

Sommaire

Etude bibliographique	1
I. Escherichia coli : métabolisme des sucres.....	3
I. 1. Escherichia coli : description générale.....	3
I. 2. Métabolisme Carboné Central d' <i>Escherichia coli</i>	5
I. 2. A Glycolyse (Embden–Meyerhof pathway) et gluconéogenèse	7
I. 2. B Cycle de Kreb (TCA).....	7
I. 2. C Voie des pentoses phosphates.....	7
I. 2. D Voie d'Entner-Doudoroff.....	8
I. 3. Métabolisme du glucose	8
I. 3. A Le glucose.....	8
I. 3. B Transport du glucose.....	9
I. 3. C Catabolisme du glucose	10
I. 3. D Régulation du catabolisme du glucose	10
I. 4. Métabolisme du xylose	12
I. 4. A Le xylose.....	12
I. 4. B Transport du xylose.....	12
I. 4. C Métabolisme du xylose	13
I. 4. D Régulation du métabolisme du xylose.....	14
I. 5. Métabolisme de l'acétate.....	15
I. 5. A L'acétate.....	15
I. 5. B Transport de l'acétate.....	16
I. 5. C Catabolisme de l'acétate	16
I. 5. D Régulation du catabolisme de l'acétate	16
I. 6. Métabolisme de l'arabinose et du lactose	17
I. 6. A Transport de l'arabinose.....	17
I. 6. B Catabolisme de l'arabinose	17
I. 6. C Régulation du catabolisme de l'arabinose	18
I. 6. D Transport et catabolisme du lactose	18
I. 6. E Régulation du métabolisme du Lactose	19
II. Réponse adaptative chez <i>Escherichia coli</i>	20
II. 1. Adaptation à l'apparition d'un substrat chez <i>Escherichia coli</i>	23

II. 2. Adaptation à la disparition d'un substrat et entrée en état stationnaire	25
II. 2. A <i>Levée de la répression catabolique</i>	25
II. 2. B <i>Mise en place de la réponse stringente</i>	25
II. 2. C <i>La réponse au stress contrôlé par le facteur σ^S</i>	27
II. 3. Adaptation à un changement de substrat	27
II. 3. A <i>La croissance diauxique</i>	27
II. 3. B <i>Motifs de régulation dans la consommation des sucres</i>	29
III. Hétérogénéité phénotypique au sein des populations	32
III. 1. Définition	32
III. 2. Présence de microenvironnements au sein des cultures	33
III. 3. Généralité sur l'hétérogénéité phénotypique non environnementale.....	34
III. 4. La persistance bactérienne	36
III. 4. A <i>Définition</i>	36
III. 4. B <i>Mécanismes</i>	38
III. 5. Hétérogénéité métabolique liée à l'utilisation de source de carbone	40
III. 5. A <i>Définition de l'hétérogénéité métabolique</i>	40
III. 5. B <i>Le cas de l'hétérogénéité métabolique en croissance diauxique</i>	41
III. 5. C <i>Mécanisme d'apparition de l'hétérogénéité métabolique</i>	43
III. 6. Le rôle de l'hétérogénéité phénotypique	47
IV. Objectifs de la thèse :	49
Chapitre 1 :	53
Mise en place de la méthodologie pour l'étude des sous populations.....	53
I. Choix expérimentaux pour la construction des marqueurs de l'hétérogénéité.....	55
I. 1. Choix expérimentaux.....	55
I. 1. A <i>Le modèle glucose/xylose</i>	55
I. 1. B <i>Zones promotrices</i>	56
I. 1. C <i>Rapporteur de l'expression: fluorescence</i>	59
I. 1. D <i>Construction plasmidique</i>	60
I. 1. E <i>Souches</i>	61
I. 2. Méthodologie de construction des premiers plasmides	62
I. 3. Test d'expression et de spécificité des promoteurs.....	64
I. 3. A <i>Test d'expression</i>	64

I. 3. B	Test de spécificité sur xylose	65
I. 4. C	Test de spécificité sur glucose	67
I. 4. D	Test de spécificité sur acétate	68
I. 5.	Test d'impact des marqueurs fluorescents lors de la croissance.....	70
II.	Création du plasmide d'étude portant une double fluorescence.....	73
Chapitre 2 :		81
La disponibilité du facteur de transcription XylR influe sur l'hétérogénéité phénotypique et le temps de transition d' <i>Escherichia coli</i> lors de l'adaptation du glucose au xylose		81
Introduction.....		87
Results.....		89
Discussion.....		109
Materials and Methods		113
Chapitre 3 :		121
Optimisation de la construction fluorescente		121
I.	Le plasmide pSB1C3 avec inversion du sens de transcription	124
I.1.	Design et Construction du plasmide.....	124
I. 2.	Test de fluorescence	125
II.	Intégration de la construction dans le génome de la souche BW25113	127
III.	Utilisation d'un plasmide à moyen nombre de copies	128
III.1.	Design et Construction du plasmide	128
II. 2.	Test de fluorescence.....	128
II. 3.	Test du nombre de copies	131
II. 4.	Test de l'impact sur la croissance	132
IV.	Caractérisation du comportement de la souche pBR322 avec inversion du sens de transcription lors de croissances en mélange glucose xylose.....	133
IV. 1.	Caractéristiques de croissance de la souche avec le plasmide pBR322	133
IV. 2.	Hétérogénéité de fluorescence de la souche avec le plasmide pBR322	136
V.	Conclusion.....	139

Chapitre 4:.....	141
Etude du phénotype des sous-populations observées lors de la transition glucose/xylose	141
I. Reprise de croissance après tri cellulaire	144
II. Analyse transcriptomique lors de culture en mélange glucose/xylose	150
II. 1. Choix des échantillons :.....	152
II.2. Etude de l'expression génique lors de la transition glucose/xylose à l'échelle de la population	154
II. 2. A Description générale des résultats d'expression	154
II. 2. B Analyse fonctionnelle des gènes différentiellement exprimés	157
II. 2. C Analyse de gènes ciblés	162
II. 3. Déconvolution des données : Etude du profil d'expression génique au niveau des sous- populations.	164
III. Conclusion	166
Discussion et Perspectives.....	169
I. Conclusions	171
II. Critique de la Méthodologie	173
II.1. Limites de notre méthodologie.....	173
II. 2. Perspectives d'amélioration.....	175
III. Le rôle de XylR et des facteurs impactant la transition	176
III.1. Le rôle de XylR.....	176
III. 2. Les autres facteurs impactants la transition.....	178
IV. Comportement de la souche BW25113 avec ou sans construction.....	180
IV. 1. La souche portant le plasmide pSB1C3.....	180
IV. 2. La souche portant le plasmide pBR322	181
IV. 3. Chez la souche sauvage.....	183
V. Conclusion finale.....	184
Matériels et méthodes	187

I.	Construction des souches.....	189
I. 1.	Clonage en ligation/digestion	189
I. 2.	Clonage Gibson.....	189
I. 3.	Intégration génomique	190
II.	Milieux de cultures.....	190
III.	Culture en bioréacteurs	190
IV.	Calcul des μ_{max} et des temps de latence	191
V.	Analyse de la cytométrie de flux.....	191
VI.	Tri cellulaire pour l'expérience de reprise de croissance.....	192
VII.	Analyse transcriptionnelle.....	192
VII. 1.	Prélevement et préparation des échantillons.....	192
VII. 2.	RNAseq.....	193
	Références.....	195

Etude bibliographique

Rapport-Gratuit.com

La bactérie *Escherichia coli* est un modèle d'étude en biologie des procaryotes, cette thèse porte sur le comportement individuel des cellules lors de croissance en mélange glucose/xylose. La première partie de cette thèse est une étude bibliographique qui présente les informations nécessaires à la contextualisation des travaux. Dans un premier temps nous allons brièvement décrire la bactérie *Escherichia coli* et son métabolisme carboné central. Puis nous nous intéresserons aux mécanismes d'adaption de ce métabolisme carboné central face à l'apparition, la disparition ou le changement de substrat. Puis dans un troisième temps nous nous intéresserons à la variabilité phénotypique des individus au sein de la population.

I. Escherichia coli : métabolisme des sucres

I. 1. Escherichia coli : description générale

La bactérie *Escherichia coli* est un bacille à gram négatif (Figure 1 A et B) de la famille des *Enterobacteriaceae*. Cette bactérie est naturellement présente dans la flore intestinale d'un grand nombre d'espèces animales dont l'homme. Elle a été isolée en 1885 par le docteur Theodore Escherich. La majorité des souches sont commensales du microbiote intestinal, c'est-à-dire qu'elles colonisent le microbiote intestinal sans nuire à l'hôte. Cependant certaines souches peuvent être responsables d'infections intestinales, urinaires, pulmonaires ou du système nerveux. Ces souches sont ainsi classées en deux catégories selon qu'elles soient capables de provoquer des pathologies intra-intestinales (InPEC) ou extra-intestinales (ExPEC).

