connexion inox est présente dans chacune des branches d'alimentation afin de pouvoir déconnecter stérilement un réacteur s'il a besoin d'être autoclavé. Juste avant le branchement au bec d'arrivée de milieu de chaque réacteur, on a mis un système de "casse-goutte" avec une arrivée d'air (figure 4.1) dans le but d'empêcher la propagation de contaminations de bactéries vers le réservoir de milieu stérile. En effet une telle contamination de l'arrivée de milieu peu être induite par des aérosols provenant de la culture dans le réacteur.

Le débit d'arrivée de milieu imposé par la consigne de la pompe est calibré grâce à la mesure de la décroissance de la masse du réservoir en fonction du temps. Pour cela le grand réservoir de milieu est posé sur une balance qui est capable de mesurer au décigramme près et qui est reliée à un ordinateur par liaison série. Les mesures sont commandées toutes les 30 secondes comme pour la mesure de densité optique (voir section 4.2.8) et sont consignées dans un fichier.

Comme la décroissance de la masse du réservoir est due au volume de milieu pompé et que la masse volumique du milieu est 1 g/mL, si on trace la masse du réservoir en fonction du temps (en secondes), la pente de la droite nous donnera le débit de la pompe en mL/s. Après avoir mesuré les débits réels Q correspondant à plusieurs consignes de la pompe, on peut tracer le débit mesuré en fonction de la consigne de la pompe et établir ainsi la relation $Q = A.c$ avec Q , le débit réel, c la consigne de la pompe et A , une constante (pente de la droite).

4.2.4 Arrivée d'air

L'air utilisé provient d'une source d'air comprimé de 10 bars qui est réduite à une pression de 1.2 bar grâce à un régulateur de pression. Un débitmètre et une vanne pointeau sont placés dans le circuit pour permettre de mesurer et ajuster le débit d'air. Le flux d'air peut être réglé entre 20 et 200 mL/min. L'air est rendu stérile grâce à des filtres 0,2 μm . L'alimentation en air du réacteur se fait principalement via le tube en verre long qui plonge jusqu'en bas du réacteur et qui se termine par une bougie filtrante permettant de faire buller de l'air dans le réacteur. En remontant à la surface, les bulles d'air formées permettent de mélanger la culture en plus de l'oxygéner. Cette alimentation principale en air est réglée sur 100 mL/min.

Des alimentations secondaires en air sont présentes. En effet de l'air arrive dans le réacteur en même temps que le milieu grâce au système de casse-goutte. De plus, pour la réinjection de bactéries après le brassage (voir section 4.3.3) ou lors des phases de rinçage avec du milieu ou de la soude (voir section 4.2.6), on a besoin de mettre des bouteilles sous pression et pour cela on utilise une autre arrivée d'air.

4.2.5 Evacuation du trop plein du réacteur vers une poubelle ou une bouteille de prélèvement

Le niveau de liquide dans le réacteur (et non le volume voir section 4.2.1) est déterminé par la position du tube en verre qui est relié soit à la poubelle soit à une

bouteille de prélèvement. En effet le trop plein de culture est éliminé du réacteur en même temps que le trop plein d'air quand le niveau de culture dépasse l'ouverture de ce tube de sortie. Le tuyau d'évacuation du trop plein bifurque en deux juste après la sortie du réacteur : la bifurcation vers la bouteille de prélèvement est fermée en dehors des phases de prélèvement ou de réinjection et le trop plein s'évacue vers une poubelle de 10L (figure 4.1). Cette poubelle a trois ports contrôlés par des vannes manuelles :

- une entrée pour le tuyau d'évacuation du trop plein du réacteur 1
- une entrée pour le tuyau d'évacuation du trop plein du réacteur 2
- une sortie pour évacuer l'air (provenant du trop plein du réacteur 1 ou 2) vers une bouteille contenant de l'eau de javel. Le tube d'entrée plonge au fond de la bouteille pour forcer un passage par bullage.

Pour prélever de grands échantillons de population afin d'extraire leur ADN ou de les transformer avec cet ADN (voir sections 4.3.1 et 4.3.2), on dévie le trop plein de milieu et d'air vers une bouteille de prélèvement de 250 mL (on ferme la bifurcation vers la poubelle et on ouvre celle vers la bouteille de prélèvement). Cette bouteille de prélèvement (ou de réinjection voir section 4.3.3) est munie d'un bouchon avec trois ports contrôlés par des vannes manuelles qui permettent de les ouvrir ou les fermer selon leur utilisation :

- un port est connecté à cette dérivation en sortie de réacteur et possède un tube plongeant jusqu'au fond de la bouteille
- un deuxième est utilisé pour injecter stérilement du nitrate via une connexion inox et qui est fermé par un bouchon inox quand il n'est pas utilisé (voir section 4.3.1)
- un troisième port est branché à un filtre à air qui permet d'évacuer le trop plein d'air tout en maintenant la stérilité.

Quand le volume prélevé est suffisant, on ouvre la bifurcation vers la poubelle et on ferme celle vers la bouteille de prélèvement tout en laissant la vanne de la bouteille de prélèvement ouverte afin de permettre d'évacuer le volume restant dans les tuyaux de la dérivation vers la bouteille. Cette évacuation est réalisée en injectant de l'air via un filtre à air présent au début de la bifurcation vers la bouteille de prélèvement. Ensuite on ferme la vanne d'entrée du trop plein afin de déconnecter stérilement la bouteille et un bouchon inox autoclavé est branché au bout de la dérivation à la place de la bouteille.

Pour prélever de petits échantillons, une deuxième dérivation plus basse (figure 4.1) se terminant par une connexion inox peut être branchée à un système de deux petites bouteilles de 10 mL autoclavées (non montrées sur le schéma du montage) : une bouteille vide permet de prélever un échantillon tandis que l'autre contenant de l'eau autoclavée permet de nettoyer la dérivation avant et après le prélèvement (en mettant sous pression la bouteille d'eau) .

4.2.6 Transfert et système de rinçage

Les deux réacteurs sont interconnectés par un tuyau de transfert qui présente à mi-distance une bifurcation vers un système de rinçage composé de deux bouteilles : une contenant du milieu, l'autre contenant de la soude 1 M (mode de rinçage inspiré de [31]).

Lorsqu'on transfère la culture d'un réacteur (par exemple réacteur 1) vers un réacteur vide (réacteur 2 qui a été autoclavé), on ouvre la section reliant les deux réacteurs mais on ferme la dérivation vers le système de rinçage ainsi que le tuyau d'évacuation du réacteur 1 vers la poubelle. Ainsi le trop plein du réacteur 1 ne peut s'évacuer que vers le réacteur 2. Celui-ci est alimenté en air et le trop plein du réacteur 2 peut s'évacuer vers la poubelle. Après le transfert, au niveau de la bifurcation de l'arrivée de milieu après la pompe, on coupe l'alimentation du réacteur 1 et on ouvre celle du réacteur 2. L'évacuation du réacteur 1 vers la poubelle est ouverte tandis que le tuyau de transfert est fermé du côté du réacteur 2 (entre le réacteur 2 et la bifurcation vers le système de rinçage). Le réacteur 2 fonctionne ainsi en mode chemostat et est déconnecté du réacteur 1 qui va être nettoyé.

Après un transfert vers un autre réacteur ou après un prélèvement pour faire du brassage de génomes, le réacteur n'est plus utilisé et doit être nettoyé voire autoclavé. Le rinçage du réacteur se fait alors en deux temps :

- on le rince d'abord avec du milieu qui arrive à très haut débit afin de vider le réacteur des bactéries
- puis on injecte de la soude afin de tuer les bactéries restantes et de garder le réacteur stérile le temps de sa non-utilisation.

Le réacteur peut aussi être autoclavé notamment en cas de formation de biofilms importants.

Lorsqu'on veut réutiliser un réacteur qui contient de la soude (pour un transfert par exemple), on rince celui-ci avec du milieu afin d'éliminer les débris et de revenir au pH du milieu de culture (mesure du pH au niveau de la sortie vers la poubelle). Lorsqu'on fait un transfert vers ce réacteur, on procède ensuite comme décrit plus haut avec un réacteur vide.

Pour rincer le réacteur avec du milieu ou de la soude, on met sous pression les petits réservoirs (1 L) de milieu ou de soude en utilisant de l'air comprimé qu'on injecte dans la bouteille par un port contenant un filtre 0,2 μm . Lorsque la bouteille est sous pression, on ouvre la vanne de sortie de la bouteille ainsi que la bifurcation vers le réacteur à rincer. Quand on veut arrêter le rinçage, on ferme l'arrivée au niveau du réacteur puis on ferme la vanne de la bouteille et on la dépressurise en déconnectant l'arrivée d'air dans la bouteille et en laissant l'air s'échapper par un filtre à air. Pour éliminer le liquide présent dans le tuyau reliant le système de rinçage au réacteur, on a branché un filtre à air dans ce tuyau du côté du système de rinçage afin de pouvoir injecter stérilement de l'air. Ainsi lorsqu'on injecte de l'air, le liquide présent dans le tuyau est éliminé vers le réacteur dont l'arrivée a été ouverte après avoir branché l'alimentation d'air sur le filtre.

4.2.7 Régulation en température

Pour cultiver les bactéries à 35°C, la portion à section carrée de chaque réacteur (contenant la culture) est inséré dans un support en alliage d'aluminium dont les parois qui font 1 cm d'épaisseur sont régulées en température. Pour chaque réacteur, il y a une résistance en platine qui sert de capteur de température et qui est reliée à un régulateur PID (pour "Proportionnel Intégral Dérivé") afin de respecter la consigne souhaitée. Avec ce système de régulation en température, on peut maintenir la culture à 35°C +/- 0,5°C.

4.2.8 Mesure de la densité optique

Pour mesurer la densité optique (DO, absorbance) de la culture dans le réacteur, on a mis en place un système de mesure composé de diodes qui émettent de la lumière à une longueur d'onde d'environ 800 nm et d'un détecteur photosensible. Les diodes émettrices et le photorécepteur se font face au niveau de la section carrée du réacteur et sont encastrés dans deux parois parallèles du support du réacteur. Les diodes émettent de la lumière qui traverse la culture contenue dans le réacteur et la lumière reçue par le photodétecteur est d'autant plus faible que la culture est opaque. Le courant du photodétecteur dépend donc de cette absorbance. Ce courant est ensuite converti en une tension, elle-même mesurée par un convertisseur analogique-numérique relié à un ordinateur. Pour faire la correspondance entre la tension mesurée et la densité optique, on a tout d'abord mis du milieu sans bactérie dans le réacteur afin de déterminer la valeur de la tension correspondante (DO "nulle"). Ensuite on peutensemencer des bactéries et établir une correspondance entre l'évolution de la tension mesurée par l'ordinateur et la DO mesurée par un appareil commercial à partir de prélèvements faits à différents temps de la croissance des bactéries.

Une fois qu'on a la correspondance entre la tension et la densité optique, on peut estimer le nombre de bactéries (densité optique proportionnelle au nombre de bactéries) dans le réacteur et éventuellement utiliser des modes de fonctionnement dépendant de la densité optique. Cet asservissement se fait grâce à des changements du débit de milieu qui ont lieu lorsqu'on passe au dessus ou en dessous des valeurs d'un des trois paliers de tension (DO).

Par exemple, au lancement du réacteur, les bactéries ne sont pas encore adaptées au milieu et ont un taux de croissance très lent qui augmente ensuite lorsque des mutations avantageuses apparaissent. Ne sachant pas quand ont lieu ces augmentations du taux de croissance, on a besoin de s'assurer que les bactéries tolèrent le taux de dilution qu'on impose afin de ne pas vider le réacteur. Pour cela on régule approximativement la densité optique entre trois valeurs seuils (paliers). En dessous du palier le plus bas, on a une consigne de pompe très basse (voir nulle au tout début). Lorsque la tension dépasse le palier intermédiaire, on impose une consigne de pompe plus élevée et si la culture tolère ce débit, alors la tension peut franchir un troisième palier pour lequel la consigne de pompe est encore plus élevée. Si les bactéries ne tolèrent pas ce taux de dilution, la densité optique va diminuer et si elle descend en dessous de la valeur du

palier intermédiaire alors la consigne de la pompe sera celle du palier intermédiaire et de même si la DO continue à descendre en dessous du palier le plus bas, on aura la consigne du palier le plus bas.

Ainsi lorsqu'on veut augmenter progressivement le taux de dilution via une augmentation du débit de milieu, on impose pour le palier intermédiaire le débit actuel qui est toléré (culture à l'équilibre) et pour le palier supérieur, on met le débit qu'on voudrait atteindre. Tant que les bactéries ne supportent pas le taux de dilution imposé au dessus du palier supérieur, la densité optique va osciller entre celle du palier intermédiaire et celle du palier supérieur. Une fois que les bactéries ont μ_m supérieur ou égal au taux de dilution qu'on veut atteindre, on peut obtenir l'équilibre où la densité optique est constante ($dx/dt=0$) et $D = \mu$.

4.3 Protocole du brassage de génomes

Le brassage de génomes a lieu une fois par semaine et est réalisé en dehors du réacteur (voir raisons chapitre 3 section 3.3.2). Pour pouvoir le faire, il faut dans un premier temps extraire l'ADN de la population du jour $j - 1$ puis transformer les bactéries du jour j avec cet ADN et enfin réinjecter les bactéries dans le réacteur après transformation.

4.3.1 Extraction de l'ADN génomique de la population du réacteur

Pour extraire l'ADN génomique de la population, on prélève 100 mL au niveau du trop plein de culture du réacteur qui a été évacué dans une bouteille de prélèvement de 250 mL (voir section 4.2.5). Comme la densité dans le réacteur est faible (DO autour de 0,3), pour relancer la culture on injecte stérilement 20 mM de nitrate via le port d'injection de la bouteille qui possède une connexion inox. Une fois le volume prélevé, on déconnecte stérilement la bouteille et sous la hotte à flux laminaire, on transfère le contenu de la bouteille dans un erlen. On incube ensuite environ 2H à 35°C 300 rpm dans un incubateur-agitateur pour obtenir une DO autour de 1,2 (soit 2 doublements successifs de la population). Après cette repousse, on transfère le prélèvement dans deux tubes de 50 mL et on centrifuge afin de culoter les bactéries. On resuspend les culots dans du tampon TE (10 mM Tris HCl pH 8, 1 mM EDTA) pour les réunir dans un volume final de 4 mL. Le protocole d'extraction d'ADN génomique est décrit dans les annexes.

4.3.2 Transformation de la population du réacteur avec l'ADN génomique extrait

Pour faire le brassage de génomes, on doit transformer les bactéries avec l'ADN génomique de la population qu'on a extrait le jour précédent. Comme on ne peut pas transformer dans le réacteur (voir raisons chapitre 3 section 3.3.2), on doit prélever

la population du réacteur pour faire le brassage. Qu'on utilise une souche compétente ou en contrôle une souche non compétente, le protocole est le même (seule la souche compétente pourra effectivement faire du brassage) afin d'avoir exactement les mêmes conditions de culture entre l'expérience d'évolution avec brassage et son contrôle sans brassage (souche non compétente). On prélève comme précédemment sauf qu'après avoir transféré dans un erlen, on ajoute en plus des 20 mM de nitrate l'ADN génomique de la population à une concentration finale de $1\mu\text{g}/\text{mL}$. On laisse transformer pendant 4H30 à 35°C 300 rpm avant de réinjecter les bactéries dans le réacteur.

Pour l'expérience d'évolution avec brassage de génomes (souche compétente), on transforme en parallèle un prélèvement de la population avec de l'ADN génomique extrait d'une souche ayant un gène de résistance à la kanamycine (voir chapitre 3 section 3.3.2). Ce test a pour but de suivre l'évolution du taux de transformation au cours de l'expérience.

4.3.3 Réinjection des bactéries transformées dans le réacteur

Si on réinjecte dans un réacteur qui a été nettoyé à la soude, après l'avoir rincé avec du milieu, il faut enlever un volume de milieu correspondant au volume de bactéries qu'on va réinjecter. Pour enlever ce volume, on utilise le tube en verre de l'arrivée d'air principale car il plonge au fond du réacteur. Le tuyau d'arrivée d'air bifurque en deux après le filtre à air et avant l'entrée dans le réacteur (non représenté sur le schéma du montage). Une des branches est connecté au réacteur et l'autre est connecté à une bouteille de prélèvement comme pour les prélèvements du trop plein. Lorsqu'on veut enlever du milieu, on doit fermer l'arrivée d'air avant la bifurcation (après le filtre à air) et ouvrir la dérivation vers la bouteille de prélèvement. Pour que le milieu s'évacue vers la bouteille, on doit mettre sous pression le réacteur en fermant tous les entrées/sorties du réacteur sauf le circuit pour évacuer le milieu et le filtre à air du réacteur où on va brancher une arrivée d'air secondaire. Lorsqu'on a enlevé un volume suffisant, on enlève l'arrivée d'air secondaire par le filtre à air et on ouvre l'arrivée d'air principale. La pression du réacteur s'évacue par le filtre à air secondaire et le reste de milieu dans les circuits est éliminé vers la bouteille ou le réacteur. Ensuite on referme la bifurcation de l'arrivée d'air vers la bouteille.

Pour réinjecter les bactéries dans le réacteur (pas entièrement plein via la manipulation décrite précédemment ou vide car autoclavé), on transfère les bactéries dans une bouteille de prélèvement autoclavée et on la rebranche stérilement à la bifurcation d'évacuation du trop plein (voir schéma du montage, figure 4.1). Ensuite on ouvre la bifurcation vers la bouteille mais on ferme celle vers la poubelle comme on fait pour un grand prélèvement du trop plein. Pour que la culture soit injectée dans le réacteur, on met sous pression la bouteille de prélèvement en injectant de l'air par le port relié à un filtre à air. Une fois qu'il n'y a plus que de l'air qui s'échappe de la bouteille de prélèvement, on ferme la bifurcation vers la bouteille de prélèvement et on ouvre celle vers la poubelle. Comme le trop plein d'air peut s'évacuer vers la poubelle, on ferme le filtre à air secondaire du réacteur. Pour compléter le niveau dans le réacteur on ouvre l'arrivée de milieu du réacteur ainsi que l'arrivée d'air secondaire par le casse-goutte. La

remise en route du réacteur en mode chemostat est une étape délicate car les bactéries transformées sont arrivées à saturation et il leur faut du temps pour se réadapter au condition du chemostat. A cette étape, le contrôle de la DO par régulation de la pompe est importante pour éviter de vider le réacteur.

Chapitre 5

Expériences d'évolution avec ou sans brassages de génomes

Trois expériences d'évolution ont été étudiées en détail dans mon travail pour cette thèse :

- une première expérience d'évolution avec une souche d'*Acinetobacter baylyi* a été réalisée en chemostat sur environ 3000 générations. Cette première expérience a eu pour but de mettre en place le système expérimental. La souche pour démarrer l'expérience est compétente pour la transformation mais les conditions expérimentales de croissance en réacteur conduisent à un taux de transformation très faible ($<10^{-9}$). Par ailleurs, on n'a pas mis en oeuvre de mode avec réinjection d'ADN (comme cela a été le cas pour les expériences suivantes). Cette expérience a permis d'obtenir des informations sur la nature des mutations et sur la diversité générée lors de l'évolution d'*Acinetobacter baylyi* dans des conditions sélectives choisies et avec des échanges horizontaux très faibles. Cette expérience pilote est présentée en détail dans un article qui est soumis (voir section 5.1 et annexes pour les données supplémentaires).
- deux expériences réalisées dans les mêmes conditions (voir sections 4.2 et 4.3) ont été mises au point afin de pouvoir comparer l'évolution avec ou sans brassage de génomes.

5.1 Expérience pilote

Cette expérience pilote a été réalisée avant mon arrivée au laboratoire par Philippe Thomen. J'ai participé à l'analyse de cette expérience en parallèle de mettre au point les expériences d'évolution avec brassage et son contrôle sans brassage.

Long-term diversity and genome adaptation of *Acinetobacter baylyi* in a minimal-medium chemostat

Nadia Jezequel^{1,2}, Marco Cosentino Lagomarsino^{1,3,4}, Francois Heslot^{1,2}, and Philippe Thomen^{1,2,*}

¹Université Pierre et Marie Curie, 4 place Jussieu, 75005, Paris, France.

² Laboratoire Pierre Aigrain, Ecole Normale Supérieure, CNRS (UMR 8551), Université P. et M. Curie, Université D. Diderot, 24 rue Lhomond, 75231 Paris Cedex 05, France

³Génophysique / Genomic Physics Group, CNRS (UMR 7238) "Microorganism Genomics" 15 rue de l'école de Médecine 75006 Paris, France

⁴Università degli Studi di Torino, Dipartimento di Fisica, Via P. Giuria 1, 10125 Torino, Italy

*Corresponding author: E-mail: philippe.thomen@upmc.fr.

Abstract

Often, laboratory evolution experiments for large populations of microorganisms that cannot recombine show two distinct phases: an initial rapid increase in fitness followed by a slower regime. In order to explore the population structure and the evolutionary tree in the later stages of adaptation, we evolved a very large population of about 3×10^{10} *Acinetobacter baylyi* bacteria for approximately 2800 generations from a single clone, while maintaining it in a chemostat at high dilution rate. Nitrate in limiting amount and as the sole nitrogen source was used as a selection pressure. Analysis via resequencing of genomes extracted from populations indicates that diversity can establish at long terms in a chemostat with very simple medium. To find out which biological parameters were targeted by adaptation, we measured the maximum growth rate, the nitrate uptake and the resistance to starvation. Overall, we find that maximum growth rate could be a reasonably good proxy for fitness. The late slow adaptation is compatible with selection coefficients spanning a typical range of 10^{-3} - 10^{-2} per generation as estimated by resequencing, pointing to a possible subpopulations structuring.

Key words: *Acinetobacter*, experimental evolution, resequencing, nitrogen limitation, chemostat, large population.

Understanding adaptation is important in a wide number of biological, ecological and technological contexts (Toft and Andersson 2010; Jackson et al. 2011). In this framework, controlled laboratory evolution experiments performed with large populations of bacteria allow to explore and better understand adaptation under well-defined selection pressure (Buckling et al. 2009; Conrad et al. 2011; Hindré et al. 2012). Through phenotypic characterization and sequencing, these experiments give a quantitative insight into the genomic adaptation of microbial populations (Barrick et al. 2009; Barrick and Lenski 2009; Tenaillon et al. 2012).

Possibly one of the best examples of the impact of this approach concerns fundamental ecology. As a result of the "competitive exclusion principle" (Hardin 1960; Gause 1934), it was previously believed that diversity cannot be created and maintained in simple, spatially and temporally homogeneous environments containing a single resource. The recent *in vitro* evolution experiments of Maharjan and coworkers (Maharjan

et al. 2006, 2012) has subverted these classic views by disproving the competitive exclusion principle and demonstrating that a clonal population of *E. coli* evolving in a chemostat can diversify into a number of subpopulations.

The mechanisms setting up this diversity are unclear, and include convergent evolution and regulatory degeneracy (Gresham et al. 2008; Wang et al. 2010), cross feeding (Friesen et al. 2004; Rosenzweig et al. 1994), and frequency-dependent selection (Levin et al. 1988; Doebeli and Ispolatov 2010). Beardmore and coworkers (Beardmore et al. 2011) recently put forward a fourth mechanism that can explain the emergence and maintenance of diversity in simple environments based on metabolic and physiological trade-offs. A fifth possibility is related to the fact that in such large population-size experiments in absence of recombination, many beneficial mutations compete to drive adaptation, through "clonal interference" (Gerrish and Lenski 1998). A recent laboratory evolution study (Hughes et al. 2012) found stable competition between

coexisting hosts with different plasmid genotypes in plasmid-host adaptation dynamics, suggesting that this diversity plays a role in the evolution of the persistence of drug resistance. Despite recent theoretical and experimental advances (Desai et al. 2007; Desai and Fisher 2007; Park and Krug 2007; Brunet et al. 2008; Good et al. 2012) this regime of adaptation is not completely understood, and many open questions concerns e.g. the dynamics of the advantage (Barrick et al 2009), the role of epistasis (Kahn et al. 2011; Chou et al 2011) and the structuring of a population into sub-populations of competing clones (Maharjan et al 2007). A recent study (Maharjan et al. 2012) used large-scale genomic and phenotypic analysis on the chemostat evolutionary data to identify the extent of diversification. They placed their results into a context of the above theoretical mechanisms of evolutionary divergence and proposed that diversity is unlikely to be explained by any one of the available theoretical models.

A criticism directed to the earlier study of Maharjan and coworkers (Maharjan et al. 2006) is that the experiment was conducted over a relatively short evolutionary time scale (26 days), which may lead some to speculate that running similar experiments for longer times would destroy the observed diversity (Conrad et al. 2011). Here we report of an evolutionary experiments for 124 days (~ 2800 generations) using a different organism, and we report stable, and possibly increasing diversity, also providing an analysis of genome adaptation.