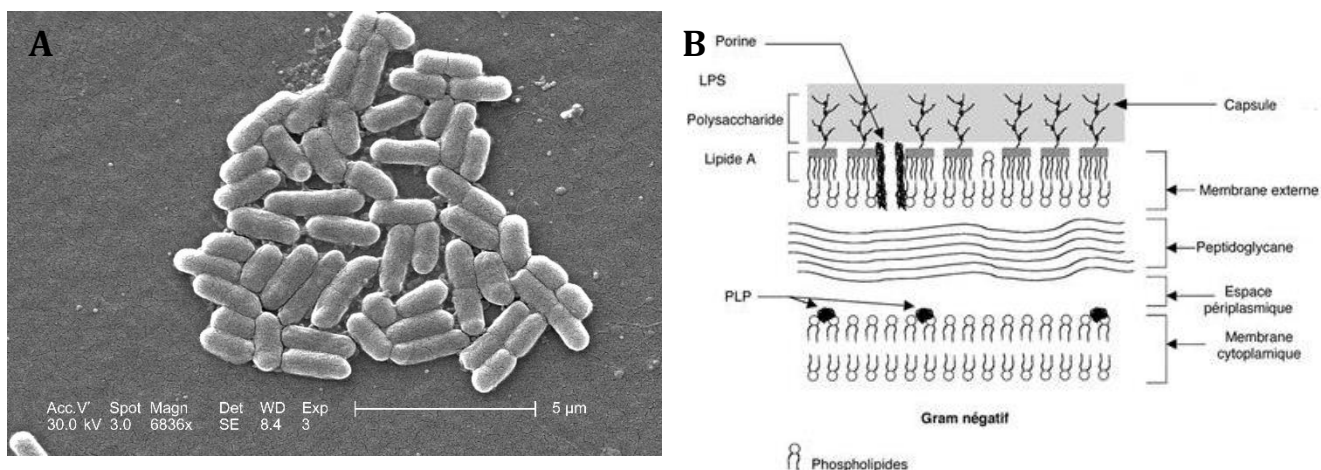


Figure 1: A) *Escherichia coli* observé en microscopie électronique à balayage (source : CDC/ National Escherichia, Shigella, Vibrio Reference Unit at CDC) B) Schéma de la paroi d'une bactérie à gram négative

La classification des souches pathogène est basée sur la présence de trois antigènes de surface : O (antigène somatique), K (antigène capsulaire) et H (antigène flagellaire) (Evans and Evans, 1983). Dans la majorité des cas les souches pathogènes sont désignées sous leur stéréotype O:H (ex : O104 :H4).

E. coli est aujourd'hui un des organismes les mieux caractérisés car elle présente les avantages d'avoir une croissance rapide et les souches modèles de laboratoire ne sont pas pathogènes. De plus son génome a été entièrement séquencé (Blattner *et al.*, 1997). L'évolution des techniques de séquençage a permis de réaliser une analyse phylogénétique basée sur différentes souches d'*Escherichia coli* isolées chez l'Homme. Cette analyse a mis en évidence quatre groupes phylogénétiques distincts (A, B1, B2 et D1) ainsi que deux sous-groupes (E et D2) (Chaudhuri and Henderson, 2012).

Escherichia coli est une bactérie anaérobie facultative. Elle est retrouvée dans l'intestin des mammifères, plus particulièrement dans la couche de mucus qui tapisse les cellules épithéliales du colon (Møller *et al.*, 2003). Son métabolisme est adapté à la consommation de sucres issus de la dégradation du mucus (mannose, galactose, fucose, ribose, L-arabinose, galacturonate, gluconate, N-acetylglucosamine...) (Chang *et al.*, 2004) mais aussi de sucres plus rarement présents dans le colon tel que le glucose, le xylose et l'acétate.

E. coli est très utilisée en châssis de production industrielle, notamment dans le domaine pharmaceutique où environ 30% des protéines recombinantes commercialisées sont produites chez *E. coli* (Huang *et al.*, 2012). C'est aussi un des organismes les plus utilisés en ingénierie métabolique et en biologie de synthèse (Pontrelli *et al.*, 2018). Cette bactérie reste un modèle d'étude en laboratoire, notamment la souche K12, souche historique de laboratoire dont le métabolisme a été intensivement étudié. Les apports nutritionnels nécessaires à la croissance de la souche K12 et de ses dérivées sont une source d'azote, de phosphore, de soufre, de carbone et des éléments minéraux (Na, Fe, Mg, K, Ca). Elle est hétérotrophe pour le carbone, c'est-à-dire qu'elle a besoin d'une source de carbone organique pour sa croissance. Son métabolisme lui permet d'utiliser un grand nombre de substrats différents comme apport de carbone (oses, acides aminés, acides faibles...) (Yoon *et al.*, 2012).

I. 2. Métabolisme Carboné Central d'*Escherichia coli*

Le métabolisme carboné central est l'ensemble des réactions biochimiques, intracellulaires permettant la dégradation des sources de carbone (catabolisme) et la synthèse des constituants de la cellule (anabolisme). Il permet de fournir l'apport énergétique majoritaire de la cellule ainsi que les précurseurs biosynthétiques indispensables (notamment à la synthèse d'acides nucléiques, d'acides aminés, de lipides...). Cette partie sera centrée sur la description du métabolisme des substrats impliqués dans la croissance sur mélange glucose/xylose.

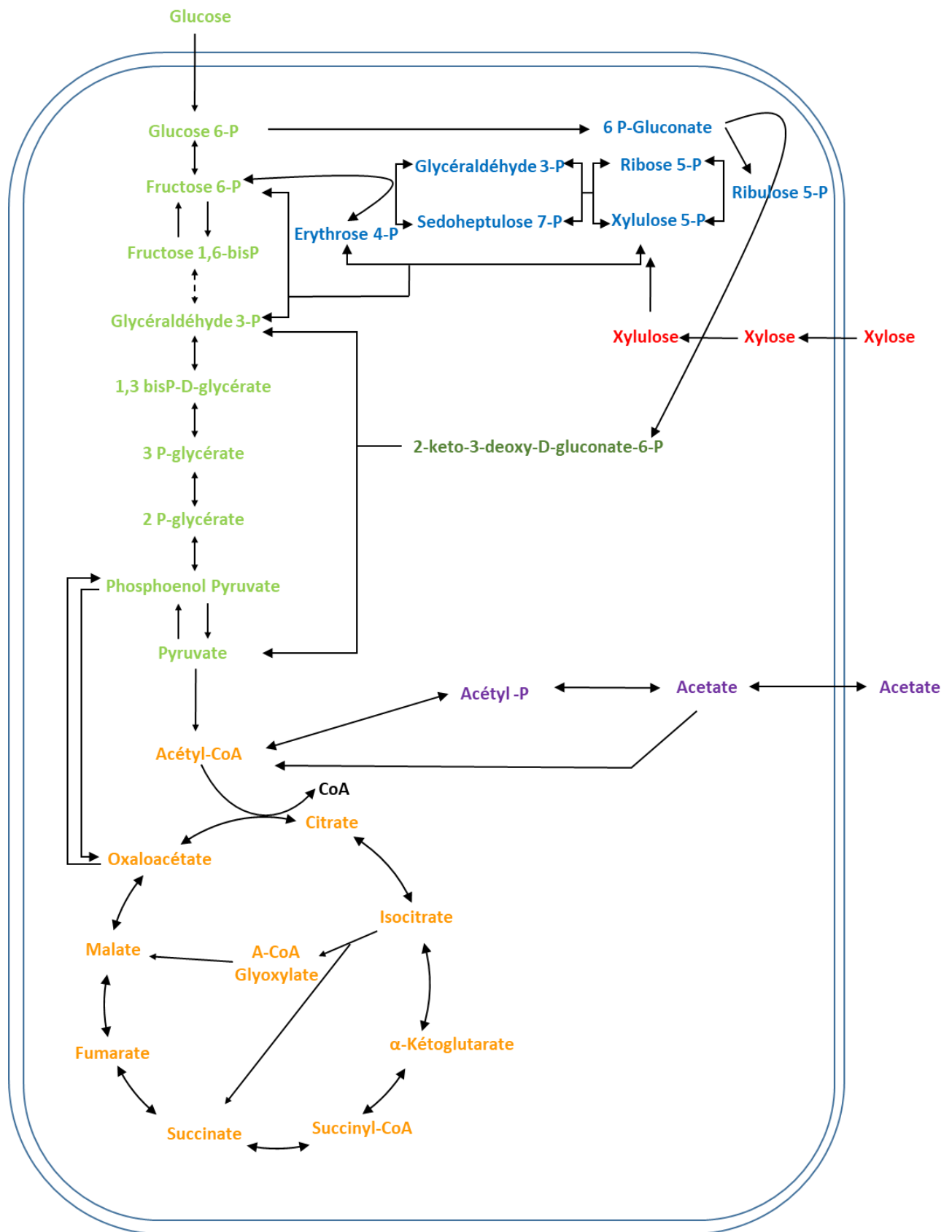


Figure 2 : Métabolisme carboné central d'*Escherichia coli*. La glycolyse est représentée en vert clair, le cycle de Krebs en jaune, la voie d'Entner-Doudouroff en vert foncé, La voie des pentoses phosphates en bleu et le métabolisme de l'acétate en violet.

I.2.A Glycolyse (Embden–Meyerhof pathway) et gluconéogenèse

La glycolyse est la voie de dégradation du glucose et de manière plus générale des hexoses. Elle permet la dégradation d'une molécule de glucose en deux molécules de pyruvate (Figure 2). Cette voie génère de l'ATP et du NADH. Elle est réversible, son fonctionnement inversé s'appelle la gluconéogénèse et permet de synthétiser différents précurseurs biosynthétiques à partir de substrats plus pauvres en carbone qui rentrent dans le métabolisme au niveau du cycle de krebs.

I.2.B Cycle de Kreb (TCA)

Le cycle de Krebs, aussi appelé cycle de l'acide citrique, est la source majoritaire de production d'énergie et de pouvoir réducteur via la réduction des co-enzymes impliquées dans la chaîne respiratoire. Il permet de former du citrate à partir d'une molécule d'acétyl-CoA liée à une molécule d'oxaloacétate. Le cycle complet est composé de 8 étapes successives, dont 2 de décarboxylation dégageant du CO₂, et 6 étapes d'oxydoréduction (générant NADH et FADH₂) (Figure 2). Ainsi un tour complet du cycle de Krebs permet de générer 3 molécules de NADH, 1 de FADH₂, 1 de GTP et dégage 2 molécules de CO₂. Les molécules de NADH et FADH₂ vont ensuite être utilisées dans la chaîne respiratoire pour générer un total de 12 molécules d'ATP. Quand l'acétate est la source de carbone, le cycle de Krebs n'est pas effectué en entier, le cycle du glyoxylate est activé. Ce cycle permet aux micro-organismes d'utiliser des substrats pauvres en carbone en s'affranchissant des étapes de décarboxylation du cycle de Krebs (Figure 2).

I.2.C Voie des pentoses phosphates

La voie des pentoses phosphates est une des voies principales du métabolisme carboné central. Cette voie est divisée en deux phases, une première partie appelée branche oxydative, et une seconde nommée voie non oxydative. La branche oxydative permet à la cellule de générer du NADPH via la transformation du glucose-6-phosphate en ribulose-5-phosphate. La seconde phase permet la synthèse de précurseurs métaboliques comme le ribose-5-phosphate (précurseur dans la synthèse des acides nucléiques) et l'érythrose-4-phosphate (un des précurseurs de la voie de synthèse des

acides aminés aromatiques). De plus cette voie permet un retour dans la glycolyse au niveau du fructose-6-P et glycéraldéhyde-3-P.

I.2.D Voie d'Entner-Doudoroff

La voie d'Entner-Doudoroff est la voie alternative à la glycolyse présente chez les procaryotes. Elle permet de dégrader le glucose en pyruvate et est composée de deux enzymes : une 6-phosphogluconate déshydratase (EDD), qui dégrade le D-gluconate 6-phosphate en 2-keto-3-déoxy-D-gluconate-6-phosphate, et une 2-keto-3-déoxy-D-gluconate-6-phosphate aldolase (EDA) (Figure 2). Cette dernière permet de dégrader l'intermédiaire précédant en glycéraldéhyde 3-phosphate et en pyruvate (Wolfe, 2015). Cette voie est notamment activée lors de croissance sur gluconate (Eisenberg and Dobrogosz, 1967).

I.3. Métabolisme du glucose

I.3.A Le glucose

Le glucose est une molécule de la famille des hexoses, il s'agit d'un monomère de glucide contenant 6 carbones, sa formule chimique est $C_6H_{12}O_6$ (Fig 3).

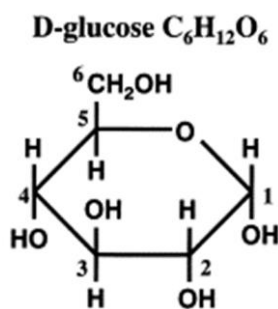


Figure 3 : Schéma d'une molécule de D-glucose

Dans la nature le glucose est le plus souvent retrouvé sous forme de polymère, soit en dimères avec lui-même ou d'autre monomère d'oses (maltose, lactose, saccharose), soit en polymères plus complexes souvent retrouvés chez les végétaux (amidon, cellulose, lignocellulose...).

I.3.B Transport du glucose

Chez les bactéries à gram négative, comme *E. coli*, le glucose diffuse à travers la membrane externe via un grand nombre de porines (OmpC, OmpF et LamB) (Death *et al.*, 1993) (Gosset, 2005). Une fois dans le périplasme, la première étape du métabolisme du glucose est l'internalisation de la molécule dans le cytoplasme. Le passage de la membrane interne nécessite un transport actif de la molécule. L'assimilation principale du glucose se fait via le système phosphoenolpyruvate phosphotransférase glucose (PTS glucose). Ce système de transport a la plus grande affinité pour le glucose (Ferenci, 1996).

Les systèmes PTS sont des systèmes de phosphorelais impliquant plusieurs protéines. Ils permettent de coupler l'internalisation et la phosphorylation des carbohydrates. Dans le système PTS glucose, le phosphoenolpyruvate (PEP) agit comme un donneur de phosphate pour une cascade de phosphorylation comprenant les protéines EI, HPr et EII. L'enzyme EI permet le transfert d'un groupement phosphate du PEP au résidu His15 de la protéine HPr, qui à son tour va transmettre le groupement phosphate au résidu histidine de la sous-unité A du complexe enzymatique EII (EIIGlc) (Figure 4). Le groupement phosphate est par la suite transmis à la sous unité EIIB_{glc}. Les sous unités EIIC et EIIB forment un complexe membranaire qui permet au glucose de franchir la membrane interne à travers l'EIIC puis d'être phosphorylé par l'EIIB pour rentrer dans la cellule sous forme de glucose-6-phosphate (Stülke and Hillen, 1999; Deutscher *et al.*, 2006; Görke and Stülke, 2008). La partie EIIBC_{glc}, du complexe enzymatique EII, est codée par le gène *ptsG* et la partie EIIGlc est codée par le gène *crr*.

D'autres voies d'entrée du glucose à travers la membrane cytoplasmique existent chez *E. coli*. En effet, le glucose peut être pris en charge par le système PTS du mannose (Curtis and Epstein, 1975) ainsi que par les différents systèmes de transport du galactose (Fig 4). Le glucose peut ainsi être internalisé via le transporteur *GalP* du galactose (McDonald *et al.*, 1997; Hernández-Montalvo *et al.*, 2003) ainsi que par l'ABC transporteur du galactose, codé par les gènes *mglA*, *mglB* et *mglC* (Zukin *et al.*, 1977; Death and Ferenci, 1993). Dans le cas de transport du glucose via les transporteurs du galactose, le glucose est alors phosphorylé par la glucokinase (Glk).

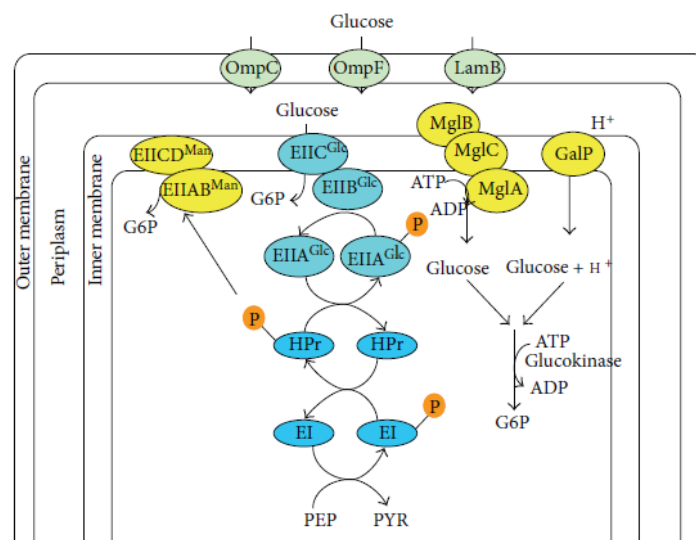


Figure 4: Schéma des différentes voies d'entrées du glucose dans la cellule (Shimizu, 2013)

I.3.C Catabolisme du glucose

Une fois entré dans la cellule sous forme de glucose-6-Phosphate, le glucose va être dégradé en pyruvate. Le glucose-6-Phosphate se distribue entre plusieurs voies métaboliques, la grande majorité étant prise en charge par la glycolyse et le reste par la voie des pentoses phosphates (Figure 2) (Leighty and Antoniewicz, 2013; Crown *et al.*, 2015), approximativement 88% du flux entrant de glucose part dans la glycolyse et 11% part dans la voie des pentoses phosphates (Hollinshead *et al.*, 2016). La voie de la glycolyse permet de générer 2 molécules de pyruvates, 2 molécules d'ATP et 2 molécules de NADH par molécule de glucose consommé. La voie des pentoses phosphates permet de générer du NADPH.