We constructed a chemostat culture using a strain of *Acinetobacter baylyi* growing in a minimal medium, in principle allowing for a limited number of distinct ecological niches. The bacteria were grown in condition of limited amount of nitrate as a sole source of nitrogen, in high dilution rate regime (between 0.6 h^{-1} and 0.9 h^{-1}), which should be the least favorable condition for the onset of diversity (Maharjan et al. 2012). Our system involves a unique growth medium supply, and is composed of interconnected alternating tandem chemostats (figure 1) so as to prevent bacterial attachment to the walls of the chemostat (de Crécy-Lagard et al. 2001). The main advantages of this setup (compared e.g. to serial dilution experiments) are that it allows constant growth conditions, with very large population sizes. The main drawback is that it is cumbersome, and (with our resources) not allowing for replicate populations in parallel. However, we will see that this does not prevent from extracting relevant information. *Acinetobacter baylyi* ADP1 has been sequenced and annotated previously (Barbe et al. 2004); we identified the mutations appearing during the experiment by genome resequencing at three time points late in the experiment. This allowed to determine the frequencies of the mutations at different points in time, and, by genotyping each mutation on isolated clones, to reconstruct a partial population tree. These data lead to approximate quantitative estimates of the selection coefficient of some mutations. We measured the change in maximum growth rate, and found that its increase is consistent with the estimated selection coefficients.

Sequencing and reconstruction of the population tree

In order to study the mutational patterns and connect them with the evolutionary dynamics, we resequenced several samples: the ancestral clone (ABWT), two endpoint clones (AB2800a, AB2800b) and three samples of populations at three different times (AB1900, AB2300, AB2800)¹. Sequencing of the ABWT genome revealed that it has differences with respect to the *Acinetobacter baylyi* ADP1 genome (Barbe et al. 2004). The list of those differences is given as Supplementary Data. The mutations that occurred via the experimental evolution from the ancestral ABWT clone are listed in table 1. In order to infer the order of appearance of the mutations and identify sub-populations carrying different mutations, we performed 378 short PCRs on 87 isolated clones targeting the mutation loci, sequenced the amplified fragments to test for the presence of the mutations, or tested for deletions or insertions by gel electrophoresis (see Supplementary Data).

In total, we detected 23 mutations on 15 chromosomal loci, of which 10 and 11 on the endpoint clones, in line with other evolution experiments with similar number of generations (Kahn et al. 2011; Bachmann et al. 2012). Only 4 point mutations were found in genes, three of which were non-synonymous. Remarkably (table 1), large events such as deletions (2), inversions (1), duplications (1) or IS transpositions (6) balanced in number small events like SNPs (11) and indels (2) ($< 30\text{bp}$), and seem to appear early on during the experiment. To our knowledge, this feature is not common to other experiments of this kind (Barrick et al. 2009; Wang et al. 2010; Tenaillon et al. 2012). Together, these data allowed to reconstruct a partial tree of the evolution in the experiment (figure 2). To this end, an important feature is the detection of multiple co-occurring insertion sequence (IS) integration events just before 1000 generations and concomitant with the disappearance of an IS present between ACIAD0290 and ACIAD0291 in ABWT (but not in ADP1). These two types of changes may be related to a single mutational event like transposition of the latter IS to a new target. However, these changes may also involve two different mutational events. This target locus was variable among different clones as five different loci were identified. This gave birth to five detectable branches (named B1, B2, B3, B4 and B5). This IS insertion was used as a genomic marker to estimate the ratio of the five sub-populations in the sequencing of sampled populations.

The dynamics of branch B2 is remarkable. Figure 2 shows the near extinction of the branch over about 1000 generations: at generations 980 and 1100, one isolated clone out of ten be-

¹ We adopt the following notations: clones/populations are labelled by the string AB followed by the numbers G and i, where G is the number of generations undergone by the clone or population and i is a tag to identify an isolated clone

| Position on the ADP1 Chromosome | Gene - region | Product | Mutational event | Comment | G=1900 | G=2300 | G=2800 |
|---------------------------------|--|--|-----------------------------------|--|--------------|--------------|--------------|
| SNP | | | | | | | |
| 244223 | glnK (a) | Regulatory protein, for nitrogen assimilation by glutamine synthetase | C > T | 41 bp upstream of the start codon | ND | 20 % | 41 % |
| 389458 | citA (a) | Citrate proton symporter | T > C | Codon frequencies GAA (0,041) > GAG (0,014) | ND | ND | 29 % |
| 1020409 | ACIAD1029 (b) | Putative lipoprotein | T > C | 100 bp upstream of the start codon | ND | ND | 23 % |
| 1903797 | ACIAD1911 | Putative nitrate transporter transmembrane protein | G > A | 12 bp upstream of the start codon | ND | ND | ND |
| 1903816 | ACIAD1911 (a, b) | Putative nitrate transporter transmembrane protein | C > T | 31 bp upstream of the start codon | 100 % | 100 % | 100 % |
| 1903844 | ACIAD1911 | Putative nitrate transporter transmembrane protein | G > A | 59 bp upstream of the start codon | ND | ND | ND |
| 1903876 | ACIAD1911 | Putative nitrate transporter transmembrane protein | A > T | 91 bp upstream of the start codon | ND | ND | ND |
| 2042174 | ACIAD2054 (a, b) | Conserved hypothetical protein | T > C (S > P) | Non synonymous mutation in a gene yet cut by the inversion event | 26 % | 65 % | 91 % |
| 2100627 | ACIAD2113 | conserved hypothetical protein | A > G | 239 bp upstream of the start codon | 3 % | 41 % | 6 % |
| 2790910 | ftsH | Cell division protein | G > A (H > Y) | Non synonymous | 33 % | 57 % | 10 % |
| 2891560 | qseB (a, b) | Quorum sensing DNA-binding response regulator | G > A (S > F) | Non synonymous | 100 % | 100 % | 100 % |
| Small indels | | | | | | | |
| 1920605 | ACIAD1931 | Putative magnesium citrate secondary transporter | Ins A | 32 bp upstream of the start codon | ND | ND | 12 % |
| 1249430 | csrA (a, b) | Carbon storage regulator | Δ8 bp | Last 15 amino acids (21% of the protein) modified | 100 % | 100 % | 100 % |
| Large rearrangements | | | | | | | |
| 1230681 | ACIAD1230 | putative transcriptional regulator | IS insertion | 79 bp upstream of the start codon (B1 branch) | 43 % | 32 % | 43 % |
| 1230766 | ACIAD1230 | putative transcriptional regulator | IS insertion | Within gene (B5 branch) | 9 % | 7 % | ND |
| 1230919 | ACIAD1230 | putative transcriptional regulator | IS insertion | Within gene (B3 branch) | 32 % | 53 % | 10 % |
| 1231207 | ACIAD1230 (b) | putative transcriptional regulator | IS insertion | Within gene (B2 branch) | 1 % | 0,5 % | 47 % |
| 1860246 | ACIAD1864 - ACIAD1865 | putative transcriptional regulators | IS insertion | Between genes (B4 branch) | 16 % | 8 % | ND |
| NA | (a, b) | NA | IS insertion | Within a DNA sequence not found in A.baylyi ADP1 | 100 % | 100 % | 100 % |
| 1848721-1858316 | ACIAD1848 ACIAD1861 (a, b) | putative phage replication initiation factor, putative phage-related protein, putative phage replication initiation factor | duplication | Spans about 10 kb | 100 % | 100 % | 100 % |
| 2041592-2046126 | ACIAD2054, gltL, gltK, gltJ, gltI (a, b) | Conserved hypothetical protein and glutamate/aspartate transport proteins | Inversion (4534 bp) | Cuts gltI and ACIAD2054; quenches glt operon | 100 % | 100 % | 100 % |
| 2046132-2046451 | gltI (b) | glutamate/aspartate transport protein | Δ319 bp (named ΔgltI in the text) | Deletion in a gene yet cut by the inversion event | ND | ND | 51 % |
| 2664637-2726455 | ACIAD2713-ACIAD2778 (a, b) | NA | Δ60 kb | 58 CDS deleted | 100 % | 100 % | 100 % |

Table 1—Mutations detected by analysis of sequencing data. List of detected mutations occurring in the experiment, separated by type (SNPs, indel and large rearrangements). The last three columns indicate the frequency of each mutation at the three different sequencing times (G stands for the number of generations). 100 % frequencies are marked in boldface. NA: not applicable. ND: not detected (<1%). The letters (a) and (b) in the second column indicate if the mutation was detected in the endpoint clones AB2800a and AB2800b respectively. Compared to other laboratory evolution experiments, there is a relevant number of large rearrangements, many of which appear early on and reach very high frequencies.

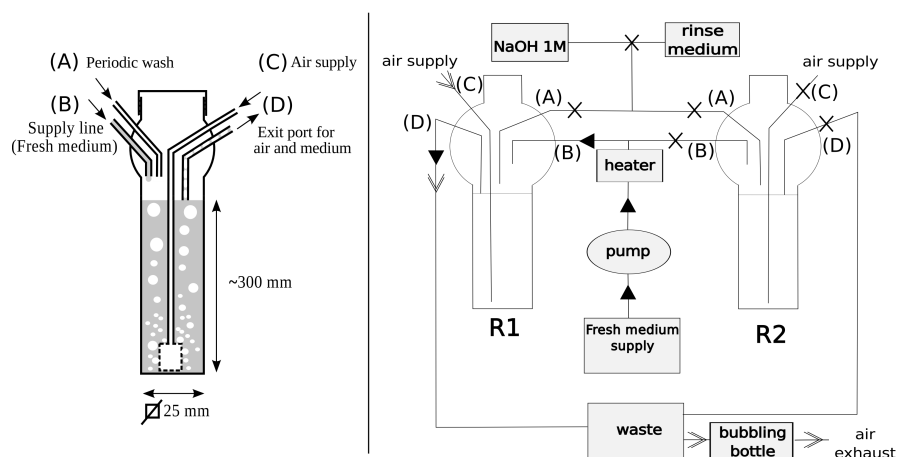


Fig. 1.—Left: Schematics of a single reactor during the culture stage. The reactor has been assembled by glass blowing, and incorporates a square cross-section glass tube in the bottom portion, and a screwed plastic cap on top. The entry and exit tubes are welded to the top of the reactor, and comprise: an air-supply connected to a tube terminated by a porous glass plug at the bottom of the reactor (generating bubbles that give both aeration and liquid mixing by air-lift), (C), a fresh medium supply (B), a wash fluid entry (A), and a fluid exit, a biphasic mixture of air and liquid (D). Right: schematics of the tubing connections during the culture phase in the reactor R1. R1 and R2 are the two reactors of the set-up; each is temperature regulated; filtered air (rate adjusted by a flow-meter/needle valve) is injected through port (C) of reactor R1. It passes through a porous glass plug and generates bubbles serving for both culture aeration and liquid mixing via air lift within the reactor. The bottle containing the fresh medium is connected to a peristaltic pump; a constant flow of medium arrives at port (B) of reactor R1. A heating resistance heats the medium before the entry (B), to prevent backward contamination of the supply of medium; The reactor wash solution (connected to port (A), and not used at this stage of operation of R1) comprises NaOH 1M and rinse medium (the same solution as the supply of medium to the reactor); A waste bottle is connected to port (D) of the reactor R1. It receives a biphasic mixture of air and medium exiting from the reactor by a common path, and traps the liquid; the air exiting from the waste bottle is bubbled through bleach in a bubbling bottle, so as to avoid lab air contamination. The full arrows indicate the circulation of liquids and the open arrows indicate the circulation of air. Crosses indicate the pinched-closed sections of tubes. Every week, the content of the operating reactor is transferred to the already cleaned tandem reactor. During the transfer phase (e.g. from R1 to R2), port (D) of R1 is closed and ports (A), (B), (C) are open; ports (B) and (C) of R2 are closed and ports (A) and (D) are open. Air and medium containing bacteria come from R1 to R2 via ports (A). When the transfer is stopped, R2 enters culture phase, and R1 is rinsed and washed, ready for the next transfer.

longs to B2 (data not shown), while at generations 1900 and 2300 the sequencing on sampled populations indicates that the B2 branch only occupies about 1% of the population. However, the same branch uprose to occupy nearly half the population at generation 2800. Two mutational events appeared after 2300 generations in the B2 branch and appear associated with this uprise: a SNP upstream of ACIAD1029 and a deletion of 319 bp in the *gltI* gene (Δ *gltI*). The frequency of these events at the end of the experiment are 23% and 51% respectively. Comparing these frequencies to 47% of the B2 branch at the end of the experiment, we conclude that essentially all B2 bacteria carried the deletion which is likely linked to a gain in fitness allowing the B2 branch to increase in relative proportion. This deletion event did not appear in the remaining of the population, suggesting that this event might be rare and/or dependent on the mutational background. Notably, the locus of this deletion appears to be highly dynamic during the whole experiment as it involves two other different mutational events assumed to give a relevant advantage at two other different

times: (i) an early inversion of 4534 bp, (ii) a point mutation (ACIAD2054) shared by several branches, with a frequency increasing from 26% to 91% from generation 1900 to generation 2800. However, it must be noted that also some genetic instability may increase the mutation frequency at that specific locus, and that in principle non-transitive interactions between sub-populations and negative frequency-dependent selection cannot be ruled out.

In brief, the combination of clonal and population sequencing with systematic mutation detection by PCR points to a complex scenario with highly dynamic coevolving sub-populations, even in the later stages of the experiment, where the main metabolic adaptation should be already accomplished.

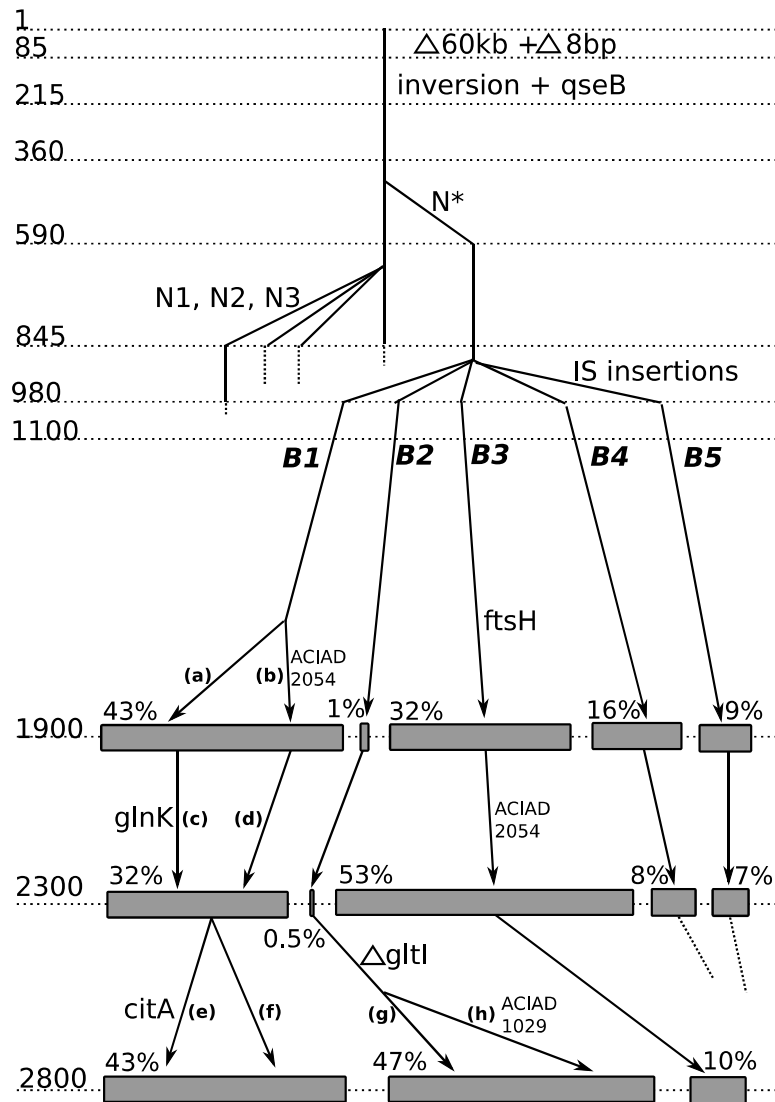


Fig. 2.—Inferred evolutionary tree of the population, deduced by combining genome sequencing data analysis and systematic sequencing of PCR fragments performed on 87 isolated clones. The numbers on the left side indicate generations, estimated from the imposed dilution rate. Each vertical line corresponds to a sub-population (not all sub-populations are displayed). Some relevant mutations are indicated at the time of appearance. N*, N1, N2, and N3 designate C1903816T, C1903797T, C1903844T, C1903876T respectively (table 1). Five branches arise before 1000 generations and 3 branches (B1, B2, B3) are still present at the end of the experiment. For the last three time marks, genomes extracted from samples of population were resequenced; the sizes of the gray rectangles indicate the estimated relative population percentage of each branch (% values indicated immediately above rectangles) as deduced from genomic data analysis. The labels (a) through (f) designate specific sub-populations as follows: (a) branch B1 with no additional mutation, (b) branch B1 with an additional ACIAD2054 mutation, (c) = (a) with mutated glnK, (d) = (b) with no additional mutation, (e) additional mutation in citA, (f) no additional mutation in citA, (g) deletion in gltI, (h) deletion in gltI, plus ACIAD1029 mutation.

The dynamics of maximum growth rate is compatible with the advantage estimated from population sequencing.

During the experiment, the dilution rate was typically kept constant² between 0.6 and 0.9 h⁻¹. However, the dilution rate was sometimes transiently increased above the wash-out threshold to measure the maximum growth rate μ_m of the population, by real-time monitoring of optical density during the wash-out. We assume that this transient increase in dilution rate did not interfere heavily with the population structure because it never washed out more than half of the population, and lasted for about one hour only³. We also measured μ_m for isolated clones in batch, by fitting the growth curve during the exponential phase. The results are plotted in figure 3. The maximum growth rate in our experiment initially increased rapidly, while after about 500 generations it became compatible with a much slower linear increase, or a plateau. A similar feature appears in other evolution experiments (Barrick et al. 2009; Chou et al 2011).

Thus, the maximum growth rate had little increase in the later stages of the experiment, where the described examples point to a highly dynamic situation. The population sequencing data gave further clues on the population dynamics occurring between generations 1900 and 2800. Some mutation frequencies increase while others decrease, the ratio of sub-population frequencies not being constant. This indicates that competitions between sub-populations still take place during the later stage of the experiment. If one assumes that the frequency $f(t)$ of a mutation is related to the selection coefficient s by $f(t) = f(0)e^{st}$, s can be roughly estimated by comparing the frequency observed at different generations. Following this reasoning, we obtained seven estimated values of s , laying between 0.001 and 0.01 per generation (table 2). If one uses the fit of μ_m vs. number of generations, one estimates values of s between 0.0075 and 0.0094 per generation, which are of the same order as the values deduced from sequencing. While both estimates are rough, their agreement suggests that a small increase in μ_m may account for the fairly high population dynamics observed during the last 2000 generations. It also suggests that μ_m might be a reasonable proxy estimate of fitness. This result is not completely surprising, as the bacteria were subjected to a regime of high dilution rate, with a dilution rate of about 80% of maximum growth rate during the last 2500 generations.

Notably, some values of s estimated from population sequencing data are still relatively high towards the end of the ex-

² The dilution rate was below 0.5 h⁻¹ during the first six days

³ One may wonder whether this procedure could remove completely some low-frequency mutants. The probability of losing a sub-population of n bacteria in a wash-out of half the population can be estimated as $1/2^n$. Let us consider a sub-population of 20 individuals (which corresponds to a frequency of about 10^{-9}), the probability of losing this sub-population is 0.0001%.

Table 2—Selection coefficient per generation estimated from re-sequencing data.

| Mutation | Selection coefficient |
|---------------|-----------------------|
| ACIAD2054 | > 0.0041 |
| ftsH | 0.0014 |
| ACIAD2113 | 0.0065 |
| glnk | > 0.0075 |
| Δ gltI | 0.0093 |
| citA | > 0.0067 |
| ACIAD1029 | > 0.0063 |

periment (e.g. the values of s for citA, Δ gltI and ACIAD1029, table 2). This suggests that (on average) s is not decreasing significantly during the late stages of the experiment. Consistently with this picture, the values of μ_m measured late in the experiment appear to be compatible with a slow increase with generation number.

Complementary contributions to fitness

Thus, variations in mutation frequencies suggest that the population is still highly dynamic during the last 2000 generations. Our rough estimates indicate that the increase in μ_m could be sufficient to capture the degree of adaptation during the experiment. Even though the ratio between the dilution rate and μ_m is about 0.8, we cannot exclude that other ways of adaptation could be operative, in view of the fact that the chemostat dilution rate is set below μ_m . In order to illustrate this aspect, let us compare two isolated clones, AB1100h and AB2800b. AB1100h was isolated and identified as belonging to the B2 branch. Its descendants are led to near extinction (between generation 1900 and 2300) but later acquire beneficial mutations that make them the largest sub-population present at the end (figure 2). It is thus expected that the fitness of AB2800b, an endpoint clone belonging to B2 branch, is larger than the fitness of AB1100h. Because of the uncertainties of the measurements of the maximum growth rate in batch of the two clones ($0.944h^{-1} \pm 0.028$ for AB1100h and $0.954h^{-1} \pm 0.035$ for AB2800b), it leaves open the possibility that other traits of selection might have been involved.

One targeted trait could be nitrate uptake: bacteria in a chemostat have to manage growth under limited nutrient availability. To measure how bacteria deal with nitrogen limitation, it is necessary to access their affinity for the limiting nutrient (nitrate). In order to do so, we grew a clone (isolated after 7 days) in a chemostat, let it reach an equilibrium for different dilution rates, and sampled extracts from the chemostat for each dilution rate. By a series of chemical reactions we were able to estimate the nitrate concentration in each sample (Supplementary Data). Plotting the nitrate concentration versus the dilution rate leads in principle to an estimate of the saturation constant. Unfortunately, the concentration was so low (<2 μ M) that we failed to detect nitrate for all the dilution rates but the highest (this value being likely above the

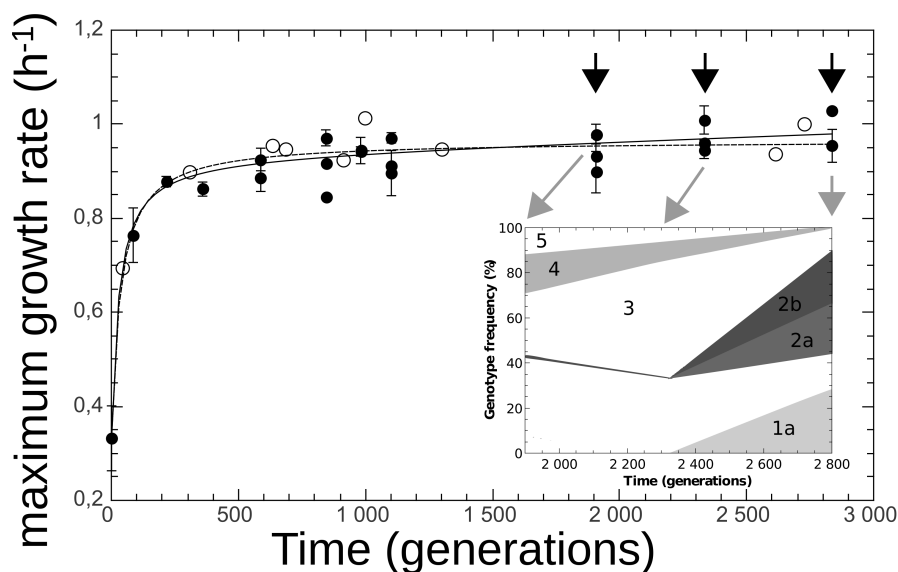


Fig. 3.—Maximum growth rate from batch (black circles) and chemostat (open circles) measurements, plotted as a function of time (in generations). Chemostat measurements were performed during the experiment by transiently increasing the dilution rate above the wash-out threshold (open circles). Lines are fits with a hyperbola (dashed: $f(t) = At/(B + t) + C$) and with hyperbolic plus linear function (solid: $g(t) = f(t) + Dt$), as in Barrick et al. 2009. The maximum growth rate initially increases rapidly, and drastically slows down after 500 generations approximately. The shaded area indicates the late stages of the experiment. Arrows indicates times of population sequencing. The inset plots the dynamics of frequencies related to some mutational events resulting from resequencing analysis in late stages of the experiment. Each number designates a branch identified by the IS insertion event. 1a, 2a and 2b designate clearly identified sub-populations validated by PCR on isolated clones. Mutations shared by several branches or not verified by PCR are not represented.

wash-out threshold).