I.3.D Régulation du catabolisme du glucose

Chez les bactéries, il existe une hiérarchisation de l'utilisation des substrats. Chez *Escherichia coli* la source de carbone préférentiellement consommée est le glucose (Monod, 1949). Ainsi la consommation du glucose entraîne une répression de l'utilisation des autres substrats. Ce phénomène est appelé la répression catabolique (Deutscher, 2008).

Chez la bactérie *Escherichia coli*, la répression catabolique est médiée par la protéine EIIGlc. En absence de glucose, la forme phosphorylée de cette protéine vient activer l'adénylate cyclase, qui produit alors de l'AMP cyclique. L'augmentation de la

concentration d'AMPc intracellulaire favorise la formation du complexe régulateur de la transcription CRP-AMPc (Figure 5). Ce complexe est le principal activateur des voies alternatives du métabolisme carboné central qu'il régule en se fixant aux promoteurs de nombreux gènes et opérons du métabolisme. En présence de glucose, EIIA_{glc} phosphorylé transmet le phosphate au glucose lors de son internalisation dans la cellule formant du Glucose-6-P et empêchant ainsi la formation du complexe CRP-AMPc. De plus EIIA_{glc} non phosphorylé interagit directement avec plusieurs transporteurs des sucres alternatifs au glucose : LacY pour le lactose, MalK pour le maltose, MelB pour le melibiose ainsi qu'avec la glycérol Kinase GlpK et inhibe leur activité (Deutscher, 2008) (Figure 5). Ce phénomène de répression catabolique est appelé exclusion de l'inducteur, car dans de nombreux cas le transporteur sert aussi d'inducteur de la voie métabolique.

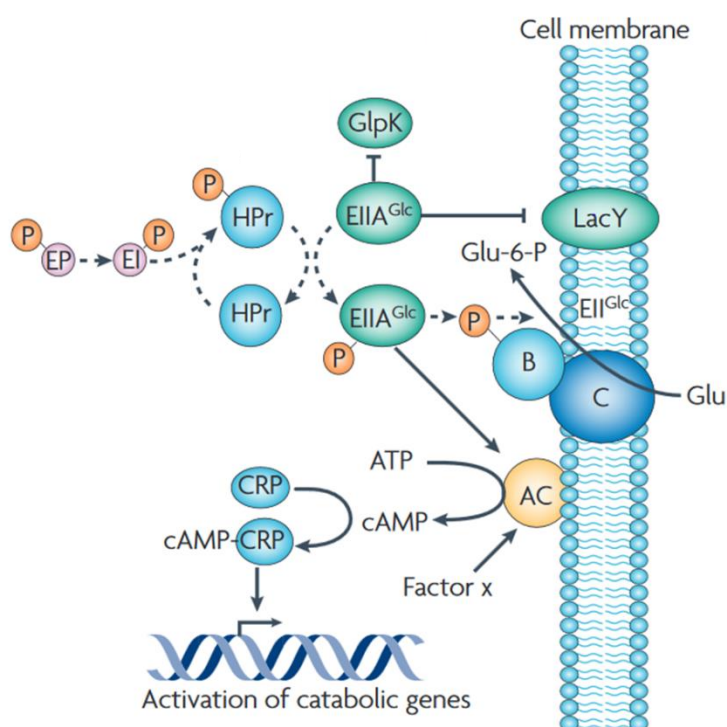


Figure 5 : Schéma de la répression catabolique (Görke and Stülke, 2008). L'internalisation du glucose sous forme de glucose-6-P entraîne la déphosphorylation des composants du système PTS. Ainsi la formation d'AMPc n'est pas stimulée et les transporteurs d'autres substrats sont inhibés (par exemple, GlpK et LacY).

Il existe d'autres régulateurs du métabolisme du glucose, notamment Mlc codée par le gène *dgsA*. Mlc est un facteur de transcription qui réprime l'expression de *ptsH*, *ptsG*, *manXYZ*, *malT* et lui-même (Deutscher, 2008). Lors de croissance sur glucose, Mlc est capturé par EIIB_{glc}.

De plus, d'autres régulateurs post-transcriptionnels sont impliqués dans le contrôle du métabolisme du glucose tel que Cra (Chavarría and Lorenzo, 2018) et Csr (Morin *et al.*, 2016).

I. 4. Métabolisme du xylose

I. 4. A Le xylose

Le xylose est une molécule faisant partie de la famille des pentoses, soit un monomère de glucide composé d'une chaîne de 5 carbones (Figure 6). Ce sucre peut être utilisé comme unique source de carbone par la bactérie *Escherichia coli* (Eiteman *et al.*, 2008). Il est présent dans la paroi des végétaux et est un des composants de la lignocellulose.

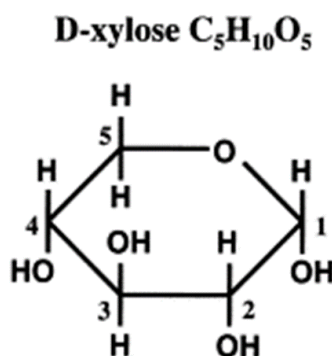


Figure 6 : Schéma d'une molécule de D-xylose

I. 4. B Transport du xylose

Deux systèmes de transport actif du D-xylose vers l'intérieur de la cellule ont été identifiés. Comme pour le galactose, le premier est un système de transport de type ABC (ATP Binding Cassette) codé par les gènes *xylF*, *xylG* et *xylH*. Les transporteurs ABC sont des systèmes de transports actifs de molécules nécessitant un apport d'énergie fourni par

l'hydrolyse de l'ATP. Ils permettent l'import d'une seule molécule spécifique dans la cellule (uniport). Les transporteurs de type ABC sont majoritairement constitués de cinq domaines : deux domaines transmembranaires permettant la formation d'un canal pour le passage de la molécule, deux domaines cytoplasmiques qui assurent la fixation et l'hydrolyse de l'ATP (Nucleotide-Binding Domains), et une partie périplasmique nécessaire à la reconnaissance et la fixation du substrat (Hollenstein *et al.*, 2007). Dans le cas du transport du xylose, le gène *xylF* code pour la protéine constituant la partie périplasmique du transporteur permettant la fixation du xylose. Ce site de liaison au substrat confère au système de transport une forte affinité pour le xylose (K_d 0.13 μM)(Sooriyaarachchi *et al.*, 2010) (K_m 0.2-4 μM) (Sumiya *et al.*, 1995). Le gène *xylG* code pour la partie cytoplasmique de fixation de l'ATP et le gène *xylH* code pour la sous-unité transmembranaire. Les gènes *xylF*, *xylG* et *xylH* sont organisés en opéron.

Le second système de transport actif du xylose est un co-transporteur de type symport, car deux molécules rentrent dans la cellule. Le principe est de coupler l'import d'une molécule se déplaçant selon son gradient de concentration (ions H^+) avec le transport de la seconde molécule (xylose). L'import des ions H^+ va permettre de fournir l'énergie nécessaire au transport des molécules de xylose. Ce transport est assuré par une protéine transmembranaire codée par le gène *xylE*. Ce système de transport a une affinité faible pour le xylose (K_m 63-169 μM)(Sumiya *et al.*, 1995).

I.4 C Métabolisme du xylose

Une fois importé dans la cellule, le xylose est dégradé via deux enzymes dont les gènes sont organisés en opéron (*xylA* et *xylB*) (Figure 7). Le gène *xylA* code pour la xylose isomérase. Cette enzyme catalyse la première étape de dégradation du D-xylose en xylulose (Wovcha *et al.*, 1983; Briggs *et al.*, 1984; Schellenberg *et al.*, 1984). La xylulokinase, codée par *xylB*, phosphoryle le D-xylulose en D-xylulose-5-phosphate (Lawlis *et al.*, 1984). Le xylulose-5-phosphate ainsi produit par la dégradation du xylose intervient dans la phase non oxydative de la voie des pentoses phosphates et rejoint la glycolyse au niveau du fructose-6-P et du glycéraldéhyde-3-P.