Thus (i) an increase in nitrate uptake after seven days could not be measured in our experiment, but (ii) we can conclude that bacteria evolved rapidly within seven days (or were pre-adapted) to be able to capture >99.8% of the nitrate entering the chemostat, in a wide range of dilution rates. The following additional result is consistent with a good affinity for nitrate: measuring nitrate concentration at different times during growth in batch, we observed that cells grow exponentially until nitrate is no longer detectable, without a smooth transition between exponential and saturation phase (data not shown). To sum up, if nitrate uptake optimization gave a contribution to fitness, it did so only in the early stages of the experiment.

Adaptation might also consist in a better management of the low nitrogen level in the chemostat, i.e. the evolved cells might be able to grow using lower amounts of nitrogen per cell. For example, an increase in protein turnover has been observed in *Pseudomonas putida* grown under nitrogen limitation (Hervas et al. 2008). The comparative catabolic activities of the starting strain ABWT and of the endpoint clone AB2800a were determined by the Phenotype MicroArray Services of Bi-

olog, Hayward, CA. The two strains were tested on different media (see Supplementary Data). New phenotypes appeared, notably the endpoint clone was able to grow faster compared to ABWT in diverse media containing different nitrogen sources. Therefore, the Biolog assay gives evidences for adaptation, but it is not easy to connect these results to a specific parameter contributing to fitness.

Indirectly, adaptation in recycling of nitrogen compounds might possibly be evidenced by survival in conditions of low nutrients. In order to test this, we performed starvation experiments with the two clones AB1100h and AB2800b discussed above. Both clones were grown separately in batch in the selection medium, the end of exponential phase was defined as the beginning of starvation, and samples were then plated to estimate the number of living cells after starvation as a function of time. Starting from about $2 \cdot 10^8$ cfu/ml, about 10^7 cfu/ml were counted after 100 hours of starvation, but no significant differences were observed between the strains. In conclusion, the increase in maximum growth rate μ_m and new metabolic phenotypes are strong proofs of adaptation, but except for μ_m , we failed to identify additional relevant parameters encapsulating fitness.

Diversity in the population appears to increase slightly at late times

We first focus on evidence of population diversity during the experiment. After the transposition of the IS that occurs after 1000 generations, diversity is detected for the whole experiment, as three branches are still present at the end (figure 2). Each branch evolved with both specific mutations (citA for B1, ACIAD1029 and Δ gltI for B2, ftsH for B3), and mutations that are shared with other branches (glnK and ACIAD2054). The latter likely appeared independently in different branches and are not related to horizontal transfer: in control experiments, the transformation rate in the chemostat was too low to be detected (Supplementary Data). It is highly probable that diversity took place before 1000 generations, as observed in short-term evolution experiments with bacteria in minimal medium (Wang et al. 2010; Herring et al. 2006; Conrad et al. 2009; Chou et al. 2011). We lack sequencing data on populations in the early stages of the experiment. However, a clue of diversity is given by the following observation. PCR on clones at the 845th generation were performed to check for the mutation C1903816T. Four other different mutations have been observed in the same locus within a sequence of 80 bp (see table 1) in 17 clones. We failed to isolate a clone carrying mutation C1903816T at generation 845, but this allele was later the only one found in the population.

Is diversity constant? The diversity can be roughly estimated by the total number of mutations detected in the population at a given time. It is expected to increase as a function of the time if the number of different lineages increases during the evolution. The number of detected mutations is 15, 16 and 18 respectively at generations 1900, 2300 and 2800 and thus appears to rise slightly during 1000 generations. We can estimate the diversity more quantitatively by calculating the Shannon index $H = -\sum_i p_i \log p_i$ where p_i is the frequency of mutation i . Note that this is a loose use of this formula, which is a natural definition for frequencies of sub-populations, while here we can only access the frequencies of mutations. The index calculated for generations 1900, 2300 and 2800 are respectively 0.88, 1.04 and 1.34, suggesting a slight increase in diversity during the last 1000 generations.

It is difficult to judge whether this diversity is large or small. Assuming constant additive s , fairly standard population genetics theories (Desai et al. 2007, Brunet et al. 2008) would predict for our conditions the coexistence of order of 10 sub-populations of fitness classes. However, detailed information on linkage and precise measurements of fitness distributions would be necessary to connect fitness classes and mutation classes. Thus, given this large uncertainty, the observed diversity might well be in the predicted range.

Conclusion

To sum up, we performed a laboratory evolution experiment with bacteria growing in a minimal medium, and focused on its

later stages. Analysis sequenced populations at different times show that the number and type of mutations are compatible with other laboratory evolution experiments, characterized by competition between variant clones (Bachmann et al. 2012; Barrick et al. 2009). Interestingly, we found many fixed large rearrangements compared to point mutations.

Additionally, we found significant and persistent diversity in a population with presumably the same ecotype (Maharjan et al. 2006; Woods et al. 2011). We thus report for the first time that diversity can be established and maintained in a very simple medium for long times, when the proxies for fitness appear to saturate. In this regime, where "adaptation" might appear to have slowed down, we observe diversity that is stable or even increasing.

This observed long-term diversity can be due to different mechanisms, which were considered carefully in (Maharjan et al. 2012). While it is beyond our scopes to identify which of these mechanisms were in place during our experiment, we can produce some qualitative and hypothetical considerations based on our observations. Metabolic trade-offs (Beardmore et al. 2011) might be witnessed by the results of the biologic assays, where we observed changes in metabolic phenotypes that appeared *a priori* very distant from the expected target (Supplementary Figure B2). An example of evolutionary convergence (Tenaillon et al. 2012; Maharjan et al. 2012) possibly can be represented by the insertion of an IS upstream or within a putative transcriptional regulator (ACIAD1230), giving birth to four sub-populations (branches B1, B2, B3, B5). At the same time, a fifth sub-population (B4 branch) arose with the insertion of an IS between putative transcriptional regulators, in a completely different region (ACIAD1864-1865). The observation of a significant number of mutations in regulatory regions might also point towards a role of convergence.

Maharjan and coworkers (Maharjan et al. 2007) identified in their experiment an evolved sub-population of *E. coli* specialized in the metabolization of by-products of other cells. In our case, reduced forms of nitrogen or amino acids could be cross-fed, and then, a role of cross-feeding cannot be excluded. More in general, frequency-dependent selection could be in place. Its effect has been observed at low dilution rates, and could be ascribed to numerous ecological phenomena, including cross-feeding (Maharjan et al. 2012). In our experiment, the preservation of the B2 branch at low frequency during more than 400 generations could result from a frequency-dependent selection; this should be checked in future investigations.

Finally, in terms of basic parameters such as selection coefficient and population size, the results should be compatible with the regime where clonal interference (in the wide sense of competition between beneficial mutations) can be observed (Park and Krug 2007; Hughes et al. 2012). However, clonal interference necessarily has to coexist with convergence (availability of alternative evolutionary pathways), and with the observed "diminishing returns" epistasis mechanism (Kahn et al. 2011; Chou et al. 2011), according to which the advantage of

beneficial mutations depends on their order of appearance.

Overall, we speculate that a possible way that the sub-populations interact could be indirect, through the effect of the limiting amounts of nitrogen in the medium. We explicitly tested for contributions to fitness, which, if acquired by one sub-population would affect the others, such as uptake and resistance to starvation. However, we were not able to identify such an effect. On the other hand, our results suggest that the maximum growth rate μ_m measured in batch might be a sufficiently good proxy for fitness *per se* as compared to the selection coefficients estimated from resequencing data, indicating that the observed sub-populations structure appears consistent with a regime in which competition between beneficial mutations could be relevant. While both measurements are rough, we can surmise that if there are selected traits that are not recapitulated by μ_m , their associated advantage cannot be very large.

Materials and Methods

Bacterial strain and medium We used a strain of *Acinetobacter baylyi* obtained from NCIMB (<http://www.ncimb.com>, NCIMB 11826). The minimal medium (pH 6.8) used in the evolution experiment comprises 50 mM Potassium Phosphate, 1 mM Potassium nitrate, 3 mM Potassium Citrate, 9 mM Citric Acid, 1 mM Magnesium Sulphate, traces elements (3 μ M Iron (III) Chloride, 1 μ M Zinc Sulphate, 1 μ M Cobalt Chloride, 0.6 μ M Copper Sulphate, 1 μ M Manganese Sulphate, 1 μ M Sodium Molybdate, 1 μ M Manganese Chloride and 1 μ M Calcium Sulphate), and Polypropylene Glycol 2000 at dilution 1:20000 as anti-foaming agent. For the initiation of the evolution experiment, the ancestral strain (ABWT) was plated, a single colony was inoculated in the minimal medium and grown at 35°C until OD reaches 0.1-0.2. A volume of 50 ml of this culture was introduced in the reactor then completed by fresh medium. At various time intervals, samples were snap-frozen as stocks with 10% DMSO. Stocks were later plated, and clones isolated. They are named ABGi, where G is the number of generations undergone by the bacteria before sampling, and i is a letter used to identify a clone.

Experimental evolution in chemostat Two interconnected reactors were used allowing aseptic periodic transfer from one reactor to the other so that the reactor being not used for growth can be cleaned in situ, preventing long term biofilm formation and bacteria evading dilution (de Crécy-Lagard et al. 2001). Such transfer is made every week. The reactors (150 ml) are made of glass and have been assembled by glass-blowing. One entry tube is used for the medium supply with its dripping opening inside the reactor positioned above the liquid level of the culture. A compressed air input tube is plunging inside the reactor down to its bottom. Air bubbles provide for aeration and mixing of the growth medium. An exit tube allows for both air and liquid outflow. Another entry is allowed for the periodic

injection of a cleaning solution (1 M NaOH) within the reactor or for the transfer of the bacterial population before a reactor cleaning. The reactors are connected to a unique growth medium supply via a flow rate controlling peristaltic pump (Masterflex). Temperature was regulated using a commercial PID regulator (ENDA ETC 442), maintaining the temperature at 35 °C \pm 0.5 °C. During the first six days, the dilution rate was kept below 0.5 h⁻¹, in order not to lose the population. Subsequently, it was kept constant between 0.6 h⁻¹ and 0.9 h⁻¹ (the reasons for this variations were both practical and related to experiments testing the maximum growth rate occurring over the evolutionary time).

Maximum growth rate measurement The batch cultures were performed in a glass vessel presenting a lower part with a transparent square cross-section so that an in-situ optical density measurement could be performed. This system consisted in (i) a Light Emitting Diode (LED) positioned in close proximity to one flat face of the reactor and emitting light inside the reactor, (ii) a first photosensitive detector PD1 placed on the other side of the reactor and facing the light emitting diode and (iii) a second photodetector PD2 placed in close proximity of the light emitter. The current of the photodetectors PD1 and PD2 were fed to a logarithmic-ratio amplifier so as to give a voltage related to the optical density of the culture. The reference value was obtained when fresh medium was first injected. The system yielded optical density measurements taken approximately every minute. The maximum growth rate in the batch mode for a set of clones was estimated from those measurements.

Genomes resequencing We detected all events, including IS movements and DNA rearrangements, through genome sequencing. Genome resequencing of two endpoint clones were performed using Illumina platform by GATC-Biotech. The genome of the ancestral clone and the genomes of three sampled populations were resequencing using Illumina platform at GRCF HTS Center at Johns Hopkins University, Baltimore. Sequence alignment against the reference genome (*Acinetobacter baylyi* ADP1) was performed using Mozaik (<http://bioinformatics.bc.edu/marthlab>) or bowtie2 (Langmead et al. 2009). Visualization was performed using Tablet (Milne et al. 2010). Detection of mutation was performed using SAMtools (Li et al. 2009) and gigaBayes (<http://bioinformatics.bc.edu/marthlab>). The de novo genomic assembler velvet (Zerbino et Birney 2008) was also used. These tools are designed to extract allele frequencies, and thus also well adapted for the detection of mutation frequencies for mixed populations.

Supplementary Material

Supplementary data are available at Genome Biology and Evolution online (http://www.oxfordjournals.org/our_journals/gbe/).

Acknowledgments

We thank Dominique Schneider and Gilles Fischer for critical reading of this manuscript. This work was supported in part by a "Convergences" grant from Université Pierre et Marie Curie, Paris.

Literature Cited

- de Crécy-Lagard et al. 2001. Long term adaptation of a microbial population to a permanent metabolic constraint: overcoming thymineless death by experimental evolution of *Escherichia coli*. *BMC biotechnology* 1:10-16.
- Barbe V. et al. 2004. Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Research* 32:5766-5779
- Bachmann H, et al. 2012 Microbial domestication signatures of *Lactococcus lactis* can be reproduced by experimental evolution. *Genome Res.* 22:115-24.
- Barrick J.E. and Lenski R.E. 2009. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. *Cold Spring Harb Symp Quant Biol.* 74:119-129
- Barrick J.E. et al. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature.* 461:1243-1247
- Beardmore R.E., Gudelj I., Lipson D.A. and Hurst L.D. 2011. Metabolic trade-offs and the maintenance of the fittest and the flattest. *Nature.* 472:342-346
- Brunet, E., Rouzine, I.M., and Wilke, C.O. 2008. The stochastic edge in adaptive evolution. *Genetics* 179:603-620
- Buckling A., Maclean R.C., Brockhurst M.A. and Colegrave N. 2009. The Beagle in a bottle. *Nature* 457:824-829
- Chou H.H., Chiu H.C., Delaney N.F, Segra D. and Marx C.J. 2011. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 332:1190-1192
- Conrad T.M. et al. 2009. Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biol.* 10:R118
- Conrad T.M., Lewis N.E. and Palsson B.O. 2011. Microbial laboratory evolution in the era of genome-scale science. *Molecular Systems Biology* 7; Article number 509; doi:10.1038/msb.2011.42
- Desai M.M. and Fisher D.S. 2007. Beneficial Mutation-Selection Balance and the Effect of Linkage on Positive Selection. *Genetics* 176:1759-98
- Desai M.M., Fisher D.S. and Murray A.W. 2007. The Speed of Evolution and Maintenance of Variation in Asexual Populations. *Current Biology* 17:385-94
- Doebeli M. and Ispolatov I. 2010. Complexity and diversity. *Science.* 328:494-497.
- Friesen ML, Saxer G, Travisano M, Doebeli M. 2004. Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in *Escherichia coli*. *Evolution* 58:245-260.
- Gause G.F. 1934. *The struggle for existence*. Baltimore, The Williams and Wilkins company.
- Gerrish P.J. and Lenski R.E. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica.* 102/103: 127-144
- Good, B.H., Rouzine I.M., Balick D.J., Hallatschek O., and Desai M.M. 2012. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proc. Natl. Acad. Sci.* 109:4950-4955
- Gresham D. et al. 2008. The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet* 4:e1000303. doi: 10.1371/journal.pgen.1000303
- Hardin G. 1960. The Competitive Exclusion Principle. *Science.* 131:1292-1297
- Herring C.D. et al. 2006. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nature Genetics.* 38:1406-1412
- Hervas A.B., Canosa I., and Santero E. 2008. Transcriptome Analysis of *Pseudomonas putida* in Response to Nitrogen Availability. *J. Bacteriol.* 190:416-420
- Hindré T., Knibbe C., Beslon G. and Schneider D. 2012. New insights into bacterial adaptation through in vivo and in silico experimental evolution. *Nature Rev. Microbio.* 10: 352-365
- Hughes JM, et al. 2012. The role of clonal interference in the evolutionary dynamics of plasmid-host adaptation. *mBio* 3(4):e00077-12. doi:10.1128/mBio.00077-12.
- Jackson RW, Johnson L.J., Clarke S.R. and Arnold D.L. 2011. Bacterial pathogen evolution; breaking news. *Trends Genet.* 27: 32-40
- Khan A.I., Dinh D.M., Schneider D., Lenski R.E. and Cooper T.F. 2011. *Science.* 332:1193-1196
- Langmead B, Trapnell C, Pop M, and Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25.
- Levin B.R., Antonovics J. and Sharma H. 1988 Frequency-dependent selection in bacterial populations (and discussion). *Phil. Trans. R. Soc. Lond. B.* 319:459-472.
- Li H. et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 25:2078-2079.
- Maharjan R., Seeto S., Notley-McRobb L. and Ferenci T. 2006. Clonal Adaptive Radiation in a Constant Environment. *Science* 313:514-517
- Maharjan R.P., Seeto S. and Ferenci T. 2007. Divergence and redundancy of transport and metabolic Rate-Yield strategies in a single *Escherichia coli* population. *J. Bact.* 189:2350-2358
- Maharjan R.P. et al. 2012. The multiplicity of divergence mechanisms in a single evolving population. *Genome Biology* 13:R41
- Milne I. et al. 2010. Tablet-next generation sequence assembly visualization. *Bioinformatics.* 26:401-402
- Park, S.C., and Krug, J. 2007. Clonal interference in large populations. *Proc Natl Acad Sci U S A.* 104:18135-18140.
- Rosenzweig R.F., Sharp R.R., Treves D.S., and Adams J. 1994. Microbial evolution in a simple unstructured environment: Genetic differentiation in *Escherichia coli*. *Genetics* 137:903-917.
- Tenaillon O. et al. 2012. The molecular diversity of adaptive convergence. *Science* 335:457-461
- Toft C. and Andersson S.G. 2010. Evolutionary microbial genomics; insights into bacterial host adaptation. *Nature Rev Genet.* 11:465-475
- Wang L. et al. 2010. Divergence involving global regulatory gene mutations in an *Escherichia coli* population evolving under phosphate limitation. *Genome Biol. Evol.* 2:478-487
- Woods RJ, et al. 2011. Second-order selection for evolvability in a large *Escherichia coli* population. *Science* 18;331(6023):1433-6
- Zerbino D.R. and Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research.* 18:821-829

5.2 Expériences d'évolution avec ou sans brassage de génomes (sem et lud)

5.2.1 Déroulement des expériences d'évolution avec ou sans brassage (sem et lud)

Pour comparer l'évolution avec ou sans brassage, on a utilisé deux souches différentes : l'une compétente, l'autre non-compétente. Pour l'expérience avec brassage de génomes, la souche d'*Acinetobacter baylyi* est la même que pour l'expérience d'évolution pilote (ABWT) tandis que pour l'expérience d'évolution sans brassage, cette souche a été rendue non compétente en délétant l'opéron comFECB (ACIAD3314 à 3318) essentiel pour la compétence [4]. Cette délétion de l'opéron comFEBC a été réalisée par transformation de notre souche ABWT avec une cassette de résistance à la kanamycine qui possède à ses extrémités des séquences homologues à comF ou comC.

Les souches compétentes et non compétentes ont été cultivées dans les mêmes conditions avec le montage et le protocole décrit précédemment dans les sections 4.2 et 4.3. Le taux de dilution a été progressivement augmenté de $0,1h^{-1}$ à $1h^{-1}$ sur les trente premières générations environ puis il a été maintenu à $1h^{-1}$ (en dehors des phases de transformation et de réinjection dans le réacteur).

Lors de l'expérience d'évolution avec brassage, des déterminations de taux de transformation avec de l'ADN génomique comportant une cassette de résistance à la kanamycine ont été réalisés. On a observé une baisse de la compétence au cours du temps (voir tableau 5.1) qui nous a conduit à arrêter les deux expériences au bout de 1300 générations.

| générations | 0 | 150 | 500 | 664 | 830 | 995 | 1160 | 1323 |
|------------------------|---|---------------------|---------------------|-------------------|-------------------|---------------------|---------------------|---------------------|
| taux de transformation | $1,5 \cdot 10^{-6} \pm 8 \cdot 10^{-7}$ | $1,6 \cdot 10^{-6}$ | $2,8 \cdot 10^{-7}$ | $2 \cdot 10^{-6}$ | $1 \cdot 10^{-6}$ | $4,4 \cdot 10^{-7}$ | $5,8 \cdot 10^{-8}$ | $2,3 \cdot 10^{-8}$ |

Tableau 5.1 – Evolution du taux de transformation au cours de l'expérience d'évolution avec brassage de génomes (sem). Les taux de transformation ont été estimés à l'aide de tests de transformations avec de l'ADNg KanR (voir chapitre 3)

Lors de l'expérience d'évolution sans brassage, on a observé l'apparition de biofilms récurrents dans le chemostat, nous conduisant à augmenter la fréquence des transferts d'un réacteur à l'autre. À 1300 générations, lorsqu'un prélèvement de population a été étalé sur boîte, des colonies avec des phénotypes différents sont apparues : la plupart étaient bien rondes et opaques (la forme "classique" d'une colonie d'*Acinetobacter baylyi*) mais de rares colonies étaient plus plates avec un anneau diffus moins épais autour d'un centre opaque. Des colonies diffuses peuvent être dues à de la "twitching motility" qui est un mode de déplacement des bactéries sur les surfaces (agar, verre

...) [55,91]. Ce mode de déplacement dépend des pili de type IV et est favorisé en milieu humide. De plus il est impliqué dans la formation de biofilm . *Acinetobacter baylyi* est connue pour être capable de faire de la "twitching motility". Un clone formant une colonie diffuse (lud1300t) et une colonie de forme "classique" (lud1300a) ont été isolés.

Pour voir si cette différence de phénotype est liée à la capacité de former des biofilms, les deux clones ont été cultivés en milieu minimum liquide soit avec agitation soit sans agitation. Dans la culture statique, lud1300t forme des agrégats au fond du tube et un biofilm remonte au dessus du niveau de la culture alors que lud1300a ne forme ni biofilm ni agrégat au bout de 24h. Dans la culture avec agitation, lud1300t forme des agrégats et un biofilm est visible sur les parois du tube au niveau de l'interface du milieu avec l'air. Pour lud1300a, la culture avec agitation est homogène et on n'observe pas de biofilm.

Pour tester si la capacité à former des biofilms est liée à une différence de vitesse de déplacement par "twitching", une goutte de culture de lud1300a et lud1300t sont déposées au centre de boîtes de milieu minimum. Après 15 heures d'incubation à 35°C en atmosphère humide, le diamètre de la colonie lud1300t est de 60 mm en moyenne contre 25 mm pour lud1300a. L'aspect des colonies est plat et diffus dans les deux cas mais sont de formes différentes : lud1300a forment des anneaux de croissance concentriques tandis qu'on observe une croissance étoilée pour lud1300t. Enfin, la comparaison des deux ancêtres (compétent et non compétent) montre que l'ancêtre non compétent semble légèrement plus rapide pour la formation de zone diffuse.

En résumé, la délétion de comFECB pour rendre *Acinetobacter baylyi* non-compétente semble avoir des effets secondaires sur le phénotype de la souche, en plus de la non-compétence. Cette non-compétence a été recherchée pour éviter tout transfert horizontal, mais le revers est une évolution résultante qui a pu dépendre des phénotypes secondaires associés à la délétion de comFECB.

5.2.2 Séquençage de génomes de clones et de populations

Toutes les semaines soit environ toutes les 165 générations, une extraction d'ADN génomique de la population est réalisée en vue de faire le brassage de génome le lendemain (voir 4.3) et des clones sont isolés de cette même population. Les populations de l'expérience sans brassage sont nommées ludX où X est le nombre de générations tandis que les populations issues de l'expérience avec brassage de génome sont nommées semX. Les clones isolés sont distingués par une lettre minuscule : par exemple sem1300a est un clone isolé de la population sem1300. Toutes les 330 générations environ, l'ADN génomique d'une population et d'un clone (deux pour lud1300) ont été purifiés pour du séquençage. Ces échantillons ont ensuite été envoyés à séquencer à la plateforme Illumina de l'Université John Hopkins à Baltimore. Les populations et les clones ont été séquencés sur 100 bases en "paired-end", c'est à dire à chaque extrémité et ce sur des fragments d'ADN génomique de 400 bases environ. Le taux de couverture moyen des génomes est de 1500 X sauf pour sem1300 et sem1300a qui ont été séquencés en "single-end" avec une couverture de 600 X.

L'alignement des fragments sur le génome de référence d'*Acinetobacter baylyi* sp.

ADP1 est réalisé par bowtie2 [78,79] ou Mozaik (<http://bioinformatics.bc.edu/marthlab>). SAMtools [84] permet de détecter les mutations tandis que Tablet [96] permet une visualisation de l'assemblage des fragments alignés sur la séquence de référence. Bien que SAMtools soit conçu pour détecter des SNPs et des Indel, les réarrangements de type insertion, délétion et inversions peuvent aussi être détectés car ils entraînent des séries de mésalignements de part et d'autre du réarrangement. Une visualisation avec Tablet permet d'identifier les séquences de part et d'autre des régions de mésalignements et l'assemblage de novo réalisé par Velvet [139] permet de rechercher les raccords de ces fragments avec des mésalignements. Blast permet d'aligner ce raccordement sur le génome de référence et permet ainsi d'identifier le type d'évènement à l'origine de ce mésalignement.

La fréquence des mutations dans la population a été calculée de deux façons. Pour les SNPs détectés par SAMtools, la fréquence des mutations peut être calculée à partir du fichier de sortie de SAMtools car pour chaque position, le nombre d'allèles de référence et d'allèles mutés sont fournis pour chaque orientation ("forward" et "reverse"). Pour les réarrangements (délétion, insertion d'IS et inversion) et les mutations connues à faible fréquence, les fréquences des mutations sont estimées d'après le nombre de fragments porteurs d'un motif spécifique de la mutation et le nombre de fragments porteurs d'un motif non présent chez les porteurs de cette mutation. Ces motifs sont recherchés dans les deux orientations afin d'avoir une meilleure statistique.

Les duplications et le nombre de copies d'IS sont identifiés grâce au taux de recouvrement des contigs obtenus lors de l'assemblage de novo.

Dans l'expérience de Lenski [8], les séquençages de populations ont été réalisés avec une couverture moyenne de 50x et leurs outils bioinformatiques sont rapportés pouvoir détecter des mutations ayant une fréquence de plus de 4%. Dans notre cas, d'après le taux de couverture plus élevé, on pourrait potentiellement détecter des mutations présentes à plus de 0,1%. Cependant, les programmes obtenus de Barrick et al. [8] ne semblent pas compatibles avec le séquençage sur 100 bases dont nous disposons. Les outils informatiques que nous avons finalement utilisés ne détectent pas les mutations à moins de 15 % environ. Par conséquent, notre recherche des mutations à plus faibles fréquences est biaisée. On sous estime donc la diversité mais contrairement à [8], cette sous estimation est biaisée par le fait qu'on recherche spécifiquement les mutations connues (lorsqu'elles ne sont pas détectées par SAMtools).

5.2.2.1 Résultats des séquençages de génomes dans l'expérience d'évolution sans brassage de génome (lud)

Les résultats des séquençages de génomes des populations et des clones à différents temps de l'expérience sans brassage de génome (lud) sont présentés sous forme de tableau dans la figure 5.1. Le détail sur ces mutations est présenté dans le tableau de la figure 5.2. Les évolutions des fréquences de certaines mutations sont représentées graphiquement dans la figure 5.3.

Trois mutations sont fixées avant 310 générations :

- la délétion d'une séquence de 40 kb non présente dans le génome de référence d'*Acinetobacter baylyi* sp. ADP1. Cette séquence a été identifiée dans notre souche ancestrale non modifiée ABWT (voir section 5.1 et annexes)
- une délétion de 11b en amont d'un gène codant pour un transporteur secondaire du citrate (qui est la seule source de carbone)
- la mutation 3410754 C>T qui est entre les gènes ACIAD3483 et 3484.

| mutations | lud310a | lud310 | lud640a | lud640 | lud970a | lud970 | lud1300a | lud1300t | lud1300 |
|---|---------|---------------|---------|---------------|---------|---------------|----------|----------|---------------|
| Délétion séquence non présente chez ADP1 | oui | oui | oui | oui | oui | oui | oui | oui | oui |
| délétion 11pb 1920624-1920634 | oui | 98,80% | oui | 99,80% | oui | 98,70% | oui | oui | 99,90% |
| 3410754 C>T | oui | 97,50% | oui | 99,30% | oui | 96,70% | oui | oui | 99,90% |
| 2333167 T>C | oui | 92,00% | oui | 89,70% | non | 14,10% | non | non | 0,30% |
| Délétion 10666 pb avant IS4 (1237527-1248193) | oui | 3,10% | non | 2,20% | non | 0,50% | non | non | non |
| 2332965 G>A | non | 2,00% | non | 6,40% | oui | 81,40% | oui | non | 95,20% |
| 2891078 A>G | non | 0,50% | oui | 77,50% | non | 9,30% | non | non | 1,00% |
| 1871062 T>C | non | non | non | 6,30% | oui | 88,40% | oui | non | 97,10% |
| 3410537 T>C | non | non | non | 6,60% | oui | 88,30% | oui | non | 98,00% |
| 99634 C>T | non | non | non | non | non | 0,30% | non | oui | 1,00% |
| 818809 T>C | non | non | non | non | non | 1,50% | oui | non | 19,20% |
| 1839168 A>G | non | non | non | non | oui | 0,30% | non | non | non |
| 2630686 T>C | non | non | non | non | non | 0,30% | non | oui | 0,90% |
| 2802047 A>G | non | non | non | non | non | 0,50% | non | oui | 0,70% |
| 3410592 T>C | non | non | non | non | non | 2,70% | non | non | 55,30% |
| 3410598 A>G | non | non | non | non | non | 0,70% | oui | non | 34,70% |
| Délétion 451 pb avant IS4 (1247742-1248192) | non | non | non | non | non | non | oui | non | 3,30% |
| 1903816 C>T | non | non | non | non | non | non | non | non | 4,90% |
| 3289591 G>A | non | non | non | non | non | non | oui | non | non |
| 3438166 A>G | non | non | non | non | non | non | non | oui | 0,50% |

Figure 5.1 – Mutations identifiées chez des clones et des populations dans l'expérience d'évolution sans brassage de génomes (lud). Les fréquences en gras ont été estimées manuellement d'après le rapport entre le nombre de fragments porteurs d'un motif spécifique de la mutation, et le nombre de fragments porteurs d'un motif non présent chez les porteurs de cette mutation. lud310a, 640a 970a, 1300a et 1300t sont des clones isolés aux générations 310, 640, 970, et 1300. Lud1300t est un mutant ayant un phénotype de formation facilitée de biofilm, décrit précédemment. Lud310, 640, 970, 1300 sont des fractions de populations totales aux générations 310, 640, 970 et 1300

Trois mutations (1871062 T>C, 3410537 T>C, 2332965 G>A) qui apparaissent entre 310 et 640 générations ont des fréquences similaires dans les populations à 640, 970 et 1300 générations. Les mutations 1871062 T>C et 3410537 T>C ont des fréquences identiques (les courbes d'évolution de leur fréquence se superposent) et sont retrouvées ensembles chez les clones lud970a et lud1300a. Ceci suggère que ces mutations sont liées. La courbe d'évolution de la fréquence de la mutation 2332965 G>A se distingue de celles des deux autres mutations car les fréquences de la mutation 2332965 G>A sont significativement inférieures à celle des deux autres dans les population lud970 et

lud1300. Cette mutation semblerait donc ne pas être présente chez tous les porteurs des mutations 1871062 T>C et 3410537 T>C.

| position | mutation | type de mutation | gène | fonction |
|-----------------|----------|---|------------------------------------|--|
| 99634 | C>T | CAA>TAA (STOP) | ACIAD0101 (pgi) | glucose-6-phosphate isomerase |
| 818809 | T>C | Y429H (TAT>CAT) | ACIAD0830 (cafA) | ribonucléase G, endoribonucléase G |
| 1237528-1248192 | Δ1066 pb | Délétion 10666 pb avant IS4 (en phase -1) | ACIAD1238 à 1248 | |
| 1247742-1248192 | Δ451 pb | Délétion 451 pb avant IS4 | ACIAD1248 (rnhB) | Ribonucléase HII |
| 1839168 | A>G | F217L (TTC>CTC) | ACIAD1831 | protéine hypothétique conservée |
| 1871062 | T>C | K127R (AAA>AGA) | ACIAD1878 (aro) | 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase (DAHP synthetase) |
| 1903816 | C>T | 31 pb avant premier gène d'un opéron | ACIAD1911 | transporteur transmembranaire du nitrate, premier gène de l'opéron codant aussi pour les nitrate et nitrite réductases |
| 1920624-1920634 | Δ11 pb | Entre 2 gènes (phase -2 / phase +1) (- 51 / -198) | ACIAD1931 / ACIAD1933 | transporteur secondaire du citrate et du magnésium / protéine hypothétique |
| 2332965 | G>A | S38F (TCT>TTT) | ACIAD2373 (hfq) | Hfq (host factor I for bacteriophage Q beta replication), régulateur du métabolisme des ARN |
| 2333167 | T>C | entre 2 gènes (phase -1 / phase -1) (-90 / +1) | ACIAD2373 (hfq) / ACIAD2374 (miaA) | Hfq / tRNA delta(2)-isopentenylpyrophosphate transferase |
| 2630686 | T>C | S20G (AGC>GGC) | ACIAD2675 | putative sous unité d'oxygénase |
| 2802047 | A>G | L1425P (CTT>CCT) | ACIAD2866 | putative hémagglutinine / hémolysine |
| 2891078 | A>G | V27A (GTC>GCT) | ACIAD2960 (qseC) | histidine kinase, senseur d'un système de régulation transcriptionnelle à deux composants (qseB - qseC) (quorum sensing) |
| 3289591 | G>A | P112L (CCG>CTG) | ACIAD3379 (czcl) | protéine du système d'efflux des cations |
| 3410537 | T>C | D61G (GAT>GGT) | ACIAD3483 | protéine hypothétique |
| 3410592 | T>C | R43G (AGA>GGA) | ACIAD3483 | protéine hypothétique |
| 3410598 | A>G | F41L (TTT>CTT) | ACIAD3483 | protéine hypothétique |
| 3410754 | C>T | Entre 2 gènes (phase -2 / phase-3) (-360 / +320) | ACIAD3483 / ACIAD3484 | protéine hypothétique/ protéine de liaison au substrat de l'ABC transporteur de l'argine / ornithine |
| 3438166 | A>G | F230L (TTT>CTT) | ACIAD3510 (lpxC) | UDP-3-O- N-acetylglucosamine déacetylase |

Figure 5.2 – Détail des mutations identifiées chez des clones et des populations dans l'expérience d'évolution sans brassage de génomes (lud). La position des mutations est indiquée par rapport au génome de référence d'*Acinetobacter sp. ADP1*. Lorsque la mutation est intergénique, sa distance par rapport aux 2 gènes est indiquée : un signe + indique que la mutation est en aval du codon stop du gène et un signe moins indique que la mutation est en amont du codon d'initiation du gène.

La mutation 2333167 T>C touche le même locus que la mutation 2332965 G>A. Les fréquences de ces deux mutations varient en sens opposé mais la somme des deux fréquences se maintient autour de 95% au cours de l'évolution. Ceci suggère que ces mutations sont présentes dans deux lignées en compétition et le séquençage de clones tend confirmer cette hypothèse car les clones porteurs de la mutation 2332965 G>A n'ont pas la mutation 2333167 T>C et vice versa. Ces deux mutations sembleraient être un exemple d'interférence clonale où la mutation 2332965 T>C serait apparue dans un clone non porteur de la mutation 2333167 T>C alors que celle-ci était sur le point de se fixer (elle était à 92 % quand la mutation 2332965 G>A était à 2 %) La mutation 2891078 A>G pourrait être apparue dans une sous-population porteuse de la mutation 2333167 T>C car la fréquence de la mutation 2891078 A>G décroît

parallèlement à celle de la mutation 2333167 T>C mais reste inférieure. Les clones séquencés porteurs de la mutation 2891078 possèdent la mutation 2333167 T>C mais pas la mutation 2332965 G>A.

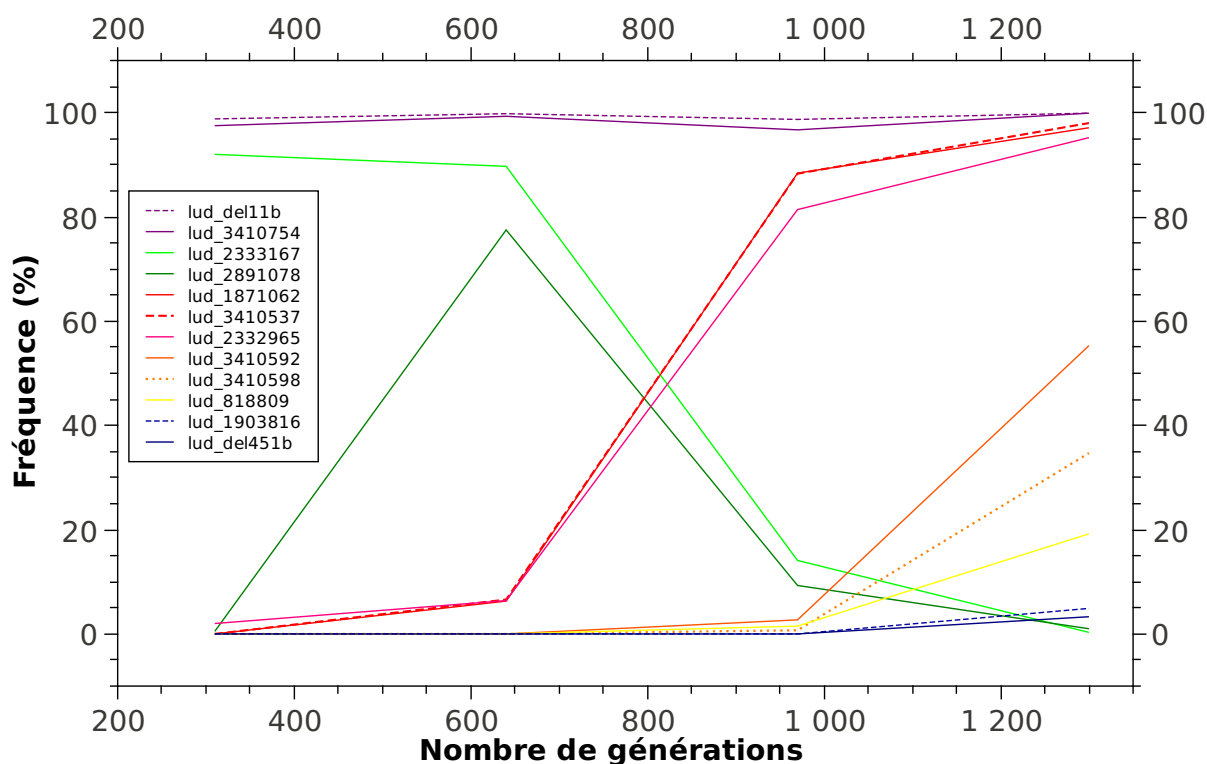


Figure 5.3 – Evolution des fréquences de mutations identifiées dans les populations de l'expérience d'évolution sans brassage de génomes (*lud*)

La divergence de la lignée du clone lud1300t avec celle de la lignée de lud1300a paraît significative car ces deux clones n'ont en commun que les trois mutations qui sont d'ailleurs fixées avant 310 générations. La sous-population dont est issue le clone lud1300t représente environ 1% de la population d'après la fréquence des mutations propres à lud1300t et serait à l'origine de la formation récurrente de biofilm au cours de l'expérience.

5.2.2.2 Résultats des séquençages de génomes dans l'expérience d'évolution avec brassage de génomes (*sem*)

Les résultats des séquençages de génomes des populations et des clones à différents temps de l'expérience avec brassage de génomes (*sem*) sont présentés sous forme de tableau dans la figure 5.4. Le détail sur ces mutations est présenté dans le tableau de

la figure 5.5. Les évolutions des fréquences de certaines mutations sont représentées graphiquement dans la figure 5.6.

| mutations | sem310a | sem310 | sem640a | sem640 | sem970a | sem970 | sem1300a | sem1300 |
|--|---------|----------------|---------|----------------|---------|----------------|----------|----------------|
| Délétion séquence pas présente chez ADP1 | oui | oui | oui | oui | oui | oui | oui | oui |
| IS en 115104 | oui | 100,00% | oui | 100,00% | oui | 100,00% | oui | 100,00% |
| 139151 A>- | oui | 27,40% | non | 10,30% | oui | 81,60% | oui | 99,70% |
| 1719568 A>G | non | 63,90% | non | 49,50% | non | 14,40% | non | non |
| Duplication 18848721-1858316 | oui | oui | non | oui | oui | oui | oui | oui |
| 1920620 A>C | non | 72,50% | non | 51,00% | non | 13,80% | non | non |
| délétion 11b | oui | 25,50% | non | 40,10% | oui | 85,80% | oui | 100,00% |
| 1920624-1920634 | oui | 25,50% | non | 40,10% | oui | 85,80% | oui | 100,00% |
| 2333167 T>C | non | 68,50% | oui | 60,80% | non | 10,00% | non | 0,80% |
| IS en 3240902 (ComP) | oui | 32,00% | non | 11,10% | oui | 85,40% | oui | 100,00% |
| 1708671 T>C | non | 3,10% | oui | 43,60% | non | 10,10% | non | non |
| 2200023 T>C | non | 3,60% | oui | 8,10% | non | 1,00% | non | non |
| 594265 A>G | non | non | non | 36,80% | non | 8,60% | non | non |
| IS en 940787 | non | non | oui | 16,00% | non | 0,90% | non | non |
| Délétion 2188 pb | non | non | non | 6,00% | non | 23,10% | oui | 23,20% |
| IS en 945970 | non | non | non | 9,00% | oui | 62,00% | non | 76,30% |
| 1903776 T>C | non | non | non | 0,10% | non | 9,10% | non | non |
| 1903797 G>A | non | non | non | 2,20% | non | 8,30% | non | 6,30% |
| 1903816 C>T | non | non | non | 1,90% | oui | 31,00% | oui | 71,00% |
| 1903838 G>A | non | non | non | 0,70% | non | 6,50% | non | 2,70% |
| 1903844 G>A | non | non | non | 2,50% | non | 12,90% | non | 6,80% |
| 1903876 A>T | non | non | non | 2,10% | non | 1,00% | non | 2,10% |
| Inversion 2041607-2046081 | non | non | non | 9,30% | oui | 81,20% | oui | 99,50% |
| IS en 2118345 | non | non | non | 8,60% | oui | 78,60% | oui | 99,80% |
| 2350095 A>G | non | non | oui | 7,90% | non | 1,30% | non | non |
| IS en 2660050 | non | non | non | 7,60% | oui | 77,60% | oui* | 100,00% |
| 2886521 A>G | non | non | non | 8,30% | oui | 80,10% | oui | 99,40% |
| 2890094 T>C | non | non | non | 1,80% | oui | 74,50% | oui | 97,00% |
| Délétion 55,984 kb | non | non | non | non | oui | 50,20% | oui | 85,40% |

Figure 5.4 – Mutations identifiées chez des clones et des populations dans l'expérience d'évolution avec brassage de génomes (sem). Les fréquences en gras ont été estimées manuellement d'après le rapport entre le nombre de fragments porteurs d'un motif spécifique de la mutation, et le nombre de fragments porteurs d'un motif non présent chez les porteurs de cette mutation. * l'IS est inséré en 2660050 mais 38 pb sont délétées après l'IS

Trois mutations sont fixées avant 310 générations :

- la délétion d'une séquence de 40 kb non présente dans le génome de référence d'*Acinetobacter baylyi* sp. ADP1. Cette séquence a été identifiée dans notre souche ancestrale non modifiée ABWT (voir section 5.1 et annexes)
- une duplication d'une région de 9,6 kb qui code pour des protéines de phage
- l'insertion d'un IS en 115104 dans un gène codant pour un régulateur transcriptionnel

| position | mutation | type de mutation | gène | fonction |
|--------------------------------|-------------------------------|--|--|---|
| 115104 | IS (-) (+3pb) | Insertion dans gène | ACIAD0114 (lrp) | régulateur transcriptionnel de l'opéron dadAX, lrp (leucine responsive protein.) |
| 139151 | Δ 1pb (A>-) | décalage du cadre de lecture d'un gène | ACIAD0135 | putatif répresseur transcriptionnel |
| 594265 | A>G | K210G (AAA>GAA) | ACIAD0604 | putative glycosyltransferase (biosynthèse LPS) |
| 940787 | IS (-) (+3pb) | Entre 2 gènes (phase +1 / phase+3) (+14 / -124) | ACIAD0949 / ACIAD0951 | protéine hypothétique / protéine hypothétique |
| 942543-943620 et 944858-945967 | Δ 2188 pb | raccordement début IS1236_2 (phase-1) en 945967 (943621-944857 correspond à l'IS_1236_3 absent chez l'ancêtre) | ACIAD0955, ACIAD0956, ACIAD0959, ACIAD0960 (-241) | Putative protéine hypothétique du transposon Tn5613 (ACIAD0955-956) ; putative protéine membranaire (ACIAD0959), putative succinate-semialdéhyde déshydrogénase (ACIAD0960) |
| 945967 | IS(-) (+3pb) | entre 2 gènes (phase -1 / phase +2) (- 348 / - 241) | ACIAD0959 / ACIAD0960 | putative protéine membranaire (ACIAD0959), putative succinate-semialdéhyde déshydrogénase (ACIAD0960) |
| 1708671 | T>C | K92R (AAA>AGA) | ACIAD1702 (pcaU) | Activateur transcriptionnel de l'opéron pca – qui (utilisation du protocatechuate et du quinate) |
| 1719568 | A>G | L479L (TTA>TTG) | ACIAD1714 (quiC) | 3-dehydroshikimate déshydratase |
| 1848721-1858316 | duplication | duplication | ACIAD1846-1861 | putatives protéines de phage |
| 1903776 | T>C | N10D (AAT>GAT) | ACIAD1911 | transporteur transmembranaire du nitrate, premier gène de l'opéron codant aussi pour les nitrate et nitrite réductases |
| 1903797 | G>A | 12 pb avant premier gène d'un opéron | | |
| 1903816 | C>T | 31 pb avant premier gène d'un opéron | | |
| 1903838 | G>A | 53 pb avant premier gène d'un opéron | | |
| 1903844 | G>A | 59 pb avant premier gène d'un opéron | | |
| 1903876 | A>T | 91 pb avant premier gène d'un opéron | | |
| 1920620 | A>C | Entre 2 gènes (phase -2 / phase +1) (- 47 / - 212) | | |
| 1920624-1920634 | Δ 11 pb | Entre 2 gènes (phase -2 / phase +1) (- 51 / - 198) | ACIAD1931 / ACIAD1933 | transporteur secondaire du citrate et du magnésium / protéine hypothétique |
| 2041607 – 2046081 | inversion | inversion dont les bornes coupent dans ACIAD2054 et ACIAD2058 (gltI, premier gène de l'opéron glt) | ACIAD2054 / ACIAD2055 à ACIAD2058 (gltL, gltK, gltJ et gltI) | Protéine hypothétique similaire à DcaP (précurseur de porine dans la catabolisme des acides dicarboxylique / sous unités de l'ABC transporteur glutamate/aspartate (opéron glt) |
| 2118342 | IS (+) (+3b) | Entre 2 gènes (phase +1 / phase -3) (+ 1504 / +464) | ACIAD2126 / ACIAD2129 | protéine hypothétique / protéine hypothétique |
| 2200023 | T>C | N402S (AAC>AGC) | ACIAD2227(dctA) | transporteur des C4-dicarboxylate |
| 2333167 | T>C | une base après le codon stop | ACIAD2374 (miaA) | tRNA delta(2)-isopentenylpyrophosphate transferase |
| 2350095 | A>G | F87S (TTT>TCT) | ACIAD2393 | putatif régulateur transcriptionnel |
| 2660047 | IS (-) (+3pb) | Insertion dans gène | ACIAD2702 | protéine hypothétique |
| 2660050-2660088 | IS (-) + Δ 38 pb après | | | |
| 2670471-2726454 | Δ 55,984 kb | | | |
| 2670471-2726454 | Δ 55,984 kb | délétion | ACIAD2724 à ACIAD2778 | |
| 2886521 | A>G | S262S (TCT>TCC) | ACIAD2956 | protéine membranaire hypothétique |
| 2890094 | T>C | N355S (AAT>AGT) | ACIAD2960 (qseC) | histidine kinase, senseur d'un système de régulation transcriptionnelle à deux composants (qseB et qseC) (quorum sensing) |
| 3240902 | IS | Insertion dans gène | ACIAD3338 (ComP) | gène de compétence, pseudopiline |

Figure 5.5 – Détail des mutations identifiées chez des clones et des populations dans l'expérience d'évolution avec brassage de génomes (sem). La position des mutations est indiquée par rapport au génome de référence d'*Acinetobacter sp. ADP1*. Lorsque la mutation est intergénique, sa distance par rapport aux 2 gènes est indiquée : un signe + indique que la mutation est en aval du codon stop du gène et un signe moins indique que la mutation est en amont du codon d'initiation du gène. Pour les insertions d'IS, le signe + et - indique l'orientation de l'IS et +3pb précise que les 3 bases suivant la position sont dupliquées de parts et d'autres de l'IS.

Un IS s'est inséré dans le gène de compétence comP avant 310 générations. Sa fréquence décroît avant de réaugmenter entre 640 et 970 générations pour finir par se fixer avant 1300 générations.

La fréquence de la délétion d'une base en 139151 suit une évolution similaire à celle de l'IS dans comP. La différence entre les fréquences de ces deux mutations reste inférieure à 5% et à 1300 générations, les deux fréquences sont identiques car les mutations sont fixées.