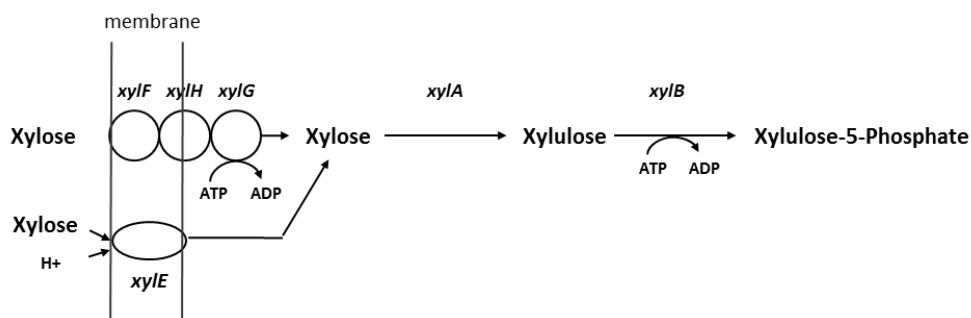


Figure 7 : Schéma de la voie de dégradation du xylose

I. 4. C Régulation du métabolisme du xylose

Le métabolisme du xylose, décrit précédemment, est majoritairement co-régulé par deux facteurs de transcription : le complexe CRP-AMPc et le régulateur XylR. En effet, le transport et la dégradation du xylose sont soumis à la répression catabolique, l'expression des opérons *xylAB* et *xylFGH* étant activée par la fixation du complexe CRP-AMPc (Figure 8). Cette fixation est possible à la disparition du glucose. Le second facteur de transcription régulant le métabolisme du xylose est XylR (Xylose régulator). Ce facteur de transcription est codé par le gène *xylR*, positionné à la suite du gène *xylH* sur le chromosome (Figure 8). Son expression est majoritairement sous le contrôle du promoteur P_{xylFGH} , même si *xylR* possède aussi son propre promoteur. Le complexe XylR-Xylose permet l'activation de l'expression des opérons *xylAB* et *xylFGH* en se fixant directement sur une séquence particulière identifiée dans les promoteurs. Cette même séquence a été retrouvée dans la zone promotrice du gène *xylE*. Il n'y a pour l'instant aucune preuve expérimentale d'une fixation de XylR au promoteur de *xylE* (Song and Park, 1997; Laikova *et al.*, 2001). Une étude structurale de XylR a montré que la molécule possède à la fois un domaine de liaison à l'ADN et un domaine de liaison au ligand sous forme de dimère. La liaison du D-xylose entraîne un réarrangement de ce dimère, libérant les domaines de liaison à l'ADN. XylR peut se lier aux séquences promotrices de *xylAB* et *xylFGH* (transcrits en sens opposés) (Ni *et al.*, 2013).

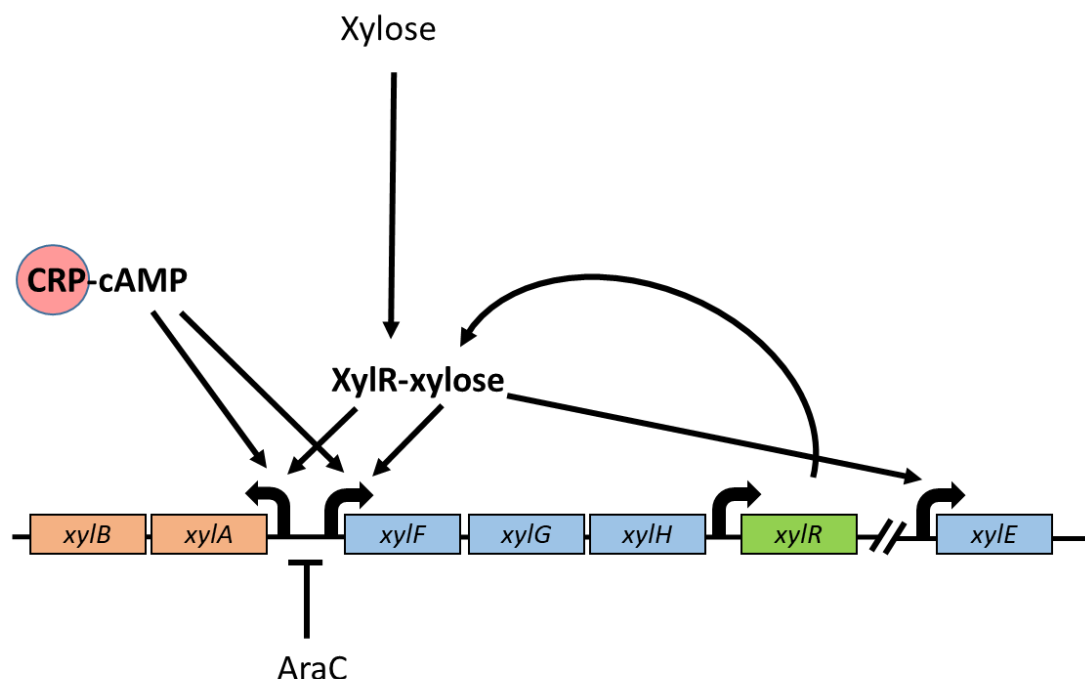


Figure 8 : Régulation du métabolisme du xylose. Les gènes codant pour le catabolisme sont représentés en orange, les gènes codant pour le transporteur en bleu et le gène codant pour le régulateur XylR en vert.

L'opéron xylose étant soumis à la répression catabolique et le glucose est consommé avant le xylose. De plus il a été prouvé que le métabolisme du xylose est réprimé en présence d'arabinose par la fixation du facteur de transcription AraC permettant la répression des opérons *xylAB* et *xylFGH* (Desai and Rao, 2010). Cela mène donc à une hiérarchisation de l'utilisation de l'arabinose (consommé en premier) et du xylose.

I. 5. Métabolisme de l'acétate

I. 5. A L'acétate

L'acétate est une molécule produite par la bactérie *Escherichia coli* en aérobie et à haut taux de croissance. Cette molécule peut aussi être consommée comme unique source de carbone par la bactérie. L'acétate est une molécule qui contient deux atomes de carbone et de formule ionique CH_3COO^- .

I.5.B Transport de l'acétate

L'acétate peut lui aussi être utilisé comme source unique de carbone par la bactérie *Escherichia coli*. L'assimilation de l'acétate peut s'effectuer via ActP, un transporteur de type symport couplant le transport actif de l'acétate avec le transport d'un ion H⁺ (Gimenez *et al.*, 2003). L'acétate est aussi capable de diffuser dans la cellule de manière passive.

I.5.C Catabolisme de l'acétate

Chez *E. coli*, il existe, deux voies métaboliques majeures qui permettent la consommation de l'acétate pour générer de l'acétyl-CoA : la voie *acs* et la voie *ack-pta* (Wolfe, 2005). La première voie est constituée de l'Acétyl-CoA synthétase (Acs) qui catalyse la transformation d'acétate en Acétyl-CoA (Figure 9). Elle nécessite la consommation d'ATP (Kumari *et al.*, 1995). Cette réaction est codée par le gène *acs* organisé en opéron avec le gène *actP* qui code pour le transporteur de l'acétate. La seconde voie métabolique permettant la consommation de l'acétate est constituée de l'acétate kinase A (AckA) qui catalyse la phosphorylation de l'acétyl phosphate, et d'une phosphate acétyl transférase (Pta) qui permet la transformation d'acétyl phosphate en Acétyl-CoA (Castaño-Cerezo *et al.*, 2009) (Figure 9). Cette voie est réversible, elle permet donc aussi la production d'acétate. Les gènes *ack* et *pta* sont organisés en opéron.

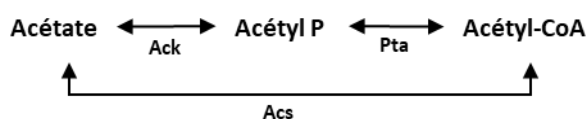


Figure 9 : Schéma des voies de catabolisme de l'acétate chez *E. coli*

I.5.D Régulation du catabolisme de l'acétate

La voie métabolique *acs* est sous contrôle de la répression catabolique et est activée par le complexe CRP-AMPc (Castaño-Cerezo *et al.*, 2011). L'opéron *ack-pta* est au contraire constitutivement exprimé (Shalel-Levanon *et al.*, 2005).

Lors de croissance aérobie en excès de glucose, *E. coli* produit et excrète de l'acétate. Ce phénomène est appelé « overflow metabolism ». En effet lors de croissance rapide sur glucose le flux de carbone entrant est plus important que la capacité du cycle de Krebs menant à l'accumulation d'Acétyl-CoA dans la cellule et entraînant un déséquilibre métabolique. Une partie de cet Acétyl-coA est alors transformé en l'acétate via la voie *ack-pta*, permettant la régénération d'ATP et de coenzyme A (Wolfe, 2005). Ce flux est bidirectionnel et dépend de la concentration extracellulaire en acétate (Enjalbert *et al.*, 2017). Ce métabolisme est aussi dépendant du taux de croissance. Ainsi, la production d'acétate n'apparaît qu'à partir d'un taux de croissance supérieur à 0.27 h^{-1} (Valgepea *et al.*, 2010). Une fois la totalité du glucose consommée et avant d'entrer en phase stationnaire, le flux net de production d'acétate s'inverse et l'acétate peut être re-consommé par la cellule (Wolfe, 2005; Enjalbert *et al.*, 2017).