Cinq autres mutations (inversion, IS en 2118345, IS en 2660050, 2886521 A>G et

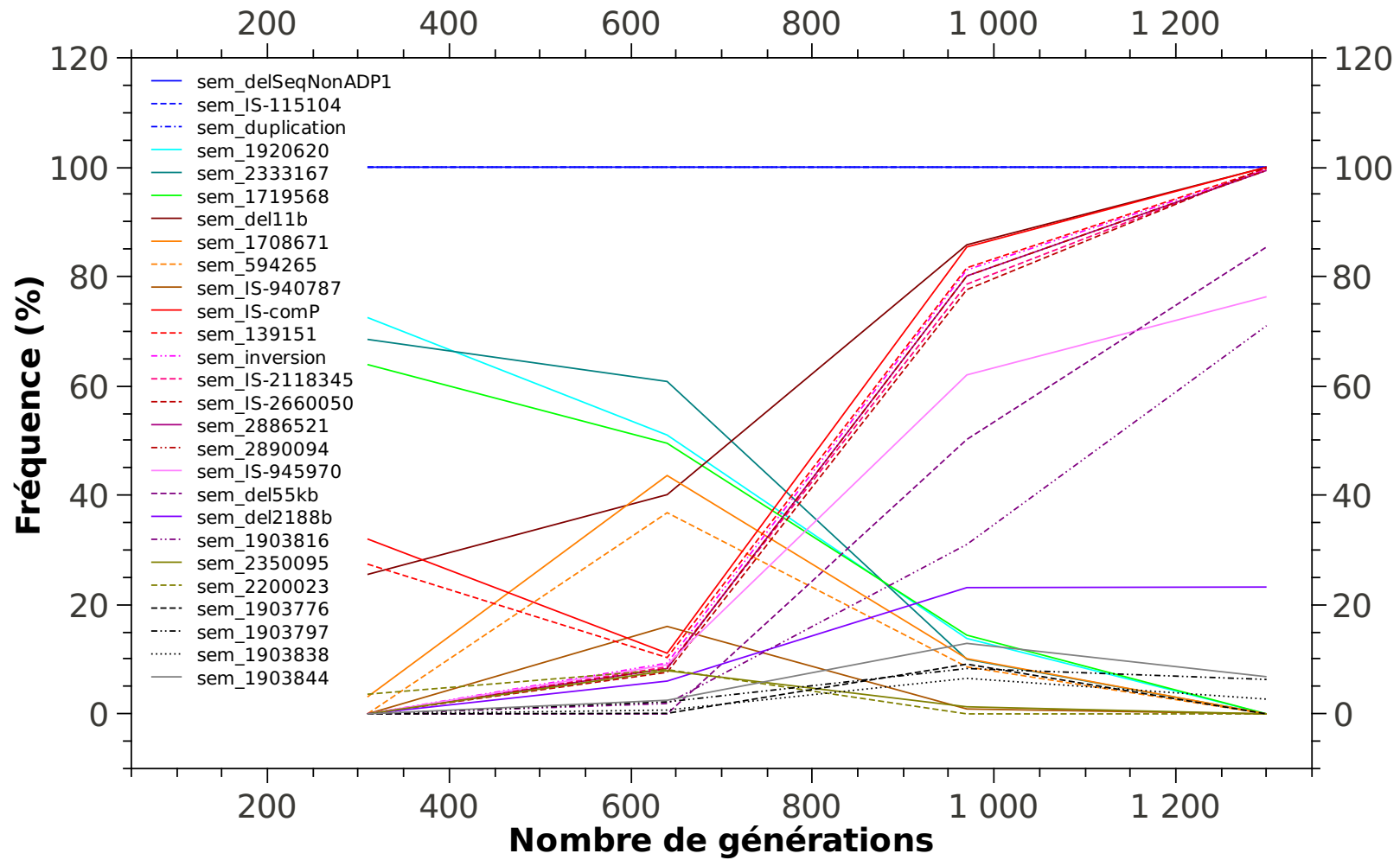


Figure 5.6 – Evolution des fréquences de mutations identifiées dans les populations de l'expérience d'évolution avec brassage de génomes (sem)

2890094 T>C) qui apparaissent entre 310 et 640 générations ont des fréquences proches de celle de l'IS dans le gène *comP*. Les fréquences de ces mutations sont très proches les unes des autres entre 640 et 1300 générations et les différences de fréquence entre ces mutations et l'IS dans *comP* restent inférieures à 11% et deviennent quasiment nulles car ces mutations se fixent.

Comme des transferts horizontaux peuvent avoir lieu, il est difficile d'établir si les changements de fréquence des mutations sont dus uniquement à de l'interférence clonales ou si des transferts horizontaux peuvent alimenter ces changements de fréquence : l'augmentation de la fréquence d'une mutation peut être due à l'envahissement d'une lignée porteuse de cette mutation et/ou à l'acquisition de cette mutation soit par mutation soit par transferts horizontaux chez d'autres lignées. Ainsi la fréquence de l'IS dans *comP* peut augmenter soit parce que les bactéries non compétentes ont acquis des mutations avantageuses qui leur confèrent un avantage par rapports aux lignées compétentes soit parce que des clones compétents ont acquis par transfert horizontal la mutation dans le gène *comP* et deviennent ainsi non compétents. Les mutations qui apparaissent entre 310 et 640 générations et qui ont ensuite des fréquences proches (inversion, IS en 2118345, IS en 2660050, 2886521 A>G ET 2890094 T>C) pourraient être apparues soit chez des lignées non compétentes soit chez des clones compétents qui auraient acquis par la suite l'IS dans le gène *comP*. La présence simultanée de ces mutations chez un clone non compétent (*sem970a* et *sem1300a*) va dans le sens de ces hypothèses mais ne permettent pas d'exclure d'autres possibilités.

La fréquence de l'IS dans *comP* diminue entre 310 et 640 générations. Cette diminution pourrait être due à des mutations avantageuses qui seraient apparues chez des clones compétents soit par mutation soit par transfert horizontal. Les mutations 1708671 T>C, 595265 A>G, 220 ou encore l'IS en 940787 pourraient être ces mutations avantageuses car leurs fréquences augmentent quand la fréquence de l'IS dans *comP* diminue et vice versa. La mutation 1708671 T>C et l'IS en 940787 sont bien apparus chez des clones compétents (présents chez *sem640a*) mais on ne peut pas déterminer avec seulement ces données si ces mutations ont permis à des bactéries compétentes d'avoir transitoirement un avantage.

5.2.3 Comparaison des expériences d'évolution avec ou sans brassage de génomes

5.2.3.1 Loci mutés en commun

Les trois expériences ont des évolutions très différentes. Il faut cependant prendre garde au fait que dans l'expérience pilote, les conditions de cultures n'ont pas été exactement identiques avec celles des expériences *sem* et *lud* (taux de dilution partiellement variable et pas de brassage en dehors du réacteur dans le cas de l'expérience pilote). Dans ce qui suit on ignore ces particularités. Seule la mutation en 1903816 est retrouvée en commun dans les trois expériences. Cette mutation est dans la région promotrice ou régulatrice d'un opéron impliqué dans l'assimilation du nitrate qui est le substrat limitant et la seule source d'azote. Bien qu'on ne sache pas si cette mutation affecte

réellement l'expression de cet opéron, une adaptation au niveau de l'efficacité de l'assimilation du nitrate est attendue même si on a montré dans l'expérience pilote qu'à 85 générations, l'efficacité de capture du nitrate présent dans le milieu était très bonne (plus de 99,8% du nitrate présent dans le milieu est consommé).

La convergence entre les répliques d'expériences d'évolution est souvent plus élevée à l'échelle du locus que de la mutation [28, 130]. Si on compare donc les trois expériences à l'échelle du locus et non de la mutation, trois loci sont mutés dans toutes les expériences. Le premier locus est la région promotrice/régulatrice de l'opéron impliqué dans l'assimilation du nitrate pour lequel on retrouve la mutation commune en 1903816 ainsi que cinq autres mutations dont trois communes à l'expérience pilote (AB) et à l'expérience d'évolution avec brassage (sem).

Le second locus muté dans les trois expériences est l'opéron *qseB/qseC* qui code pour un système à deux composants avec *qseC*, un senseur membranaire de type histidine kinase, et *qseB* un régulateur transcriptionnel activé par cette histidine kinase. Dans l'expérience d'évolution pilote, le gène *qseB* est muté tandis que dans les deux autres expériences c'est le gène *qseC*.

Enfin le troisième locus muté dans les trois expériences est la région promotrice/régulatrice du gène *ACIAD1931* codant pour un transporteur secondaire du citrate (la source unique de carbone). Une délétion de 11b dans ce locus est retrouvée à la fois chez *sem* et *lud* et une seconde mutation en 1920620 est en compétition avec la délétion de 11b (la mutation 1920620 se situe dans les 11b délétés) chez *sem* alors que pour l'expérience pilote, il s'agit d'une autre mutation (insertion d'un A en 1920605).

En conclusion, les mutations de ces trois loci sont des exemples de convergence évolutive et semblent être une adaptation au milieu car deux de ces loci sont liés à l'assimilation du substrat limitant (nitrate) ou au transport de la source de carbone (citrate).

Des réarrangements sont retrouvés à l'identique ou presque dans l'expérience d'évolution pilote (AB) et l'expérience d'évolution avec brassage :

- la duplication 18848721-1858316 est retrouvée à l'identique entre ces deux expériences,
- l'inversion possède quasiment les mêmes bornes (2041592-2046126 dans l'expérience pilote et 2041607-2046081)
- une grande délétion est retrouvée avec une même borne de fin en 2726454 mais celle de l'expérience d'évolution pilote est plus grande (60 kb contre 56kb)

Plus de loci mutés sont retrouvés en commun entre l'expérience d'évolution pilote et l'expérience d'évolution avec brassage (sem). Est-ce que cette différence est due au fait que dans l'expérience d'évolution sans brassage (*lud*), la souche de départ a été modifiée pour la rendre non compétente ? Les interactions épistatiques entre les mutations ont éventuellement pu être modifiées chez la souche non compétente et ont pu ainsi changer la trajectoire adaptative. Pour tester cette hypothèse, il faudra réaliser des répliques des expériences d'évolution avec ou sans brassage afin d'estimer la divergence entre les répliques d'une même expérience. Pour le contrôle sans brassage, une souche non compétente mutée dans un autre locus (par exemple *comP*) et/ou mutée avec une autre

cassette peut aussi être utilisée . Un contrôle avec une souche compétente mais pour lequel on ne met pas d'ADN pendant la transformation pourrait éventuellement être envisagée mais ce contrôle n'est pas parfait non plus. En effet si on reproduit une expérience exactement dans les mêmes conditions que l'expérience d'évolution avec brassage mais sans injection d'ADN, on ne peut pas être assuré qu'il n'y ait pas de brassage car la présence éventuelle d'ADN relargué par les bactéries (mortes ou vivantes) ne peut pas être exclue dans un contexte où la phase de manipulation des bactéries hors réacteur est conçue pour induire la compétence. A titre de comparaison, l'expérience "pilote" décrite précédemment a été menée sans ces inductions de compétence hors réacteur.

5.2.3.2 Comparaison de la diversité

Pour comparer l'évolution de la diversité dans chaque expérience, la diversité a été estimée en utilisant l'index de Shannon $H = -\sum_i p_i \log p_i$ avec p_i la fréquence de la mutation i au lieu de la fréquence de la sous-population i car on ne connaît que les fréquences des mutations (figure 5.7). La diversité est calculée pour l'ensemble des mutations (figure 5.7 en haut) en distinguant aussi la contribution de chaque type de mutation dans la diversité totale (figure 5.8).

La diversité dans l'expérience d'évolution avec brassage (*sem*) est plus élevée que dans l'expérience d'évolution sans brassage (*lud*) : 6 fois plus à 310 générations, 5,4 fois plus à 640 générations et 4 fois plus à 970 générations (figure 5.7 en haut).

Cependant après une importante augmentation de la diversité jusqu'à 640 générations, la diversité dans l'expérience avec brassage augmente peu entre 640 et 970 générations (stagne entre 2,1 et 2,2). Elle finit même par diminuer d'un facteur 3 entre 970 et 1300 générations et devient ainsi similaire à celle de la population *lud*1300 . Cette baisse de la diversité dans l'expérience d'évolution avec brassage est probablement liée à la fixation des mutations ayant des fréquences très proches de la mutation dans *comP* (voir section 5.2.2.2).

Dans l'expérience pilote (*AB*) et l'expérience d'évolution sans brassage (*lud*), la diversité augmente linéairement avec le temps et la diversité à des temps plus tardifs dans l'expérience d'évolution pilote semble être dans le prolongement de l'augmentation de la diversité dans l'expérience d'évolution sans brassage (*lud*). Dans l'expérience d'évolution pilote, on peut supposer qu'il n'y a pas de brassage car lors de test de transformation en réacteur, aucun transformant n'a pu être obtenu (taux de transformation inférieur à 10^{-9}). Bien que la diversité augmente linéairement dans ces deux expériences d'évolution sans brassage (*AB* et *lud*), ce n'est pas toujours le cas dans les expériences d'évolution asexuée comme le montre l'évolution de la diversité dans l'expérience de Lenski où la diversité diminue fortement entre 10 000 et 20 000 générations (figure 5.7 en bas, [8]).

Pour déterminer quelle est la contribution de chaque type de mutations à la diversité, l'index H a été calculé pour chaque type de mutation (figure 5.8). La diversité dans l'expérience d'évolution sans brassage *lud* est essentiellement due à des SNP tandis que dans les deux autres expériences, les réarrangements contribuent aussi à la diversité.

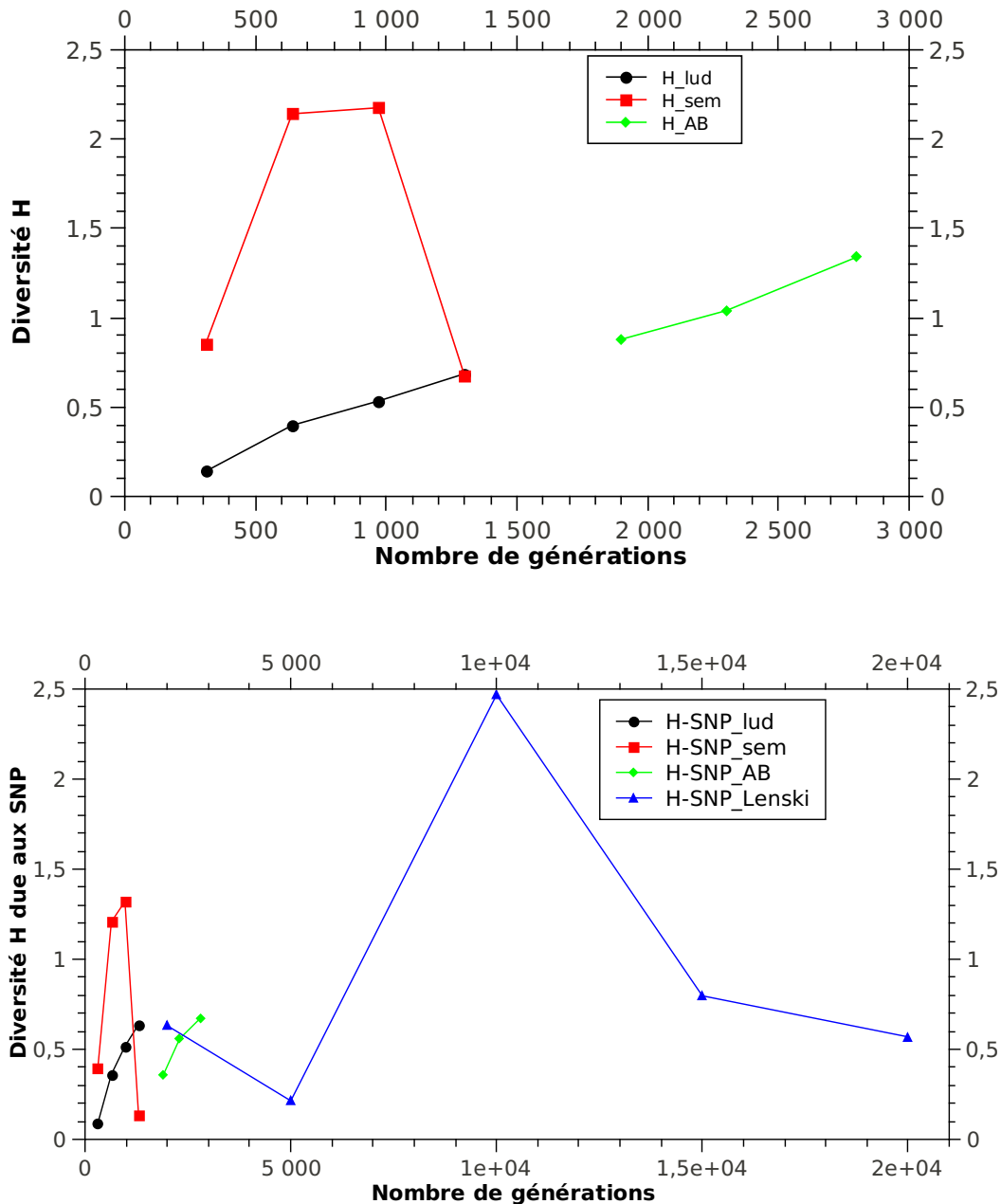


Figure 5.7 – Évolution de la diversité dans l'expérience d'évolution pilote (AB), les expériences d'évolution avec ou sans brassage (sem et lud) et l'expérience d'évolution de Lenski [8]. La diversité H est estimée à l'aide de l'index de Shannon $H = -\sum_i p_i \log p_i$ avec p_i la fréquence de la mutation i (au lieu de la fréquence de la sous-population i car on n'a accès qu'à la fréquence des mutations). Dans la figure du haut, la diversité est calculée en prenant en compte toutes les mutations détectées tandis que dans la figure du bas, la diversité est calculée uniquement pour les SNP afin de pouvoir comparer à l'expérience de Lenski où seule la fréquence des SNP a été recherchée [8].

En effet les SNP constituent en moyenne 82% de la diversité dans les populations lud tandis que les SNP ne contribuent à la diversité pour seulement 67 % en moyenne dans les populations sem et 47% en moyenne dans les populations tardives de l'expérience d'évolution pilote.

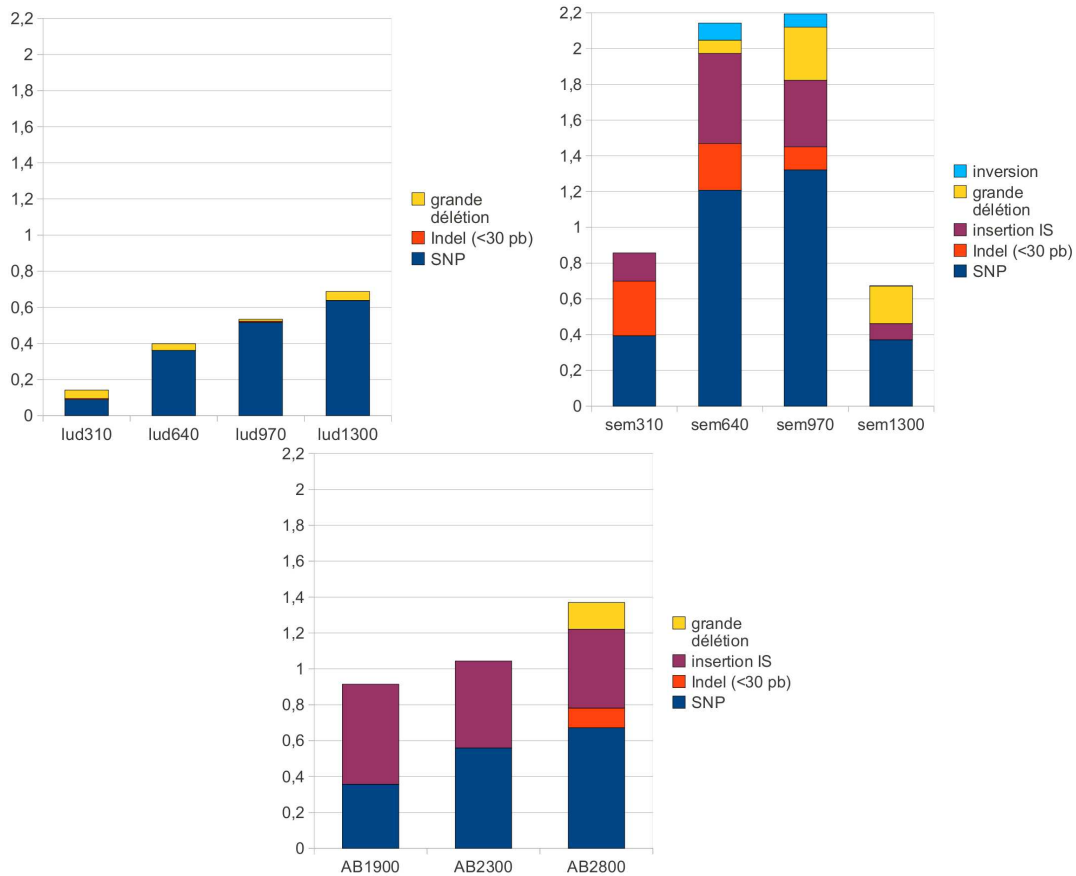


Figure 5.8 – Contribution des différents types de mutations à la diversité H pour les expériences d'évolution avec ou sans brassage (sem et lud) et l'expérience d'évolution pilote (AB). La diversité H est estimée à l'aide de l'index de Shannon $H = -\sum_i p_i \log p_i$ avec p_i la fréquence de la mutation i (au lieu de la fréquence de la sous-population i car on n'a accès qu'à la fréquence des mutations). La diversité est calculée pour chaque catégorie de mutation (SNP, Indel, grande délétion, duplication, insertion d'IS, inversion).

L'estimation de la diversité dans les expériences d'évolutions sont très dépendantes du taux de couverture du génome et de la sensibilité des outils de recherche de SNP [8]. De plus la comparaison de la diversité entre les expériences n'est pas évidente car elle peut dépendre de la phase dynamique de l'évolution où ont été prélevés les échantillons de populations qu'on compare. En effet comme on a vu précédemment, la diversité

n'augmente pas toujours de façon monotone avec le temps car des remplacements de populations peuvent faire diminuer la diversité.

5.2.3.3 Comparaison des types de mutations

Comme vu précédemment, les SNP sont la principales sources de diversité dans les populations lud et sem. Afin de prendre en compte toutes les mutations même celles qui ne contribuent pas à la diversité (mutations fixées), le "nombre effectif" de mutations est calculé correspondant pour chaque type de mutation à la somme des fréquences des mutations (figure 5.9).

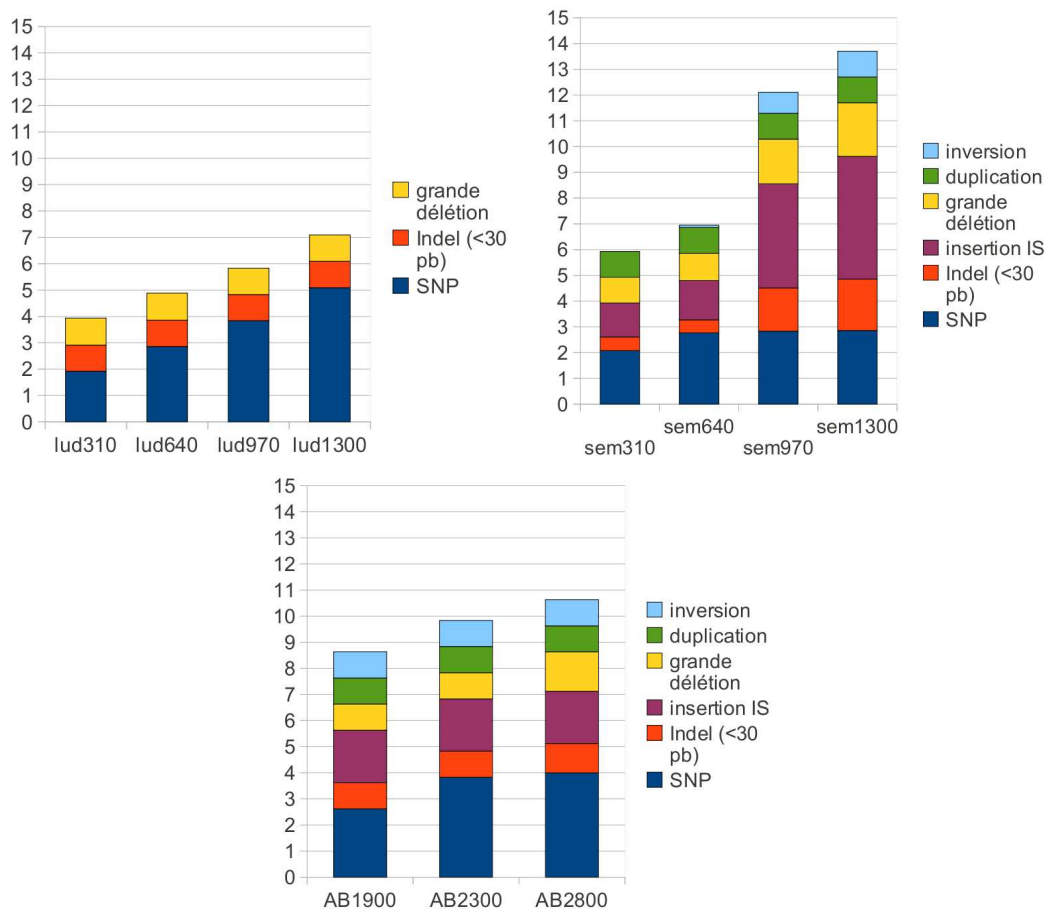


Figure 5.9 – Evolution du nombre de chaque type de mutations dans les expériences d'évolution avec ou sans brassage (sem et lud) et dans l'expérience d'évolution pilote (AB). Le nombre de mutations correspond à la somme des fréquences des mutations.

Dans l'expérience d'évolution sans brassage (lud), la proportion de SNP augmente linéairement au cours de l'évolution passant de 49% à 72% alors que dans l'expérience

d'évolution avec brassage(*sem*), la proportion de SNP diminue après 640 générations et passe d'une moyenne de 37% à 20%. Cette baisse de la proportion de SNP chez *sem* est probablement due à l'envahissement de la lignée non compétente qui cause une diminution de la diversité. Dans l'expérience d'évolution pilote la proportion de SNP est minoritaire et reste stable avec une moyenne de 36%.

Chez *sem*, la majorité des mutations sont dues à des réarrangements et en particulier à des insertions d'IS. Ce n'est pas une spécificité du brassage car les IS et les réarrangements ont aussi activement participé à l'adaptation des bactéries dans l'expérience d'évolution pilote et dans d'autres expériences d'évolution sans brassage [1, 23, 47, 112, 129, 132].

Chez *lud*, la proportion des SNP est majoritaire comme dans la plupart des expériences d'évolution [5, 9, 36, 56, 130, 135, 136].

5.2.3.4 Comparaison des SNP dans les expériences d'évolution avec ou sans brassage de génome (*sem* et *lud*)

Bien qu'il ne soit pas facile de classer les mutations selon leur effet pour de multiples raisons (fonction du gène muté inconnu, effet de la mutation pas facile à déterminer ...), une classification des SNP a été tentée selon leur type (intergénique, non synonyme et synonyme) ou selon les effets de ces mutations sur la régulation (non régulateur, régulateur pour les SNP dans des régions potentiellement régulatrices ou dans des gènes codant pour des régulateurs et enfin inconnu dans le cas où l'effet de la mutation ne peut pas être déterminé) (figure 5.10).

Sous un régime de sélection positive, on s'attend à ce que dans les séquences codantes, les mutations non-synonymes (changeant un acide aminé par un autre) soient plus nombreuses que les mutations synonymes (mutations géniques ne changeant pas d'acide aminé par un autre) car ces dernières sont souvent neutres. Dans l'expérience d'évolution sans brassage (*lud*), 100% des SNP géniques sont non-synonymes alors que dans l'expérience d'évolution avec brassage (*sem*), une mutation synonyme (2886521 A>G) apparaît à 640 génération et se propage pour se fixer (la proportion de SNP géniques qui sont non-synonymes passe de 100 % à 49 % entre 310 générations et 1300 générations). D'après la revue [36], dans la grande majorité des expériences d'évolution avec séquençage de génome, les SNP non synonymes représentent presque 100% des SNP géniques (avec une moyenne de 92%).

Les SNP intergéniques et les SNP synonymes ont plus de chance d'être neutres que les mutations non synonymes et un ratio élevé de mutations non synonymes peut donc être un signe de sélection positive. Chez *lud*, la proportion de SNP qui sont non-synonymes augmente de 1% à 79% alors qu'elle est en moyenne de 40 % chez *sem* (avec un minimum de 33 % et un maximum de 53%).

Pour tenter de savoir si les adaptations observées sont principalement guidées par des mutations régulatrices, les SNP ont aussi été classifiés selon leur effet sur la régulation. On a supposé ici que les mutations sont régulatrices si elles touchent un gène codant pour une protéine régulatrice ou si elles sont situées dans des régions

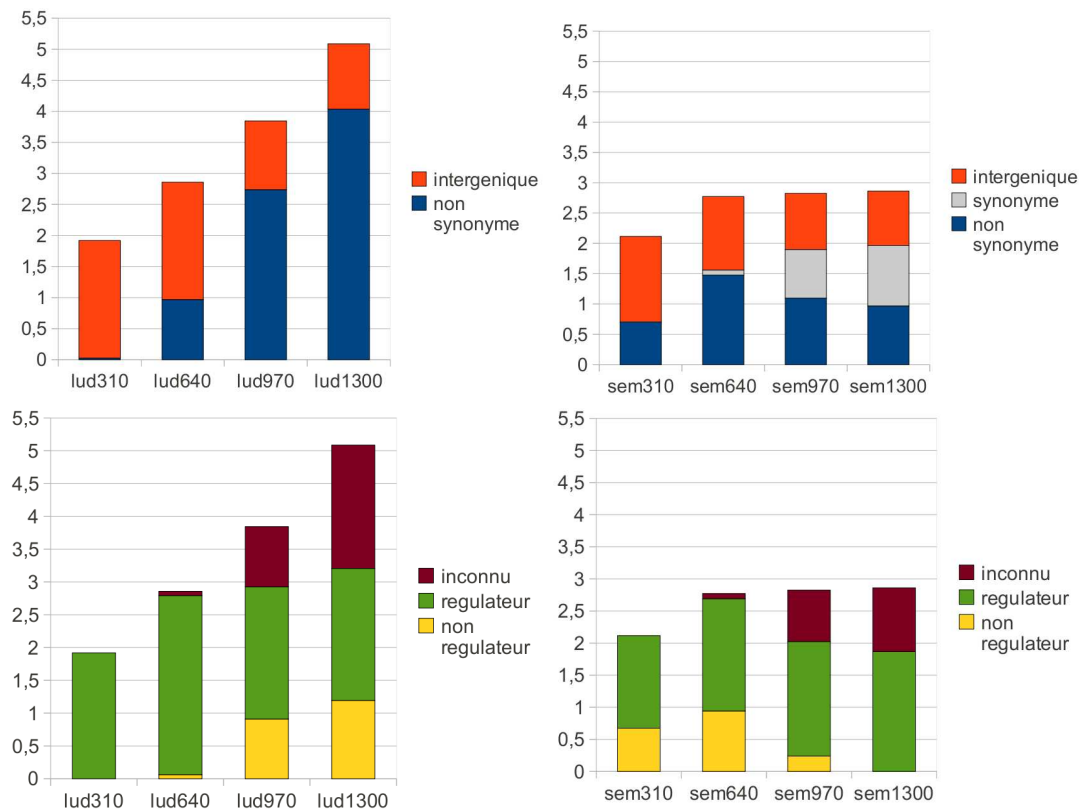


Figure 5.10 – Evolution du nombre de SNP en fonction de leur type ou de leur effet dans les expériences d'évolution avec ou sans brassage (sem et lud). Le nombre de mutations correspond à la somme des fréquences des mutations.

potentiellement régulatrices (intergéniques). Chez lud, la proportion de SNP qui sont régulateurs passe de 100 % à 40 % entre 310 et 1300 générations alors que chez sem, elle se maintient avec une moyenne de 65%. Pour lud les mutations régulatrices semblent plus importantes au début de l'évolution alors que chez sem, l'adaptation semble principalement guidée par des mutations régulatrices. Cependant on surestime peut être le nombre de mutations régulatrices car toutes les mutations intergéniques ne sont pas forcément régulatrices et certaines peuvent être neutres. De plus pour sem, les SNP ne constituent qu'une minorité des mutations (en moyenne 36%).

5.2.3.5 Résumé

Pour les trois expériences menées (i) expérience pilote sans réinjection d'ADN (brassage faible), (ii) expérience sem avec réinjection d'ADN et souche compétente (avec brassage), et (iii) expérience lud avec réinjection d'ADN et souche non-compétente (sans brassage), le nombre de loci mutés commun (voir section 1.2.3.1) se limite à 3

, (a) région promotrice/régulatrice de l'opéron nitrate, (b) opéron *qseB/qseC*, et (c) région promotrice/régulatrice de transporteur citrate.

Par ailleurs en plus de ces 3 loci, des réarrangements similaires (3 en tout) sont retrouvés entre expérience pilote et expérience avec brassage *sem* (voir 1.2.3.1).

On observe par ailleurs (voir 1.2.3.2) une grande différence entre l'évolution *lud* et l'évolution *sem* caractérisée d'abord (A) par une diversité (voir figure 1.8) qui est dans une phase initiale (de 310 à 970 générations) nettement plus élevée chez *sem* et fait intervenir délétions, insertions d'IS et SNP. La diversité chez *sem* décroît ensuite (entre 970 et 1300 générations) pour rejoindre vers 1300 générations une diversité du même ordre de grandeur que chez *lud*1300.

Une comparaison de diversité sur une autre expérience d'évolution a pu être faite avec une expérience de Barrick-Lenski [8], mais en se limitant au cas documenté de la diversité liée uniquement aux SNP. La diversité en ce cas semble être globalement dans un même ordre de grandeur entre *sem*, *lud* et l'expérience barrick-Lenski (voir section 1.2.3.2).

L'autre différence nette entre *sem* et *lud* est (B) le nombre effectif de mutations (figure 1.9 et 1.10). Ce nombre à 1300 générations est plus grand de près d'un facteur 2 (13.5 contre 7) chez *sem* par rapport à *lud*. La répartition en type de mutations est également différente avec pour *lud*1300, 5 SNP (4 non-synonymes, 1 intergénique), 1 indel, 1 grande délétion, et pour *sem*1300, 3 SNP (1 non-synonyme, 1 synonyme, 1 intergénique), 2 indel, 5 insertions d'IS, 2 délétions, 1 duplication, 1 inversion.

On a donc des différences très marquées entre les expériences *sem* et *lud*, caractérisées tant du point de vue de la diversité que de la nature des mutations. Une caractéristique étonnante est par ailleurs l'apparition rapide de mutants *sem* non-compétents. Cette sous-population finit par envahir finalement la population *sem* à 1300 générations, et cet envahissement semble concomitant avec la baisse tardive de la diversité chez *sem*.

Au final, ces expériences semblent très encourageantes en ce sens que des différences très notables d'évolution sont obtenues en condition avec transfert horizontal d'ADN par rapport à une situation sans transfert. Ces premiers résultats nécessitent des répliques d'expérience afin de déterminer si ces différences sont robustes, et si la disparition progressive de compétence est également un effet robuste. On a par ailleurs identifié que l'expérience de contrôle avait potentiellement un point faible, le mutant non-compétent utilisé montrant un phénotype de formation facilitée de biofilm, soulevant la question d'une évolution de *lud* éventuellement partiellement dépendante de phénotypes autres que la non-compétence. Il sera donc nécessaire dans la suite de ce travail de procéder à différents types d'expériences de contrôle.

Conclusion

Cette thèse porte sur des expériences d'évolution de bactéries, avec le but d'explorer expérimentalement comment des transferts horizontaux d'ADN au sein d'une population de bactéries issues d'un ancêtre commun, peuvent ou non modifier l'évolution de cette population. La bactérie modèle choisie est *Acinetobacter baylyi* ADP1, et la méthode utilisée pour permettre des échanges faciles d'ADN à l'intérieur de la population repose sur la propriété de "transformation naturelle" de cette bactérie, c'est à dire sa capacité à capturer puis intégrer dans son génome de l'ADN qui serait présent dans le milieu de culture. Un réacteur en flux continu est utilisé, permettant le maintien en conditions de croissance constante d'une très grande population de bactéries (10^{10}) dans un milieu où la seule source d'azote est du nitrate en quantité limitante.

Nous avons dans une première étape mené une expérience pilote, avec pour but de mettre en place le système expérimental. Cette expérience a été initiée avec une souche compétente d'*Acinetobacter baylyi* et avec les conditions expérimentales de croissance en réacteur qui conduisent à un taux de transformation très faible ($< 10^{-9}$). Cette expérience d'évolution pilote, réalisée sur 3000 générations environ, a permis d'obtenir des informations sur la nature des mutations et sur la diversité générée lors de l'évolution d'*Acinetobacter baylyi* dans des conditions sélectives choisies et avec des échanges horizontaux très faibles. Une structure en multiples sous populations en compétition active a été observée et semble compatible avec un régime d'interférence clonale. Une telle structuration est donc susceptible de fournir une source d'ADN avec variations en cas d'échange horizontal.

Pour l'expérience avec brassage de génomes, on réalise une "boucle de rétroaction". On commence par récupérer/purifier périodiquement l'ADN des bactéries du trop-plein (sortie de réacteur), pour préparer de l'ADN génomique de population (ADNg). Les bactéries du réacteurs sont ensuite prélevées et on manipule hors réacteur les conditions de croissance de façon à favoriser la compétence. L'ADNg purifié est enfin introduit avec les bactéries rendues plus compétentes et après une phase d'incubation la population de bactéries est ré-introduite dans le réacteur. Ce procédé a permis d'introduire des étapes d'échanges horizontaux d'ADN. Avec ce protocole d'échanges intervenant tous les 165 générations environ, on a mené une expérience d'évolution avec *Acinetobacter baylyi* sur 1300 générations. En tant que contrôle, on a réalisé dans les mêmes condi-

tions une expérience avec un mutant non-compétent d'*Acinetobacter baylyi*. Pour ces deux expériences (brassage et expérience contrôle), on a procédé à des reséquençages tous les 330 générations, à la fois sur la population et sur un clone isolé choisi aléatoirement. L'exploitation de ces données de séquençage a permis de mettre en évidence les mutations apparues au cours de l'évolution expérimentale ainsi que leurs fréquences relatives dans la population.

Les résultats obtenus dans ces deux dernières expériences sont très différents. Dans l'expérience contrôle (c'est à dire sans transfert horizontal), la diversité est essentiellement due à des SNP et augmente de façon monotone avec le nombre de générations. En contraste, dans l'expérience d'évolution avec brassage, la diversité augmente dans une première phase (de 310 à 970 générations) pour ensuite diminuer et rejoindre (vers 1300 générations) une valeur similaire à la diversité de l'expérience de contrôle. Nous avons pu relier cette diminution de la diversité est à l'envahissement de la population par une lignée devenue non compétente à cause d'une mutation dans le gène *comP* essentiel pour la transformation naturelle. Ces résultats préliminaires sont encourageants car tendent à montrer que l'échange horizontal d'ADN semble avoir modifié transitoirement l'évolution, avec une diversité initialement plus importante dans le cas de transferts horizontaux. De plus la disparition progressive de la compétence, tendrait à indiquer que la compétence présente un coût qui devient trop important à long terme dans notre contexte expérimental.

Des répliques de ces expériences sont cependant nécessaires pour confirmer la répétabilité de ces effets. Des expériences de compétition devront aussi être réalisées afin de comparer les deux expériences. D'autres expériences contrôles devront aussi être envisagées car la délétion de l'opéron *comFECB* pour rendre *Acinetobacter baylyi* non-compétente semble avoir des effets secondaires sur le phénotype de la souche et a pu modifier la trajectoire adaptative. Une possibilité pourrait être de déléter le gène *comP* au lieu de *comFECB*.

Enfin il sera important d'améliorer notre méthode de recherche des mutations afin d'avoir une meilleure détection des mutations à faibles fréquences. Une détection des mutations à un seuil plus bas et avec une méthode non biaisée permettrait une meilleure estimation de la diversité.

Annexes

Protocole d'extraction d'ADN génomique

Pour extraire de grande quantité d'ADN génomique (1 à 2 mg), on part soit d'une culture sur la nuit dans 50 mL de LB ou soit pour les extractions d'ADN de population de 100 mL de culture du réacteur avec une densité optique d'environ 1,2. Les cultures sont centrifugées dans des tubes de 50 mL afin de récupérer un ou deux culots de bactéries. Les différentes étapes de l'extraction d'ADN génomiques sont les suivantes :

1) Lyse des cellules et dégradation de l'ARN

Après resuspension des culots dans 4mL de TE (10 mM Tris HCl pH 8, 1 mM EDTA) (quand deux culots, réunion des deux culots dans un volume de 4 mL de TE), on ajoute 8mL de tampon ES (EDTA 0,5M , Sarkosyl 10%), 25 μ L de protéinase K (225 U) et 30 μ L de RNase A (300 μ g) pour dégrader l'ARN. Une incubation de 5 min au bain marie à 65°C est réalisée pour que les enzymes agissent.

2) Séparation au chloroforme

Après la lyse des cellules, on ajoute de 12 mL de chloroforme. Après agitation pour bien mélanger, une centrifugation à 10 000 rpm est réalisée à 4°C pendant 60 min (temps à optimiser selon la qualité de séparation des phases). On récupère ensuite la phase supérieure contenant l'ADN dans un tube de 50 mL propre

3) Précipitation à l'acétate de sodium (NaAc)

Pour un volume de phase supérieure récupérée, on ajoute un volume de NaAc 3M. Après avoir bien mélanger, on centrifuge à 10 000 rpm et à 4°C pendant 90 min (temps à optimiser selon stabilité du culot). On récupère ensuite le surnageant contenant l'ADN (cette étape permet d'éliminer une grande partie des protéines restantes après le chloroforme).

4) Précipitation Ethanol de l'ADN

Pour un volume de surnageant récupéré, on ajoute un volume d'éthanol 100% -20°C. On mélange délicatement par inversion jusqu'à voir la "pelotte d'ADN" puis on

incube à -20°C pendant au moins 3 heures. Après cette incubation, on centrifuge à 10 000 rpm et à 4°C pendant 20 min. On élimine le surnageant et on rince le culot d'ADN avec 7,5 mL éthanol 70% -20°C (Volume à ajuster pour que culot soit immergé). Après une centrifugation à 10 000 rpm pendant 10 min à 4°C , on sèche le culot d'ADN sous une hotte avant de le resuspendre dans 600 μL TE. Cette resuspension est facilitée par une incubation à 30°C 300 rpm (tube incliné coté culot).

5) Dosage de l'ADN au spectrophotomètre

Récupération en général de 1 à 2 mg d'ADN avec un rapport absorbance à 260 nm / absorbance à 280 nm égal à 2.

Amplification d'ADN génomique par l'ADN polymérase du bactériophage phi29

L'ADN polymérase du bactériophage phi29 (ADN polymérase phi29) est une enzyme très processive pouvant synthétiser des fragments de plus de 10 kb de long avec un taux d'erreur 100 à 1000 fois plus faible que l'ADN polymérase Taq couramment utilisée en PCR. L'ADN polymérase phi29 est souvent utilisée pour amplifier des génomes à partir de très peu de matériel et est une alternative à l'extraction d'ADN génomique qui nécessite plus de matériel de départ [34, 45, 61, 63].

Après une étape de dénaturation de l'ADN à 95°C , l'ADN polymérase phi29 polymérise à 30°C à partir d'hexamères aléatoires modifiés en 3' afin de ne pas être dégradé par l'activité exonucléase 3'-5' de l'enzyme. Quand l'ADN polymérase phi29 rencontre un autre site où s'est initié une polymérisation, elle est capable de déplacer le brin néosynthétisé devant elle afin de continuer sa polymérisation. Ces déplacements multiples de brin libèrent ainsi de nouveau site pour l'hybridation des hexamères qui vont initier d'autre polymérisation.

Le protocole utilisé pour faire de l'amplification phi29 est le suivant : 4, 3.10^7 bactéries sont centrifugées et resuspendues dans un volume final de 42 μL contenant 50 mM Tris-HCl pH 7,5, 50 mM KCl, 10 mM MgCl_2 , 10 mM $(\text{NH}_4)_2\text{SO}_4$, 50 μL d'hexamères avec les 3 dernières bases en 3' avec des groupements phosphothiorate. Le tout est incubé 3 min à 95°C pour lyser les bactéries et dénaturer l'ADN. Ensuite sont ajoutés pour avoir un volume final de 50 μL , de l'ADN polymérase phi 29 (25 U), de la pyrophosphatase et 10 mM de chaque dNTP. Le tout est incubé plus de 15 heures à 30°C pour amplification de l'ADN génomique des bactéries. Les enzymes sont inactivés à 65°C pendant 10 min. Une précipitation NaAc/éthanol est réalisée pour récupérer et concentrer l'ADN amplifié.

Données supplémentaires de l'article sur l'expérience d'évolution pilote

Supplementary Data

1. Genome resequencing of ancestral strain ABWT and comparison with ADP1 strain

The genome of ABWT was resequenced and showed significant differences with *Acinetobacter baylyi* ADP1. These differences are summarized in table S1. A sequence of about 40 kb was present in ABWT but absent in ADP1. ORFs were searched using ORF Finder on line tool and ORFs were blasted. The results are summarized in the table S2.

Table S1. Mutational event in ABWT when compared to *A. baylyi* ADP1

| Index ADP1 | Event | Comment |
|-----------------|-------------|---|
| 178468 | G → – | 74 bp upstream start codon of ACIAD0178 atpI ATP synthase protein I |
| 289081 | insertion | Insertion of IS1236 between ACIAD 0290 and ACIAD 0291 |
| 321573-322809 | deletion | Deletion of IS1236_1 |
| 364108 | G → A | L → F; ACIAD0371 pH adaptation potassium efflux system protein G |
| 943621-944857 | deletion | Deletion of IS1236_3 |
| 968073 | T → C | 500 bp upstream start codon of ACIAD0980 vanA (vanillate O-demethylase oxygenase subunit) |
| 1139725 | – → A | 75 bp upstream start codon of ACIAD1147 putative APC family, S-methylmethionine transporter |
| 1179108 | C → T | D → N; ACIAD1182 phrB deoxyribodipyrimidine photolyase (photoreactivation), FAD-binding |
| 1277087 | T → – | ACIAD1276 putative amino acid transporter; the frameshift results in a stop codon about 100bp downstream the deletion |
| 1278182 | T → – | ACIAD1277; the frameshift results in a stop codon about 100bp downstream the deletion |
| 1444653 | T → G | 62bp upstream start codon of ACIAD1446 catB (muconate cycloisomerase I) |
| 1468567 | C → T | S → L; ACIAD1471 putative transmembrane sensor protein (iron transport) |
| 1648116 | C → T | 157 bp upstream codon start ACIAD1649 putP (SSS family, major sodium/proline symporter) |
| 1794961-1795034 | | Putative duplication and putative divergence of the copy |
| 1803858-1803939 | Duplication | ACIAD1795 |
| 2045478 | A → C | 69 bp upstream start codon of ACIAD2059 gltJ glutamate/aspartate transport protein (membrane) |
| 2420279 | – → C | 56 bp upstream start codon of ACIAD2457 putative membrane protein |
| 2542652 | G → A | Silent mutation in ACIAD2584 lepA GTP binding protein |
| 2737671-2737674 | | Putative duplication and putative divergence of the copy |
| 2745958-2749970 | | Putative duplication and putative divergence of the copy |
| 2765010 | A → C | S → R; ACIAD2827 putative periplasmic binding protein of transport/transglycosylase |
| 2803531-2805196 | | Putative duplication and putative divergence of the copy |
| 2876467 | C → T | S → L; ACIAD2944 putative resistance-nodulation-cell division (RND) family (AdeB-like) |
| 3357842 | G → A | 237 bp upstream start codon ACIAD3437 |
| 3392500 | C → T | A → V; ACIAD3465 putative 2-component sensor |

Table S2. Annotation of a sequence of ABWT absent in ADP1.

| index | frame | size | product | organism | Evalue | % Identity |
|-------------|-------|------|---|--|--------|------------|
| 222-1523 | +3 | 1301 | Phage integrase | <i>Acinetobacter radioresistens SK82</i> | | 90% on |
| 315-1544 | -3 | 1229 | Mutator family transposase | <i>Acinetobacter johnsonii</i> | | 80% on |
| 3127-3933 | -1 | 806 | AraC family transcriptional regulator | <i>Pantoea sp. SLI-M5</i> | | 63% |
| | | | | <i>Klebsiella pneumoniae 342</i> | | 62% |
| 6405-7790 | +3 | 1385 | Citrate transporter | <i>Acinetobacter junii SH205</i> | | 70% |
| | | | | <i>Acinetobacter baylyi ADP1</i> | | 68% |
| 6612-7598 | -3 | 986 | NmrA family protein | <i>Klebsiella oxytoca</i> | | 80% |
| 7854-8795 | +3 | 941 | GntR family transcriptional regulator | <i>Pseudomonas syringae Cit7</i> | | 56% |
| 9022-10089 | -1 | 1067 | Luciferase family protein | <i>Sphingobacterium spiritivorum</i> | | 79% |
| 9120-10520 | +3 | 1400 | amidohydrolase | <i>Pantoea stewartii</i> | | 90% |
| | | | | <i>Klebsiella pneumoniae</i> | | 89% |
| 10123-10974 | -1 | 851 | Hypothetical protein | <i>Leeuwenhoekiiella blandensis</i> | | 58% |
| 11025-11975 | +3 | 950 | SMP30/ gluconolactonase/LRE family | <i>Pantoea sp.</i> | | 79% |
| | | | | <i>Klebsiella pneumoniae</i> | | 80% |
| 13381-14685 | -1 | 304 | CaiB/BaiF family class III CoA transferase | <i>Alcaligenes faecalis</i> | | 85% |
| 14686-15594 | -1 | 908 | Enoyl-CoA hydratase family | <i>Alcaligenes faecalis</i> | | 92% |
| 16004-17734 | -2 | 1730 | SSS sodium solute transporter superfamily | <i>Acidovorax delafieldii 2AN</i> | | 90% |
| 17532-18314 | +3 | 782 | IclR family transcriptional regulator | <i>Acinetobacter sp. P8-3-8</i> | | 75% |
| 18410-19591 | -2 | 1181 | Iron-containing alcohol dehydrogenase | <i>Acinetobacter johnsonii</i> | | 81% |
| 19663-21198 | -1 | 1535 | Methylmalonate-semialdehyde dehydrogenase | <i>Acinetobacter johnsonii SH046</i> | | 83% |
| 22280-22942 | -2 | 662 | Regulatory protein, TetR | <i>Acinetobacter radioresistens SK82</i> | | 81% |
| 23012-24001 | -2 | 989 | Putative dehydrogenase | <i>Acinetobacter baumannii AYE</i> | | 94% |
| 25238-26212 | -2 | 974 | Oxidoreductase | <i>Klebsiella pneumoniae 342</i> | | 88% |
| 27212-28432 | +2 | 1220 | Dimethyl sulfoniopropionate lyase | <i>Alcaligenes faecalis</i> | | 82% |
| 30925-31845 | +1 | 920 | Transcriptional regulator | <i>Acinetobacter calcoaceticus</i> | | 91% |
| 33303-35948 | -3 | 2645 | Aconitate hydratase | <i>Pseudomonas syringae Cit7</i> | | 74% |
| 34143-34754 | +3 | 611 | Modulator of drug activity B | <i>Shewanella sp. HN-41</i> | | 74% |
| 34980-35621 | +3 | 641 | TetR family transcriptional regulator | <i>Serratia odorifera</i> | | 62% |
| | | | | <i>Klebsiella pneumoniae</i> | | 60% |
| 35942-37114 | -2 | 1172 | AcnD accessory protein PrpF | <i>Pseudomonas syringae</i> | | 74% |
| 36415-37725 | +1 | 1310 | Transporter (MFS family) | <i>Klebsiella pneumoniae 342</i> | | 65% |
| 37115-37870 | -2 | 755 | Extracellular solute-binding protein family 3 | <i>Pseudomonas syringae Cit7</i> | | 55% |

2. Identification of mutational events and building of the tree

All mutations were identified by analysis of the genome resequencings with standard tools (cited in text) except the duplication (ACIAD1848 – ACIAD1861), and the insertion loci for branches B3, B4 and B5.