I. 6. Métabolisme de l'arabinose et du lactose

I. 6. A Transport de l'arabinose

De la même manière que pour le xylose, il existe chez *E. coli* deux transporteurs de l'arabinose. Le premier est un transporteur de type ABC ayant une haute affinité pour l'arabinose et étant codé par l'opéron *araFGH*. AraF constitue la partie périplasmique permettant la liaison à l'arabinose, AraG est la sous-unité permettant la fixation de l'ATP, et AraH constitue la sous-unité membranaire. Le second système de transport de l'arabinose est un co-transporteur de type symport ayant une faible affinité pour l'arabinose, codé par le gène *araE* (Figure 10).

I. 6. B Catabolisme de l'arabinose

Une fois internalisé dans la cellule par un de ces transporteurs, l'arabinose est dégradé. La première étape de dégradation est catalysée par une l'arabinose isomérase codée par le gène *araA*, qui permet la transformation du l-arabinose en l-ribulose. Le l-ribulose est ensuite phosphorylé par la ribulokinase AraB pour former du l-ribulose-5-phosphate. La troisième étape de dégradation de l'arabinose est catalysée par l-ribulose-5-phosphate 4-epimerase, AraD. AraD convertit le ribulose-5-phosphate en D-xylulose-5-phosphate, qui est ensuite pris en charge par la voie des pentoses phosphates (Mayer and Boos, 2005).

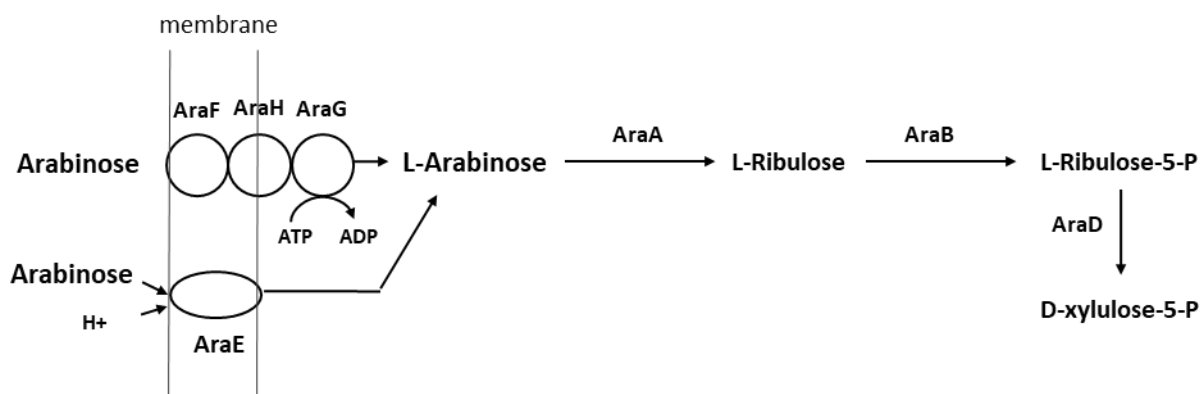


Figure 10 : Schéma du transport et du catabolisme de l'arabinose

I. 6. C Régulation du catabolisme de l'arabinose

Les gènes codant pour AraA, AraB et AraD sont organisés en l'opéron *araBAD*. L'ensemble des gènes responsable de la catalyse et du transport du L-arabinose sont sous le contrôle du facteur de transcription AraC. Les opérons *araBAD*, *araFGH* et *araE* sont activés par AraC en présence d'arabinose (Schleif, 2003). De plus en absence d'arabinose AraC est capable de réprimer l'expression de son propre promoteur ainsi que celle du promoteur de l'opéron *araBAD* (Lobell and Schleif, 1990). Le métabolisme de l'arabinose est soumis à la répression catabolique, c'est-à-dire que les quatre promoteurs (*ParaBAD*, *ParaFGH*, *ParaC* et *ParaE*) sont activés par le complexe CRP-AMPc.

I. 6. D Transport et catabolisme du lactose

Le transport du lactose au sein de cellule est effectué par la perméase LacY, il s'agit d'un transport de type symport. Une fois internalisé dans la cellule le lactose est dégradé. L'hydrolyse d'une molécule de lactose permet de donner une molécule de galactose et une molécule de glucose. Cette hydrolyse est catalysée par la β -galactosidase LacZ. Les gènes *lacZ*, *lacY* et *lacA* sont organisés en opéron. *lacA* code pour une galactoside acétyltransférase dont la fonction est pour l'instant encore incertaine.

I. 6. E Régulation du métabolisme du Lactose

La dégradation et le transport du lactose sont sous le contrôle de la répression catabolique ainsi que du facteur de répression de la transcription *LacI*. En absence de lactose dans le milieu, *LacI* inhibe l'expression de l'opéron lactose. En croissance sur lactose, *LacI* se complexe avec l'allolactose et devient non-opérationnel. La répression de l'opéron lactose est alors levée (Lewis, 2005). Le mode de fonctionnement de l'opéron lactose a été l'un des premiers découverts, il a servi de modèle dans la compréhension de la régulation du métabolisme.

II. Réponse adaptative chez *Escherichia coli*

La niche naturelle d'*E. coli* est le microbiote intestinal, plus précisément le colon des mammifères. Ainsi, la bactérie est soumise à de fréquents changements et carences de substrats. Le métabolisme carboné central, décrit précédemment, se doit d'être adaptable afin de lui permettre de tirer le plus de bénéfices possibles de chaque substrat. De ce fait, des mécanismes de réponses ont été mis en place par la cellule. L'ensemble des réactions enzymatiques qui compose le métabolisme n'est pas statique : il est soumis à une régulation contrôlée dépendant de l'environnement et de l'état cellulaire. Cette régulation intervient à différents niveaux. Elle peut être au niveau transcriptionnel et post-transcriptionnel jouant sur l'expression des gènes. Elle peut aussi être au niveau traductionnel et post-traductionnel, par le contrôle des activités enzymatiques faisant intervenir par exemple des senseurs de flux comme le fructose-1,6-biphosphate (FBP) qui impacte l'activité d'enzymes de la glycolyse (Kochanowski *et al.*, 2013), ou encore par le contrôle de la stabilité des protéines, impliquant des protéases ATP-dépendantes ainsi que des protéines chaperones.

Dans cette partie nous allons nous focaliser sur la régulation de l'expression des gènes, soit la régulation transcriptionnelle et post-transcriptionnelle. Il existe chez *E. coli* plusieurs régulateurs globaux permettant des réponses différentes en fonction de l'environnement de croissance et du stress subi par la cellule (Figure 11). Le métabolisme carboné central est régulé à différents niveaux, le premier niveau de régulation est le niveau transcriptionnel, soit la régulation de la transcription du gène en ARNm (Figure 11).

Chez les bactéries, la transcription est effectuée par l'ARN polymérase qui est formée d'un complexe de 5 sous-unités (deux sous-unités α , une β , une β' et une ω). Ce complexe permet l'élongation et la terminaison de la transcription. Pour l'initiation de la transcription, le recrutement d'un facteur σ est nécessaire (Burgess *et al.*, 1969; Paget and Helmann, 2003). Les facteurs σ permettent une régulation de l'expression génique par la reconnaissance de séquences spécifiques dans les régions promotrices et le recrutement de la polymérase sur ces séquences. Il existe, chez *E. coli*, 7 facteurs sigma contrôlant chacun et spécifiquement l'activation d'un set de gènes en réponse à un stimulus

environnemental (situation de stress, condition de carence ou d'infection par exemple) (Gourse *et al.*, 2006). En plus des facteurs sigma, il existe plusieurs centaines d'autres facteurs de transcription permettant la régulation fine de l'expression génique. Ces protéines se lient spécifiquement à l'ADN du promoteur et interagissent avec l'ARN-polymérase. Ils peuvent avoir la fonction d'activateur ou de répresseur du complexe de transcription en reconnaissant des séquences régulatrices dans la zone promotrice des gènes à transcrire (Browning and Busby, 2004). *E. coli* possède 285 facteurs de transcription, dont 58 pour lesquels la fonction n'a pas encore été caractérisée (Ishihama *et al.*, 2016). Ces facteurs peuvent être spécifiques comme le facteur de transcription XylR impliqué dans la régulation du métabolisme du xylose ou globaux, c'est-à-dire régulant l'expression de plus d'une centaine de gènes comme CRP entre autres (Figure 11).