Duplication was identified by the difference in coverage for this region; as a result the duplication is present in all genome resequencing (endpoint clones, and population sampled at 1900, 2300 and 2800th generation) but absent in ABwt.

The insertion loci for branches B3, B4 and B5 were identified as now described: in the fasta files provided by resequencing of sampled populations, we looked for fragments carrying the beginning sequence and the end sequence of IS1236. We identified all joint sequences, allowing to discover new ones. We blasted them on *Acinetobacter baylyi* ADP1 to find out the insertion loci corresponding to B3, B4 and B5. The insertion of another IS was located in a region different from ADP1 strain (see 1). To estimate the ratio of sub populations, we compared for each locus, the

number of fragments founded in fasta files matching with IS insertion with the number of fragments matching with no insertion (wild type profile). For the five branches (B1, B2, B3, B4 and B5), the sum of the ratios is about 100% for the three sequencing of population, justifying that there are no other relevant insertion loci, and that each bacteria underwent only one of such five insertion events.

How we obtained the data in table S3:

From frozen samples of population, some bacteria were plated and allowed to grow overnight; single colonies were piked and allowed to grow overnight in LB, snap-frozen and stored at -70°C. To extract genomic DNA, bacteria issued from a single colony were grown overnight in LB, centrifuged and resuspended in small volume of Tris EDTA. The samples were then heated at 100°C during 10 minutes; lysed cells were centrifuged, the supernatant recovered, aliquoted and stored at -20°C.

To check the presence of point mutations and small indels (glnK, Ins 9bp, ACIAD1029, Δ 8bp, C1903816T, ACIAD1931, ACIAD2054, and qseB in table S3), primers were designed to amplify fragments of about 800 bp (point mutation or indels in the middle of the fragment). The PCR products were purified on electrophoresis gel and sequenced by Eurofin-MWG.

The presence of IS1236 (1236 bp) at two loci (branches B1 and B2) were checked by amplifying each region and PCR product sizes were checked on electrophoresis gel.

To check the large deletion Δ 60kb two different PCR were performed per genome: one PCR with sense primer upstream the deleted sequence and antisense primer within the deleted sequence, and one PCR with antisense primer downstream the deleted region. Only one PCR gives a product which could be checked on electrophoresis gel.

To check the large deletion Δ gltI, one PCR were performed per genome with sense primer upstream the deleted sequence and antisense primer downstream the deleted region. PCR product sizes were checked on electrophoresis gel.

To check the inversion, two different PCRs were performed per genome: the same sense primer were used for both PCRs, and two different antisense primers hybridizing in the inverted sequence were used, one giving product only in the case of an inverted sequence and the other giving product only in the case of a no-inverted sequence. The products were checked on electrophoresis gel.

Table S3. Detection of mutations in several isolated clones¹.

¹ In ABWT, an IS1236 copy is present in ACIAD0290. The transposition of this IS (event “ACIAD0290 Ins 9bp”) results in an insertion of 9 bp in ACIAD0290 when compared to ADP1.

| | glnK | ACIAD0290 | citA | ACIAD1029 | IS-B1 | IS-B2 | Δbbp | C1903816T | ACIAD1931 | Inversion | ACIAD2054 | ΔgltI | Δ60kb | qseB |
|---------|--------|-----------|--------|-----------|---------|---------|---------|-----------|-----------|-----------|-----------|---------|---------|---------|
| | 244223 | 289081 | 389458 | 1020409 | 1230681 | 1231207 | 1249430 | 1903816 | 1920605 | 2041592 | 2042174 | 2046132 | 2664637 | 2891803 |
| | C → T | Ins 9bp | T → C | T → C | | | | | Ins A | | T → C | | | G → A |
| AB85a | no | | no | | | | yes | no | | no | | | yes | no |
| AB85b | | | | | | | yes | | | | | | | |
| AB85c | | | | | | | yes | | | | | | | |
| AB85d | | | | | | | yes | | | | | | | |
| AB85e | | | | | | | | | | | | | yes | |
| AB85f | | | | | | | | | | | | | yes | |
| AB215a | no | | no | | no | | yes | no | | yes | | | yes | yes |
| AB215b | | | | | no | | | | | yes | | | | |
| AB215c | | | | | no | | | | | yes | | | | |
| AB215d | | | | | no | | | | | yes | | | | |
| AB360d | no | | no | | | | | no | | yes | | | yes | yes |
| AB360f | no | | no | | | | | no | | yes | | | yes | yes |
| AB590a | | no | | | no | | | no | | | no | | | |
| AB590b | | | | | no | | | no | | | | | | |
| AB590c | | no | | | no | | yes | yes | | | no | no | | |
| AB590d | | | | | no | | | no | | | | | | |
| AB590e | | | | | no | | | no | | | | | | |
| AB590f | | | | | no | | | no | | | | | | |
| AB590g | | | | | no | | | no | | | | | | |
| AB590h | | | | | no | | | no | | | | | | |
| AB590i | | | | | no | | | no | | | | | | |
| AB590k | | | | | no | | | no | | | | | | |
| AB845a | no | | no | | no | | yes | G1903844A | | yes | | | yes | yes |
| AB845b | no | | no | | no | | | G1903797A | | | | | | |
| AB845c | | | | | no | | | G1903797A | | | | | | |
| AB845d | | | | | no | | | G1903844A | | | | | | |
| AB845e | | | | | no | | | G1903844A | | | | | | |
| AB845f | | | | | no | | | G1903797A | | | | | | |
| AB845g | | | | | no | | | G1903844A | | | | | | |
| AB845h | | | | | no | | | G1903844A | | | | | | |
| AB845i | | | | | no | | | G1903844A | | | | | | |
| AB845j | | | | | no | | | G1903844A | | | | | | |
| AB845k | | | | | no | | | A1903876T | | | | | | |
| AB845l | | | | | no | | | A1903876T | | | | | | |
| AB845m | | | | | no | | | A1903876T | | | | | | |
| AB845n | | | | | no | | | A1903876T | | | | | | |
| AB845p | | | | | no | | | A1903876T | | | | | | |
| AB845q | | | | | no | | | A1903876T | | | | | | |
| AB845r | | | | | no | | | no | | | | | | |
| AB845s | | | | | no | | | A1903876T | | | | | | |
| AB980a | | no | | | no | | | G1903844A | | | no | | | |
| AB980b | | yes | | | yes | | | yes | | | no | | | |
| AB980c | | no | | | no | | | G1903844A | | | | | | |
| AB980d | | yes | | | yes | | | yes | | | no | | | |
| AB980e | | | | | yes | | | yes | | | no | | | |
| AB980f | | | | | yes | | | yes | | | no | | | |
| AB980g | | | | | yes | | | yes | | | no | | | |
| AB980h | | yes | | no | no | yes | | yes | | | no | no | | yes |
| AB980i | | | | | yes | | | yes | | | no | | | |
| AB980k | | | | | yes | | | yes | | | no | | | |
| AB1100a | | yes | | | no | no | | yes | no | | no | | | yes |
| AB1100b | | yes | | | no | no | | yes | no | | no | | | yes |
| AB1100c | | yes | | | yes | | | yes | | | no | | | |
| AB1100d | | yes | | | no | no | | yes | no | | no | | | yes |
| AB1100e | | yes | | | no | ? | | yes | no | | no | | | yes |
| AB1100f | | yes | | | yes | | | yes | | | no | | | |
| AB1100g | | yes | | | no | no | | yes | no | | no | | | yes |
| AB1100h | | yes | | no | no | yes | | yes | | | no | no | | |
| AB1100i | | yes | | | no | no | | yes | no | | no | | | yes |
| AB1100k | | yes | | | yes | | | yes | | | no | | | |
| AB1900a | no | yes | no | | yes | | yes | yes | | | no | | | |
| AB1900b | no | | | | yes | | | yes | | | no | | | |
| AB1900c | no | yes | | | no | no | yes | yes | no | | no | | | yes |
| AB1900d | no | yes | | | no | no | | yes | no | | no | | | yes |
| AB1900e | no | | | | yes | | | yes | | | no | | | |
| AB1900f | no | | | | yes | | | yes | | | no | | | |
| AB1900g | no | | | | yes | | | yes | | | yes | | | |
| AB1900h | no | | | | yes | | | yes | | | no | | | |
| AB1900i | no | | | | yes | | | yes | | | yes | | | |
| AB1900j | no | | | | yes | | | yes | | | yes | | | |
| AB2300a | no | yes | no | | no | no | | | no | | yes | | | yes |
| AB2300b | no | yes | no | | no | no | | | no | | no | | | yes |
| AB2300c | yes | yes | | | no | no | | | no | | yes | | | yes |
| AB2300d | no | yes | | | no | no | | | no | | no | | | yes |
| AB2300e | no | | | | yes | | | | | | yes | | | |
| AB2300f | no | yes | | | no | no | | | no | | no | | | yes |
| AB2300g | yes | yes | | | no | no | | | no | | no | | | yes |
| AB2300h | yes | | | | yes | | | | | | no | | | |
| AB2300i | no | | | | yes | | | | | | yes | | | |
| AB2300k | no | yes | | | no | no | | | no | | yes | | | yes |
| AB2800a | yes | yes | yes | | yes | no | yes | yes | | yes | | no | yes | yes |
| AB2800b | no | | no | yes | no | yes | | yes | | yes | | yes | yes | yes |
| AB2800c | no | yes | no | yes | no | yes | | | | | yes | yes | yes | yes |
| AB2800d | yes | | yes | | yes | | | | | | yes | | | |
| AB2800e | yes | | no | | yes | | | | | | no | | | |

3. Transformation rates measured within and out reactors

A genomic DNA containing a resistance to kanamycin (gDNAkan^R) was used in competence assays. gDNAkan^R was obtained by transforming wild type bacteria with a cassette containing the resistance to kanamycin (see below). Genomic DNA was extracted from resistant cells by ethanol precipitation after lysis and separation with chloroform.

Competence was tested with a 20 days evolved bacteria during a control experiment. Four different conditions were tested: within the chemostat (air bubbles aeration + glass wall), in 50ml Falcon with air input tube (air bubbles aeration + plastic wall), in glass tube in shaker (aeration by shaking + glass wall) and in 2ml eppendorf microtubes (aeration by shaking + plastic wall).

For the three last conditions, the bacteria subjected to transformation were sampled from the chemostat. In each case, transformation assay was performed in selection medium supplemented with gDNAkan^R 1 µg/ml and 20 mM KNO₃ pulse enhancing transformation rate. Bacteria were incubated at 30°C for 4.5 h. Samples were plated at different dilutions on LB supplemented with 50 µg/ml kanamycin, and incubated at 30°C overnight. The OD measured before plating allowed to estimate the number of plated cells. The rate of transformation was estimated by the ratio between the number of resistant colonies and the number of plated cells.

In chemostat and in Falcon supplied with air by an input tube, the competence was not detectable (below 10⁻⁹). Transformation only took place in condition of aeration by shaking (glass tube and eppendorf microtube), suggesting that air bubbles limits the transformation process.

Protocols used to obtain bacteria resistant to kanamycin: A cassette containing kanamycin resistance was first produced as follows: the gene conferring resistance to kanamycin (kan) was amplified by PCR from the plasmid pUC4K (GE Healthcare), and two fragments (named here H1 and H2) overlapping the extremities of the mutS gene were PCR-amplified with the *Acinetobacter baylyi* genome as template. Primers were designed to introduce sequences allowing subsequent hybridization of the fragments for the later cassette assembly (the sequences added for hybridization are underlined): ttttgatgatgcacgaaaagcgcaaac (left primer H1), cgcgggcgggcgttcttaagtcacgtggtc (right primer H1), gcgggcccgcccttgaattggctacaatc (left primer H2), tttgtgcttccttgatcacgcttgaggt (right primer H2), cggccgcccgcgttgcgggaagatgcgtga (left primer kan), ggcgggcccgcttgggaaagccacgttgt (right primer kan). Each fragment were about 1 kb long. One extremity of each fragment (H1, H2), and the two extremities of the kan fragment were partially digested on the 3' end via the partial exonuclease action of the T4 polymerase in the presence of a limited set of dNTP's: the fragments were each treated via 0.5 µl T4 DNA polymerase, DTT 2.5 mM, dATP 2.5 mM, 10 µl PCR product in 20 µl buffer T4 DNA polymerase during 30 min at 22°C. In such condition, the T4 polymerases chews away the 3' ends unless it encounters an A. The polymerase was then inactivated by a 20 min incubation at 75°C. The fragments with neo-formed cohesive protruding single-stranded ends were then hybridized by incubating 6µl H1, 6 µl H2 and 2 µl kan during 3 hours at 45°C, and ligated overnight at 4°C (0.5 µl T4 DNA ligase, 14 µl hybridized mix in 20 µl ligase buffer). The ligation was inactivated by a 10 min incubation at 65°C. The ligation of the three fragments was confirmed by gel electrophoresis. The cassette was then amplified by PCR with the pfu DNA polymerase (Fermentas).

ABwt was transformed as follows: 20 µl saturated culture of bacteria were incubated with 300 µl LB 2.5 h at 30°C with shaking; 95 µl PCR products were added and incubated with shaking 2 h at 30°C; cells were selected on LB plate supplemented with 50 µg/ml kanamycin. A few clones were isolated and the presence of the cassette in the genome was verified by PCR.

4. Calibration and measurements of nitrate concentration in chemostat

In order to estimate the concentration of nitrate within the reactor during the experiment, samples

were collected from the output of the reactor, centrifuged to remove bacteria and snap-frozen. For measurements, samples were thawed and the nitrate was reduced to nitrite; samples underwent then a Griess test. In detail, the Griess Reaction Nitrate Kit from World Precision Instrument was used: the reduction of nitrate to nitrite is obtained by a reaction using metallic particles of Cadmium; a Griess test follows, which results in a red pink color (540 nm) which level is related to the nitrite concentration. Spectral measurements were done at 600 nm using a photometer (Biophotometer, Eppendorf). Phosphate was found to inhibit the reduction of nitrate. We circumvented this effect by first precipitating Phosphate via the addition of 15% CaCl_2 1 M, and with 10% NaOH 1 M to samples. Samples were then centrifuged one minute in micro-centrifuge at 13000 RPM, the supernatant was recovered and underwent a reduction by cadmium as described in the Nitralyzer kit protocol (World Precision Instrument).

Standards containing minimal medium with different concentration of nitrate were used to calibrate measurements (figure S1). The data were well fitted by polynomial function. Each sample was duplicated and the polynomial fit was used to estimate nitrate concentration in samples. The minimal level of detectable nitrate was estimated to be $\sim 2 \mu\text{M}$ nitrate.

In an effort to accurately measure a saturation constant, we started another chemostat experiment under the same conditions of the above evolution experiments but with the clone of AB85a for initial seeding. The reactor flow-out was sampled for different dilution rates, the bacteria were eliminated by centrifugation and the sample were snap-frozen and stored for subsequent analysis. This experiment was led during less than 25 generations to limit the evolution of the strain. At four dilution rates (0.19, 0.50, 0.54 and 0.60 h^{-1}), no nitrate was detectable with our experimental procedure. For $D=0.70 \text{ h}^{-1}$ a value of $82 \pm 2 \mu\text{M}$ nitrate was measured.

The same type of experiment has been made with the clone of AB215b, during about 55 generations. For $D=0.60 \text{ h}^{-1}$ and $D=0.78 \text{ h}^{-1}$, no nitrate was detectable in the flow-out of the reactor, and for $D=0.88 \text{ h}^{-1}$, we measured a concentration of $82 \pm 1 \mu\text{M}$. However, the chemostat was possibly out of equilibrium for this latter value of D because the draining threshold has been estimated to be about 0.85 in this period. It can be deduced from the very low nitrate ($< 2 \mu\text{M}$) in the reactor in a wide range of dilution rate that $> 99.8\%$ of this nitrate present in the fresh medium (1mM) is captured by the bacteria.

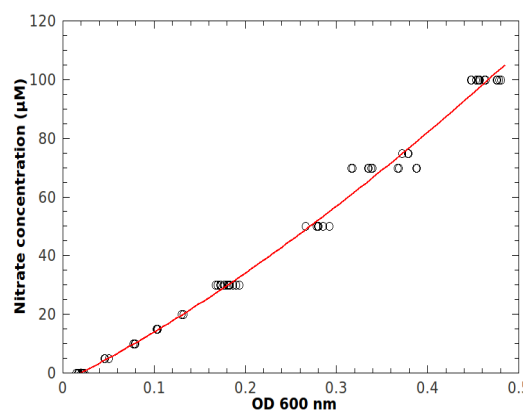


Figure S1: calibration; each point corresponds to a standard sample undergoing the Griess protocol and measured by spectroscopy at 600 nm.

5. Biolog assay

The comparative catabolic activities of the starting strain ABwt and of the endpoint clone AB2800a were determined by the Phenotype MicroArray Services of Biolog, Hayward, CA. The two strains were tested on different carbon sources, nitrogen sources, phosphorus and sulphur sources, nutrient

supplements, and peptides nitrogen sources.

Concerning carbon sources, there were lost phenotypes for L-rhamnose, D-psicose, glucuronamide, D-Galactose and L-Fucose. Concerning nitrogen sources, there were globally positive differences; AB2800a had a preference for sugar amines and amino acids (including L-tryptophan); most of the utilized dipeptides contain the amino acid tryptophan. There were notably gained phenotypes for L-tryptophan, pro-trp, phe-trp, and D-aspartic acid. Although just below threshold, trp-phe also exhibited a positive difference. There were however no difference found for growth on nitrate between the ancestor and the evolved strain: this is understood by the absence of molybdenum traces in the Biolog medium (other than as possible impurity). The need of Traces elements containing Molybdenum for growth was indeed verified in a batch culture experiment (data not shown) for which we compared the growth of strain AB2800a in a nitrate medium with or without Traces elements: after 24 hours, there was no growth in the medium without Traces elements whereas the culture with Traces elements was saturated.

There were negative differences or lost phenotypes for parabanic acid, guanine, and ala-leu. Concerning phosphorus and sulphur sources, there were positive differences or gained phenotypes for cysteamine-S-phosphate, dithiophosphate (P-sources), thiophosphate, taurocholic acid, and D,L-lipoamide (S-sources).

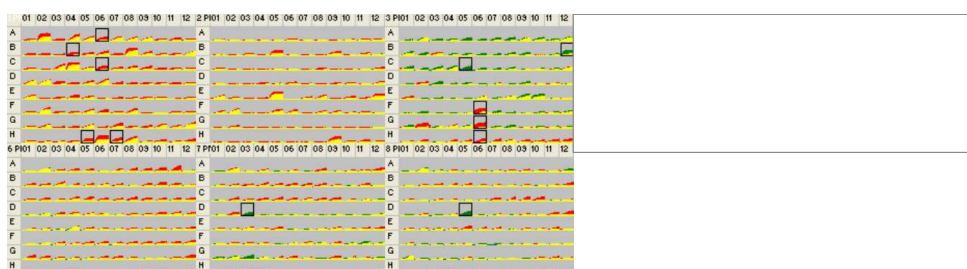


Figure S2: Each plate corresponds to a type of nutrient source tested; each square on a plate corresponds to a different nutrient. Top, from left to right: the two first plates test Carbon sources, the following plates test Nitrogen sources, Phosphorus and Sulfur sources, Nutrients supplements. Bottom: the three plates test Peptide Nitrogen sources.

AB2800a was compared to ABwt. When there are similar phenotypic responses in both cell lines the kinetic curve is yellow, different phenotypic responses are expressed respectively in red for ABwt and green for AB2800a. The relevant differences are black squared by Biolog. For more details, please visit the web site of Biolog (<http://www.biolog.com/index.shtml>).

Bibliographie

- [1] S. Atsumi, TY Wu, IM Machado, WC Huang, PY Chen, M. Pellegrini, and JC Liao. Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*. *Molecular systems biology*, 6 :449, 2010.
- [2] B. Averhoff and A. Friedrich. Type IV pili-related transformation systems : DNA transport in mesophilic and thermophilic bacteria. *Archives of Microbiology*, 180 :385–393, 2003.
- [3] B. Averhoff and I. Graf. *Acinetobacter Molecular Biology*, chapter The natural transformation system of *Acinetobacter baylyi* : A unique transport machinery, pages 119–139. Caister Academic Press,, 2008.
- [4] J.M. Bacher, D. Metzgar, and V. Crécy-Lagard. Rapid evolution of diminished transformability in *Acinetobacter baylyi*. *Journal of Bacteriology*, 188(24) :8534–8542, 2006.
- [5] H. Bachmann, M.J.C. Starrenburg, D. Molenaar, M. Kleerebezem, and J.E.T. van Hylckama Vlieg. Microbial domestication signatures of *Lactococcus lactis* can be reproduced by experimental evolution. *Genome research*, 22(1) :115–124, 2012.
- [6] D.A. Baltrus, K. Guillemin, and P.C. Phillips. Natural transformation increases the rate of adaptation in the human pathogen *Helicobacter pylori*. *Evolution*, 62(1) :39–49, 2008.
- [7] V. Barbe, D. Vallenet, N. Fonknechten, A. Kreimeyer, S. Oztas, L. Labarre, S. Cruveiller, C. Robert, S. Duprat, P. Wincker, L.N. Ornston, J. Weissenbach, P. Marlière, G.N. Cohen, and C. Médigue. Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Research*, 32(19) :5766–5779, 2004.
- [8] JE Barrick and RE Lenski. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 74, pages 119–129, 2009.
- [9] J.E. Barrick, D.S. Yu, S.H. Yoon, H. Jeong, T.K. Oh, D. Schneider, R.E. Lenski, and J.F. Kim. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature*, 461(7268) :1243–1247, 2009.