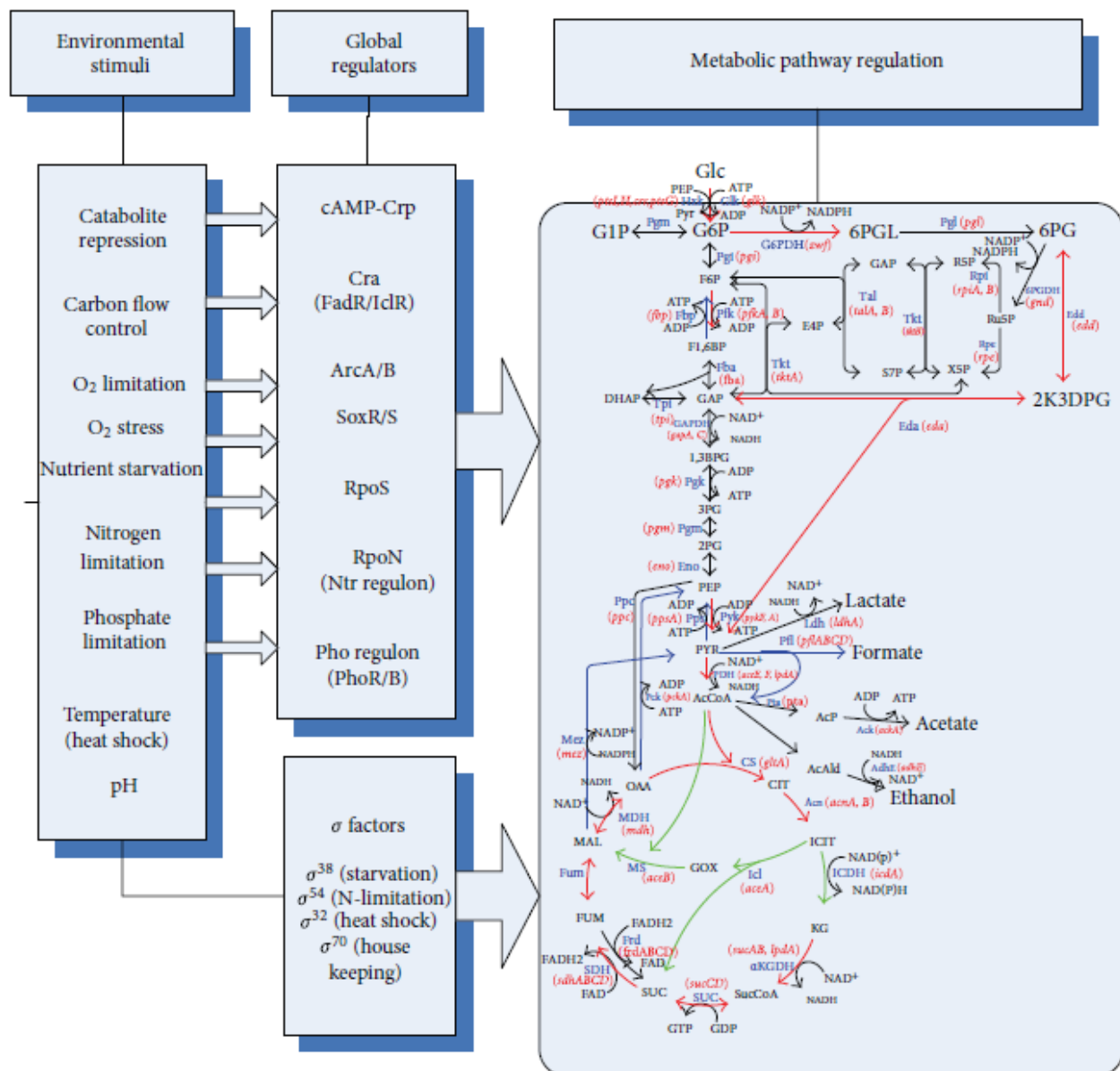


Figure 11 : Schéma des régulateurs de l'expression génique en réponse à un stimulus environnemental (Shimizu, 2013)

Conjointement aux facteurs de transcription, de petites molécules signal peuvent intervenir dans la régulation de l'expression génique. Par exemple, l'AMPc (adénosine monophosphate cyclique) se lie directement au facteur de transcription CRP pour former le complexe CRP-AMPc, acteur majeur de la répression catabolique. Ainsi à la disparition du glucose, la synthèse d'AMPc est stimulée menant à une accumulation du complexe CRP-AMPc qui vient activer l'expression des opérons de sucres alternatifs. On peut aussi citer

le ppGpp qui est le cœur de la réponse stringente (décrite en partie II. 2. B), cette molécule est accumulée en cas de carence en acides aminés.

Outre la régulation transcriptionnelle, il existe aussi des mécanismes de régulation du métabolisme au niveau post-transcriptionnel. Il existe chez *E. coli*, ainsi que chez de nombreuses autres espèces, des petits ARN non codants (ARNs) qui sont impliqués dans la régulation de la transcription, de la traduction mais aussi dans la dégradation des ARNm (Gottesman, 2004; Storz *et al.*, 2004; Negrete and Shiloach, 2017). D'autres protéines régulatrices interagissent avec les ARNm et peuvent impacter sur leur stabilité. Par exemple le régulateur post-transcriptionnel global CSR (Carbon Storage Regulator), dont la protéine cœur CsrA inhibe la traduction et/ou modifie la stabilité d'un grand nombre d'ARNm cible, notamment, les enzymes du métabolisme de la gluconéogenèse (Esquerré *et al.*, 2016).

La régulation du métabolisme est ainsi complexe impliquant différents niveaux de contrôle. Par la suite l'accent sera mis sur la réponse adaptative transcriptionnelle d'*Escherichia coli* à la disparition du glucose, que le glucose soit la seule source de carbone disponible, ou que la croissance s'effectue dans un environnement présentant plusieurs sources de carbone différentes.

II. 1. Adaptation à l'apparition d'un substrat chez *Escherichia coli*

Lors de l'apparition d'une nouvelle source de nutriment, après une carence plus ou moins prolongée, les courbes de croissance bactériennes présentent une phase de latence dite de « lag ». Cette phase de latence correspond à un temps d'ajustement de la population au nouvel environnement, avant d'initier la croissance exponentielle à partir des nouveaux substrats disponibles (Swinnen *et al.*, 2004). Un temps de latence similaire est aussi observé lorsque des cellules sont transférées artificiellement d'un environnement à un autre. Il existe plusieurs définitions de ce temps de latence, mais il est le plus communément vu comme le temps requis pour atteindre la phase exponentielle de croissance (Figure 12).

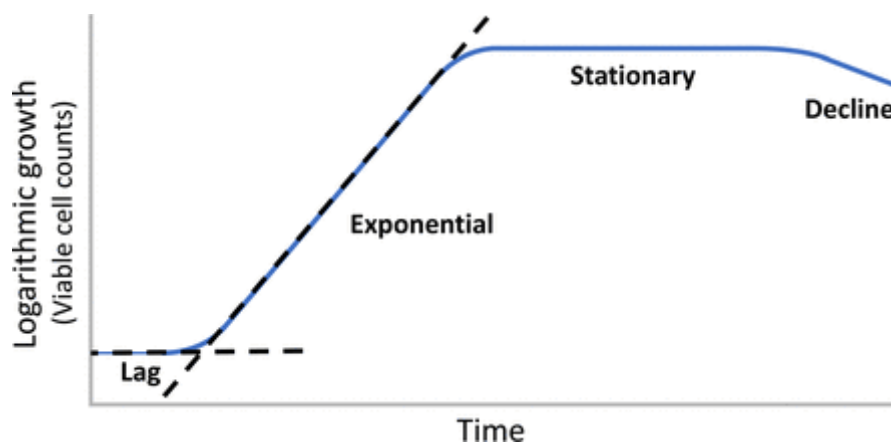


Figure 12 : Représentation graphique d'une culture bactérienne (Bertrand, 2019).

Durant ce temps de latence les cellules doivent produire les composés cellulaires nécessaires à l'accumulation de biomasse et à la division cellulaire. Une étude a montré que pour des cellules d'*E. coli* remises en croissance sur milieu riche (LB), les gènes de réponse aux stress acides et oxydatifs sont réprimés durant cette phase de latence, ainsi que les gènes de réponse à d'autres stress environnementaux médiés par le facteur σ^E (Pin *et al.*, 2009). A l'inverse sont induits des gènes impliqués dans le catabolisme, la génération d'énergie, le transport de molécules, les facteurs de transcription. Ainsi, l'introduction d'un nouveau substrat capable de soutenir la croissance induit une réponse globale permettant de mettre en route la machinerie nécessaire à la croissance. Une étude indique que l'expression de ces gènes n'est pas simultanée. Ainsi Madar *et al.*, (2013) divisent le temps de latence en deux temps distincts qu'ils appellent lag 1 et lag 2. Le lag 1 est le temps durant lequel il n'y a aucune production de biomasse et le lag 2 est le temps durant lequel la biomasse augmente sans qu'il y ait de division cellulaire. Ces auteurs ont notamment montré que, chez *E. coli* en reprise de croissance sur arabinose, seuls les promoteurs des gènes nécessaires à l'utilisation de l'arabinose sont activés lors du lag 1, alors que l'activation globale des promoteurs de tous les gènes nécessaires à la croissance et notamment la production de ribosomes ne se produit que lors du lag 2.