- [10] G. Bell. *The masterpiece of nature : the evolution and genetics of sexuality*. Springer, 1982.
- [11] Z.D. Blount, C.Z. Borland, and R.E. Lenski. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *PNAS*, 105(23) :7899–7906, 2008.
- [12] E.T. Boder, K.S. Midelfort, and K.D. Wittrup. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proceedings of the National Academy of Sciences*, 97(20) :10701, 2000.
- [13] M. Brigulla and W. Wackernagel. Molecular aspects of gene transfer and foreign dna acquisition in prokaryotes with regard to safety issues. *Applied microbiology and biotechnology*, 86(4) :1027–1041, 2010.
- [14] W. Broothaerts, H.J. Mitchell, B. Weir, S. Kaines, L.M.A. Smith, W. Yang, J.E. Mayer, C. Roa-Rodríguez, and R.A. Jefferson. Gene transfer to plants by diverse species of bacteria. *Nature*, 433(7026) :629–633, 2005.
- [15] A. Buckling, R.C. Maclean, M.A. Brockhurst, and N. Colegrave. The Beagle in a bottle. *Nature*, 457(7231) :824–829, 2009.
- [16] S. Busch, C. Rosenplänter, and B. Averhoff. Identification and characterization of ComE and ComF, two novel pilin-like competence factors involved in natural transformation of *Acinetobacter* sp. BD413. *Applied and Environmental Microbiology*, 65(10) :4568–4574, 1999.
- [17] L. Chao. Fitness of RNA virus decreased by Muller’s ratchet. *Nature*, 348(6300) :454–455, 1990.
- [18] L. Chao. Evolution of sex in RNA viruses. *Trends in Ecology & Evolution*, 7(5) :147–151, 1992.
- [19] L. Chao, Thu T. Tran, and Thutrang T. Tran. The advantage of sex in the RNA virus $\phi 6$. *Genetics*, 147(3) :953–959, 1997.
- [20] I. Chen and D. Dubnau. DNA uptake during bacterial transformation. *Nature Reviews Microbiology*, 2 :241–249, 2004.
- [21] L.M. Chevin. On measuring selection in experimental evolution. *Biology letters*, 7(2) :210–213, 2011.
- [22] H.H. Chou, H.C. Chiu, N.F. Delaney, D. Segrè, and C.J. Marx. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science*, 332(6034) :1190–1192, 2011.
- [23] Hsin-Hung Chou, Julia Berthet, and Christopher J. Marx. Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation. *PLoS Genet*, 5(9) :e1000652, 09 2009.
- [24] N. Colegrave. Sex releases the speed limit on evolution. *Nature*, 420 :664–666, 2002.
- [25] T.M. Conrad, A.R. Joyce, M.K. Applebee, C.L. Barrett, B. Xie, Y. Gao, and B.Ø. Palsson. Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biology*, 10(10) :R118, 2009.

- [26] T.F. Cooper. Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli*. *PLoS biology*, 5(9) :e225, 2007.
- [27] T.F. Cooper, D.E. Rozen, and R.E. Lenski. Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 100(3) :1072–1077, 2003.
- [28] E. Crozat, C. Winkworth, J. Gaffé, P.F. Hallin, M.A. Riley, R.E. Lenski, and D. Schneider. Parallel genetic and phenotypic evolution of DNA superhelicity in experimental populations of *Escherichia coli*. *Molecular biology and evolution*, 27(9) :2113–2128, 2010.
- [29] M.H. Dai and S.D. Copley. Genome shuffling improves degradation of the anthropogenic pesticide pentachlorophenol by *Sphingobium chlorophenolicum* atcc 39723. *Applied and environmental microbiology*, 70(4) :2391–2397, 2004.
- [30] C. Darwin. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. John Murray, London, 1859.
- [31] V. de Crécy-Lagard, J. Bellalou, R. Mutzel, and P. Marlière. Long term adaptation of a microbial population to a permanent metabolic constraint : overcoming thymineless death by experimental evolution of *Escherichia coli*. *BMC Biotechnology*, 1(1) :10, 2001.
- [32] F. De La Iglesia and S.F. Elena. Fitness declines in tobacco etch virus upon serial bottleneck transfers. *Journal of virology*, 81(10) :4941–4947, 2007.
- [33] J. De Vries, T. Herzfeld, and W. Wackernagel. Transfer of plasmid dna from tobacco to the soil bacterium *Acinetobacter* sp. by natural transformation. *Molecular microbiology*, 53(1) :323–334, 2004.
- [34] F.B. Dean, S. Hosono, L. Fang, X. Wu, A.F. Faruqi, P. Bray-Ward, Z. Sun, Q. Zong, Y. Du, J. Du, et al. Comprehensive human genome amplification using multiple displacement amplification. *Proceedings of the National Academy of Sciences*, 99(8) :5261, 2002.
- [35] M.M. Desai and D.S. Fisher. Beneficial mutation–selection balance and the effect of linkage on positive selection. *Genetics*, 176(3) :1759–1798, 2007.
- [36] J.R. Dettman, N. Rodrigue, A.H. Melnyk, A. Wong, S.F. Bailey, and R. Kassen. Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Molecular Ecology*, 21(9) :2058–2077, 2012.
- [37] W. Joseph Dickinson. Synergistic fitness interactions and a high frequency of beneficial changes among mutations accumulated under relaxed selection in *Saccharomyces cerevisiae*. *Genetics*, 178(3) :1571–1578, 2008.
- [38] J.W. Drake. A constant rate of spontaneous mutation in dna-based microbes. *Proceedings of the National Academy of Sciences, USA*, 88(16) :7160–7164, 1991.
- [39] J.W. Drake. Contrasting mutation rates from specific-locus and long-term mutation-accumulation procedures. *G3 : Genes/ Genomes/ Genetics*, 2(4) :483–485, 2012.

- [40] Santiago F. Elena, Vaughn S. Cooper, and Richard E. Lenski. Punctuated evolution caused by selection of rare beneficial mutations. *Science*, 272(5269) :1802–1804, 1996.
- [41] S.F. Elena, R.E. Lenski, et al. Evolution experiments with microorganisms : the dynamics and genetic bases of adaptation. *Nature reviews : Genetics*, 4(6) :457–469, 2003.
- [42] A. Eyre-Walker and P.D. Keightley. The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, 8(8) :610–618, 2007.
- [43] T. Ferenci. Bacterial physiology, regulation and mutational adaptation in a chemostat environment. *Advances in microbial physiology*, 53 :169, 2008.
- [44] R.A. Fisher. The genetical theory of natural selection. 1930.
- [45] S.J. Foster and B.J. Monahan. Whole genome amplification from filamentous fungi using phi29-mediated multiple displacement amplification. *Fungal Genetics and Biology*, 42(5) :367–375, 2005.
- [46] A. Friedrich, T. Hartsch, and B. Averhoff. Natural transformation in mesophilic and thermophilic bacteria : identification and characterization of novel, closely related competence genes in *Acinetobacter* sp. strain BD413 and *Thermus thermophilus* HB27. *Applied and Environmental Microbiology*, 67(7) :3140–3148, 2001.
- [47] J. Gaffe, C. McKenzie, R. Maharjan, E. Coursange, T. Ferenci, and D. Schneider. Insertion sequence-driven evolution of *Escherichia coli* in chemostats. *Journal of Molecular Evolution*, 72(4) :398–412, 2011.
- [48] P.J. Gerrish and R.E. Lenski. The fate of competing beneficial mutations in an asexual population. *Genetica*, 102 :127–144, 1998.
- [49] M.R. Goddard, H.C.J. Godfray, and A. Burt. Sex increases the efficacy of natural selection in experimental yeast populations. *Nature*, 434(7033) :636–640, 2005.
- [50] O. Gohl, A. Friedrich, M. Hoppert, and B. Averhoff. The thin pili of *Acinetobacter* sp. strain BD413 mediate adhesion to biotic and abiotic surfaces. *Applied and Environmental Microbiology*, 72(2) :1394–1401, 2006.
- [51] S.J. Gould. Wonderful life : The burgess shale and the meaning of history, 1989.
- [52] S.J. Gould. *The structure of evolutionary theory*. Belknap Press, 2002.
- [53] J.C. Gray and M.R. Goddard. Sex enhances adaptation by unlinking beneficial from detrimental mutations in experimental yeast populations. *BMC Evolutionary Biology*, 12(1) :43, 2012.
- [54] J.A. Heinemann, G.F. Sprague Jr, et al. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature*, 340(6230) :205–209, 1989.
- [55] J. Henrichsen. Twitching motility. *Annual review of microbiology*, 37 :81–93, 1983.
- [56] C.D. Herring, A. Raghunathan, C. Honisch, T. Patel, M.K. Applebee, A.R. Joyce, T.J. Albert, F.R. Blattner, D. Van den Boom, C.R. Cantor, et al.

- Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nature genetics*, 38(12) :1406–1412, 2006.
- [57] R. Hershberg and D.A. Petrov. Evidence that mutation is universally biased towards at in bacteria. *PLoS genetics*, 6(9) :e1001115, 2010.
- [58] C. Herzberg, A. Friedrich, and B. Averhoff. ComB, a novel competence gene required for natural transformation of *Acinetobacter* sp. BD413 : identification, characterization, and analysis of growth-phase-dependent regulation. *Archives of Microbiology*, 173(3) :220–228, 2000.
- [59] F. Hildebrand, A. Meyer, and A. Eyre-Walker. Evidence of selection upon genomic gc-content in bacteria. *PLoS genetics*, 6(9) :e1001107, 2010.
- [60] N. Hülter and W. Wackernagel. Double illegitimate recombination events integrate DNA segments through two different mechanisms during natural transformation of *Acinetobacter baylyi*. *Molecular Microbiology*, 67(5) :984–995, 2008.
- [61] S. Hosono, A.F. Faruqi, F.B. Dean, Y. Du, Z. Sun, X. Wu, J. Du, S.F. Kingsmore, M. Egholm, and R.S. Lasken. Unbiased whole-genome amplification directly from clinical samples. *Genome Research*, 13(5) :954–964, 2003.
- [62] M. Imhof and C. Schlötterer. Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. *Proceedings of the National Academy of Sciences*, 98(3) :1113, 2001.
- [63] T. Ishoey, T. Woyke, R. Stepanauskas, M. Novotny, and R.S. Lasken. Genomic sequencing of single microbial cells from environmental samples. *Current opinion in microbiology*, 11(3) :198–204, 2008.
- [64] O. Johnsborg, V. Eldhom, and L.S. Havarstein. Natural genetic transformation : prevalence, mechanisms and function. *Research in Microbiology*, 158(10) :767–778, 2007.
- [65] O. Kaltz and G. Bell. The ecology and genetics of fitness in *Chlamydomonas*. xii. repeated sexual episodes increase rates of adaptation to novel environments. *Evolution*, 56(9) :1743–1753, 2002.
- [66] R. Kassen and T. Bataillon. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nature genetics*, 38(4) :484–488, 2006.
- [67] E. Kay, T.M. Vogel, F. Bertolla, R. Nalin, and P. Simonet. In situ transfer of antibiotic resistance genes from transgenic (transplastomic) tobacco plants to bacteria. *Applied and environmental microbiology*, 68(7) :3345–3351, 2002.
- [68] P.J. Keeling and J.D. Palmer. Horizontal gene transfer in eukaryotic evolution. *Nature Reviews Genetics*, 9(8) :605–618, 2008.
- [69] D. Kersulyte, H. Chalkauskas, and D.E. Berg. Emergence of recombinant strains of *Helicobacter pylori* during human infection. *Molecular microbiology*, 31(1) :31–43, 1999.

- [70] A.I. Khan, D.M. Dinh, D. Schneider, R.E. Lenski, and T.F. Cooper. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science*, 332(6034) :1193–1196, 2011.
- [71] T.T. Kibota and M. Lynch. Estimate of the genomic mutation rate deleterious to overall fitness in *e. coll.* *Nature*, 381 :694–696, 1996.
- [72] C. Kimchi-Sarfaty, J.M. Oh, I.W. Kim, Z.E. Sauna, A.M. Calcagno, S.V. Ambudkar, and M.M. Gottesman. A silent polymorphism in the MDR1 gene changes substrate specificity. *Science*, 315(5811) :525–528, 2007.
- [73] M. Kimura. Recent development of the neutral theory viewed from the wrightian tradition of theoretical population genetics. *Proceedings of the National Academy of Sciences*, 88(14) :5969, 1991.
- [74] J.C. Koella. The tangled bank : the maintenance of sexual reproduction through competitive interactions. *Journal of Evolutionary Biology*, 1(2) :95–116, 1988.
- [75] A.S. Kondrashov. Deleterious mutations as an evolutionary factor : 1. the advantage of recombination. *Genetical Research*, 44(02) :199–217, 1984.
- [76] A.S. Kondrashov. Deleterious mutations and the evolution of sexual reproduction. *Nature*, 336(6198) :435–440, 1988.
- [77] E.V. Koonin, K.S. Makarova, and L. Aravind. Horizontal gene transfer in prokaryotes : Quantification and classification 1. *Annual Reviews in Microbiology*, 55(1) :709–742, 2001.
- [78] B. Langmead and S.L. Salzberg. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9 :357–359, 2012.
- [79] B. Langmead, C. Trapnell, M. Pop, S.L. Salzberg, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*, 10(3) :R25, 2009.
- [80] M. Le Gac, J. Plucain, T. Hindré, R.E. Lenski, and D. Schneider. Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 109(24) :9487–9492, 2012.
- [81] B. Lehner. Molecular mechanisms of epistasis within and between genes. *Trends in Genetics*, 27(8) :323–331, 2011.
- [82] J. Lehtonen, M.D. Jennions, and H. Kokko. The many costs of sex. *Trends in Ecology & Evolution*, 27(3) :172–178, 2011.
- [83] R.E. Lenski. Evolution in action : a 50,000-generation salute to Charles Darwin. *Microbe*, 6 :30–33, 2011.
- [84] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16) :2078–2079, 2009.
- [85] C. Link, S. Eicknerjager, D. Porstendörfer, and B. Averhoff. Identification and characterization of a novel competence gene, comC, required for DNA binding and uptake in *Acinetobacter* sp. strain BD413. *Journal of Bacteriology*, 180(6) :1592–1595, 1998.

- [86] Laurence Loewe, Volker Textor, and Siegfried Scherer. High deleterious genomic mutation rate in stationary phase of escherichia coli. *Science*, 302(5650) :1558–1560, 2003.
- [87] M.G. Lorenz, K. Reipschläger, and W. Wackernagel. Plasmid transformation of naturally competent *Acinetobacter calcoaceticus* in non-sterile soil extract and groundwater. *Archives of Microbiology*, 157 :355–360, 1992.
- [88] M.G. Lorenz and W. Wackernagel. Bacterial gene transfert by natural genetic transformation in the environment. *Microbiological Reviews*, 58(3) :563–602, 1994.
- [89] S.T. Lovett. Encoded errors : mutations and rearrangements mediated by misalignment at repetitive dna sequences. *Molecular microbiology*, 52(5) :1243–1253, 2004.
- [90] R.L. Malmberg. The evolution of epistasis and the advantage of recombination in populations of bacteriophage T4. *Genetics*, 86(3) :607–621, 1977.
- [91] J.S. Mattick. Type IV pili and twitching motility. *Annual Reviews in Microbiology*, 56(1) :289–314, 2002.
- [92] J. Maynard Smith. A short-term advantage for sex and recombination through sib-competition. *Journal of theoretical biology*, 63(2) :245–258, 1976.
- [93] J. Maynard Smith. *The evolution of sex*. Cambridge Univ Press, 1978.
- [94] H.H. McAdams, B. Srinivasan, and A.P. Arkin. The evolution of genetic regulatory systems in bacteria. *Nature Reviews Genetics*, 5(3) :169–178, 2004.
- [95] D. Metzgar, J.M. Bacher, V. Pezo, J. Reader, V. Döring, P. Schimmel, P. Marlière, and V. Crécy-Lagard. *Acinetobacter* sp. ADP1 : an ideal model organism for genetic analysis and genome engineering. *Nucleic Acids Research*, 32(19) :5780–5790, 2004.
- [96] I. Milne, M. Bayer, L. Cardle, P. Shaw, G. Stephen, F. Wright, and D. Marshall. Tablet-next generation sequence assembly visualization. *Bioinformatics*, 26(3) :401–402, 2010.
- [97] J Monod. La technique de culture continue ; théorie et applications. *Annales de l'Institut Pasteur*, 79(4) :390–410, 1950.
- [98] JC Moore and FH Arnold. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nature Biotechnology*, 14(4) :458–467, 1996.
- [99] H.J. Muller. Some genetic aspects of sex. *The American Naturalist*, 66(703) :118–138, 1932.
- [100] H.J. Muller. The relation of recombination to mutational advance. *Mutation Research*, 106(1) :2–9, 1964.
- [101] H. Ochman, S. Elwyn, and NA Moran. Calibrating bacterial evolution. *Proceedings of the National Academy of Sciences, USA*, 96(22) :12638–12643, 1999.
- [102] H. Ochman, J.G. Lawrence, and E.A. Groisman. Lateral gene transfer and the nature of bacterial innovation. *Nature*, 405 :299–304, 2000.

- [103] T. Ohta. The nearly neutral theory of molecular evolution. *Annual Review of Ecology and Systematics*, 23 :263–286, 1992.
- [104] H.A. Orr. The genetic theory of adaptation : a brief history. *Nature Reviews Genetics*, 6(2) :119–127, 2005.
- [105] Elizabeth A Ostrowski, Robert J Woods, and Richard E Lenski. The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. *Proceedings of the Royal Society B : Biological Sciences*, 275(1632) :277–284, 2008.
- [106] S.P. Otto. The evolutionary enigma of sex. *The American Naturalist*, 174 :S1–S14, 2009.
- [107] S.P. Otto, T. Lenormand, et al. Resolving the paradox of sex and recombination. *Nature Reviews Genetics*, 3(4) :252–261, 2002.
- [108] C. Pal, M.D. Maciá, A. Oliver, I. Schachar, and A. Buckling. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature*, 450(7172) :1079–1081, 2007.
- [109] R. Palmen, P. Buijsman, and K. J. Hellingwerf. Physiological regulation of competence induction for natural transformation in *Acinetobacter calcoaceticus*. *Archive of Microbiology*, 162 :344–351, 1994.
- [110] R. Palmen and K. J. Hellingwerf. Uptake and processing of *Acinetobacter calcoaceticus* - a review. *Gene*, 192 :179–190, 1997.
- [111] R. Palmen, B. Vosman, P. Buijsman, C. K. D. Breek, and K. J. Hellingwerf. Physiological characterization of natural transformation in *Acinetobacter calcoaceticus*. *Journal of General Microbiology*, 139 :295–305, 1993.
- [112] D. Papadopoulos, D. Schneider, J. Meier-Eiss, W. Arber, R.E. Lenski, and M. Blot. Genomic evolution during a 10,000-generation experiment with bacteria. *Proceedings of the National Academy of Sciences*, 96(7) :3807–3812, 1999.
- [113] S. Paterson, T. Vogwill, A. Buckling, R. Benmayor, A.J. Spiers, N.R. Thomson, M. Quail, F. Smith, D. Walker, B. Libberton, et al. Antagonistic coevolution accelerates molecular evolution. *Nature*, 464(7286) :275–278, 2010.
- [114] R. Patnaik, S. Louie, V. Gavrilovic, K. Perry, W.P.C. Stemmer, C.M. Ryan, S. Del Cardayré, et al. Genome shuffling of *Lactobacillus* for improved acid tolerance. *Nature biotechnology*, 20(7) :707–712, 2002.
- [115] J.R. Peck. A ruby in the rubbish : beneficial mutations, deleterious mutations and the evolution of sex. *Genetics*, 137(2) :597–606, 1994.
- [116] L. Perfeito, L. Fernandes, C. Mota, and I. Gordo. Adaptive mutations in bacteria : high rate and small effects. *Science*, 317(5839) :813–815, 2007.
- [117] Gabriel G. Perron, Alexander E. G. Lee, Yun Wang, Wei E. Huang, and Timothy G. Barraclough. Bacterial recombination promotes the evolution of multi-drug-resistance in functionally diverse populations. *Proceedings of the Royal Society B : Biological Sciences*, 279(1733) :1477–1484, 2012.

- [118] P. C. Phillips. Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Reviews Genetics*, 9(11) :855–867, 2008.
- [119] S. J. Pirt. *Principles of microbe and cell cultivation*. Blackwell Scientific Publications, Oxford, 1975.
- [120] F.J. Poelwijk, D.J. Kiviet, D.M. Weinreich, and S.J. Tans. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature*, 445(7126) :383–386, 2007.
- [121] D. Porstendörfer, U. Drotschmann, and B. Averhoff. A novel competence gene, comP, is essential for natural transformation of *Acinetobacter* sp. strain BD413. *Applied and Environmental Microbiology*, 63(11), 1997.
- [122] D. Porstendörfer, O. Gohl, F. Mayer, and B. Averhoff. ComP, a pilin-like protein essential for natural competence in *Acinetobacter* sp. strain BD413 : regulation, modification, and cellular localization. *Journal of Bacteriology*, 182(13) :3673–3680, 2000.
- [123] PB Rainey and M. Travisano. Adaptive radiation in a heterogeneous environment. *Nature*, 394(6688) :69–72, 1998.
- [124] L.R. Ray, K. Harms, O.G. Wikmark, I. Starikova, and P.J. Johnsen. Sexual isolation in *Acinetobacter baylyi* is locus-specific and varies 10,000 fold over the genome. *Genetics*, 182(4) :1165–1181, 2009.
- [125] R.J. Redfield. Do bacteria have sex? *Nature Reviews Genetics*, 2(8) :634–639, 2001.
- [126] D.E. Rozen, J. De Visser, and P.J. Gerrish. Fitness effects of fixed beneficial mutations in microbial populations. *Current biology*, 12(12) :1040–1045, 2002.
- [127] P.D. Sniegowski and P.J. Gerrish. Beneficial mutations and the dynamics of adaptation in asexual populations. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 365(1544) :1255–1263, 2010.
- [128] W.P.C. Stemmer. DNA shuffling by random fragmentation and reassembly : In vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA*, 91 :10747–10751, 1994.
- [129] Daniel M. Stoebel and Charles J. Dorman. The effect of mobile element IS10 on experimental regulatory evolution in *Escherichia coli*. *Molecular Biology and Evolution*, 27(9) :2105–2112, 2010.
- [130] O. Tenaillon, A. Rodríguez-Verdugo, R.L. Gaut, P. McDonald, A.F. Bennett, A.D. Long, and B.S. Gaut. The molecular diversity of adaptive convergence. *Science*, 335(6067) :457–461, 2012.
- [131] M Travisano, JA Mongold, AF Bennett, and RE Lenski. Experimental tests of the roles of adaptation, chance, and history in evolution. *Science*, 267(5194) :87–90, 1995.
- [132] D S Treves, S Manning, and J Adams. Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Molecular Biology and Evolution*, 15(7) :789–797, 1998.

- [133] L. Van Valen. A new evolutionary law. *Evolutionary theory*, 1(1) :1–30, 1973.
- [134] W. Wackernagel and J. De Vries. Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *PNAS*, 99(4) :2094–2099, 2002.
- [135] L. Wang, B. Spira, Z. Zhou, L. Feng, R.P. Maharjan, X. Li, F. Li, C. McKenzie, P.R. Reeves, and T. Ferenci. Divergence involving global regulatory gene mutations in an *Escherichia coli* population evolving under phosphate limitation. *Genome Biology and Evolution*, 2 :478–487, 2010.
- [136] S. Wielgoss, J.E. Barrick, O. Tenaillon, S. Cruveiller, B. Chane-Woon-Ming, C. Médigue, R.E. Lenski, and D. Schneider. Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*. *G3 : Genes, Genomes, Genetics*, 1(3) :183–186, 2011.
- [137] G.C. Williams. *Sex and evolution*. Princeton University Press (Princeton, NJ), 1975.
- [138] S. Wright. Evolution in mendelian populations. *Genetics*, 16(2) :97, 1931.
- [139] D.R. Zerbino and E. Birney. Velvet : algorithms for de novo short read assembly using de bruijn graphs. *Genome research*, 18(5) :821–829, 2008.
- [140] C. Zeyl and G. Bell. The advantage of sex in evolving yeast populations. *Nature*, 388(6641) :465–468, 1997.
- [141] C. Zeyl and J.A.G.M. DeVisser. Estimates of the rate and distribution of fitness effects of spontaneous mutation in *Saccharomyces cerevisiae*. *Genetics*, 157(1) :53–61, 2001.
- [142] Y.X. Zhang, K. Perry, V.A. Vinci, K. Powell, W.P.C. Stemmer, and S.B. Del Cardayré. Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature*, 415(6872) :644–646, 2002.