Enfin, pendant cette période de latence toute la machinerie de division est recrutée pour permettre la première génération. Il semblerait que la concentration intracellulaire de FtsZ, protéine responsable de la mise en place de l'anneau de division cellulaire, soit importante. Ainsi, cette protéine est accumulée durant la phase de latence jusqu'à une concentration seuil permettant de déclencher la première division (Bertrand, 2019).

II. 2. Adaptation à la disparition d'un substrat et entrée en état stationnaire

II. 2. A Levée de la répression catabolique

Si le substrat présent dans le milieu est le glucose, les cellules sont soumises à la répression catabolique, empêchant l'expression des gènes nécessaires à l'utilisation de sucres alternatifs (Deutscher, 2008). La disparition du glucose dans le milieu entraîne une accumulation du complexe CRP-AMPC, qui régule l'expression de plus de 100 gènes différents, notamment impliqués dans le transport et le métabolisme des sucres, dans la production d'énergie, dans le métabolisme des acides aminés, dans le métabolisme des nucléotides mais aussi dans les systèmes de transport d'ions (Zheng *et al.*, 2004).

II. 2. B Mise en place de la réponse stringente

La réponse stringente est une réponse induite en condition de stress chez la bactérie. Elle a été découverte dans le cas de carence en acides aminés (Cashel and Gallant, 1969). Cependant, cette réponse est beaucoup plus générale et intervient dans le cas de nombreux stress et de nombreuses carences (carbone, acides gras, phosphate, fer, choc thermique...) (Atkinson and Hauryliuk, 2012). Lors d'une croissance en milieu minimum où la seule source de carbone est le glucose, la réponse stringente est induite à la disparition de ce dernier et amorce l'entrée en phase stationnaire de croissance. Elle permet d'inhiber la transcription de la majorité des gènes induits lors de la phase exponentielle de croissance et d'induire les gènes de réponse au stress lors d'arrêt de croissance.

Les acteurs principaux de la réponse stringente sont la guanosine 5'triphosphate, 3'diphosphate (pppGpp) et la guanosine 5'diphosphate, 3'diphosphate (ppGpp), regroupés sous le nom de (p)ppGpp (Potrykus and Cashel, 2008). La synthèse de

(p)ppGpp s'effectue à partir d'ATP et de GTP ou GDP (donnant respectivement du pppGpp ou du ppGpp). Le niveau de (p)ppGpp chez *Escherichia coli* est contrôlé par les enzymes RelA et SpoT qui font parties de la famille RSH (RelA et SpoT Homologue). Cette famille est très conservée chez les procaryotes et les plantes (Atkinson *et al.*, 2011). La protéine RelA est une ppGpp synthetase (Haseltine *et al.*, 1972). La fixation d'un ARNt non chargé sur son site ribosomal active RelA, elle aussi liée au ribosome (Agirrezabala *et al.*, 2013). En effet, quand il n'y a pas de limitation en acides aminés, les ARN de transfert (dit ARNt) transportent les acides aminés jusqu'au ribosome sous la forme aminoacyl-ARNt. Lors de carence, la concentration cellulaire d'ARNt non chargé en acide aminé augmente. Ils peuvent alors entrer sur le site ribosomique, ce qui entraine l'activation de RelA et ainsi la synthèse de (p)ppGpp (Haseltine and Block, 1973). De plus, il a récemment été montré que la présence de ppGpp active RelA, entraînant ainsi une boucle de rétrocontrôle positive, permettant la mise en place rapide de la réponse stringente (Shyp *et al.*, 2012).

La concentration intracellulaire de (p)ppGpp est aussi contrôlée par la protéine SpoT, qui possède une double activité de synthétase et d'hydrolase. Elle est capable de synthétiser le (p)ppGpp, mais aussi de l'hydrolyser, générant du GTP ou du GDP et du pyrophosphate (Potrykus and Cashel, 2008). Dans le cas de carence nutritive, l'enzyme SpoT bascule vers son activité de synthétase plutôt que d'hydrolase. Cette balance est médiée par la molécule signal ACP (Acyl Carrier Protein) (Battesti and Bouveret, 2006).

Le (p)ppGpp a une action globale sur la régulation de l'expression des gènes (Traxler *et al.*, 2008 ; Durfee *et al.*, 2008). Chez *E. coli*, le (p)ppGpp se lie directement aux sous-unités β et β' de l'ARN polymérase (Touloukhonov *et al.*, 2001). Il inhibe la transcription des gènes requis pour la synthèse de l'ARNr et des protéines ribosomales, entraînant ainsi l'inhibition de la synthèse des protéines (Lemke *et al.*, 2011). En parallèle, il agit aussi comme un activateur de nombreux promoteurs de gènes impliqués dans la réponse au stress, de manière directe ou indirecte, via l'implication de facteurs sigma alternatifs au facteur σ^{70} (par exemple σ^S , σ^{32} et σ^{54}) (Magnusson *et al.*, 2005).

L'initiation de la réplication chromosomique est inhibée par une haute concentration en (p)ppGpp. Cependant, l'élongation de la synthèse d'ADN n'est pas affectée. Cela permettrait d'avoir un contrôle environnemental de la réplication (Ferullo and Lovett, 2008; Kraemer *et al.*, 2019). Il a aussi été montré plus récemment que la

réponse stringente a un rôle dans la virulence bactérienne, la résistance antibiotique, ainsi que dans la persistance (Hauryliuk *et al.*, 2015).

II. 2. C La réponse au stress contrôlé par le facteur σ^S

Le facteur σ^S (aussi appelé σ^{38}), permet de réguler l'expression d'une cinquantaine de gènes impliqués dans la réponse aux stress environnementaux et dans l'entrée en phase stationnaire de croissance (Zhou and Gottesman, 1998). Il est codé par le gène *rpoS*. DksA principalement et le (p)ppGpp dans une moindre mesure, induisent la transcription du gène *rpoS* (Brown *et al.*, 2002; Hirsch and Elliott, 2002). DksA est le second régulateur transcriptionnel impliqué dans la réponse au stress de type carence nutritive (Lemke *et al.*, 2011). Ce facteur de transcription de 17 kDa, se fixe directement à l'ARN polymérase. Les régulateurs DksA et ppGpp ont des effets similaires sur la cellule, le mutant *dksA* et le double mutant *relA spoT* partagent un grand nombre de phénotypes communs et la surexpression de DksA peut compenser la suppression de (p)ppGpp dans la cellule (Magnusson *et al.*, 2007).

II. 3. Adaptation à un changement de substrat : Le cas de la croissance en mélange de sucre

II. 3. A La croissance diauxique

Lors de croissance en mélange de sucre, les micro-organismes peuvent soit co-utiliser les deux substrats (Hermsen *et al.*, 2015; Wang *et al.*, 2019), soit les utiliser de manière séquentielle (Monod, 1949; Aidelberg *et al.*, 2014; Ammar *et al.*, 2018). Dans ce dernier cas, la croissance d'une population clonale sur une source de carbone multiple est décrite comme diauxique, c'est-à-dire constituée de deux périodes de croissances successives, entrecoupées d'une phase de latence (Figure 13) (Monod, 1949; Siegal, 2015). Chaque phase de croissance s'accompagne de la disparition d'un substrat. Les substrats sont ainsi consommés de manière successive et préférentielle, par la population bactérienne. Ces comportements ont pour la première fois été décrits par Jacques Monod en 1942 (Monod, 1942) sur des populations d'*Escherichia coli* en croissance sur un milieu minimum contenant du glucose et un autre carbohydrate (Figure 13).

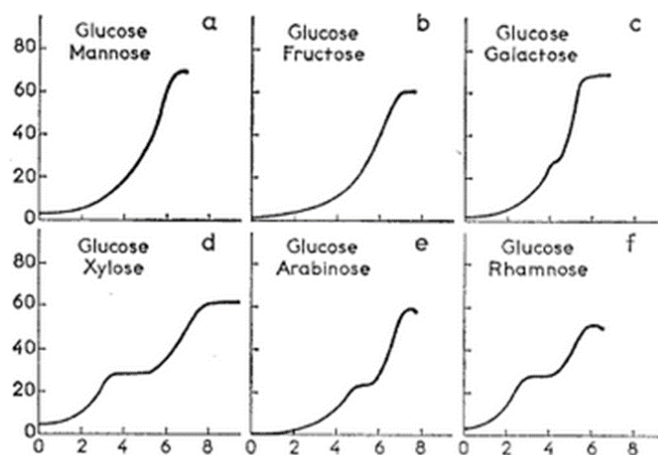


Fig. 1. Growth of *Escherichia coli* in the presence of different carbohydrate pairs serving as the only source of carbon in a synthetic medium (50).

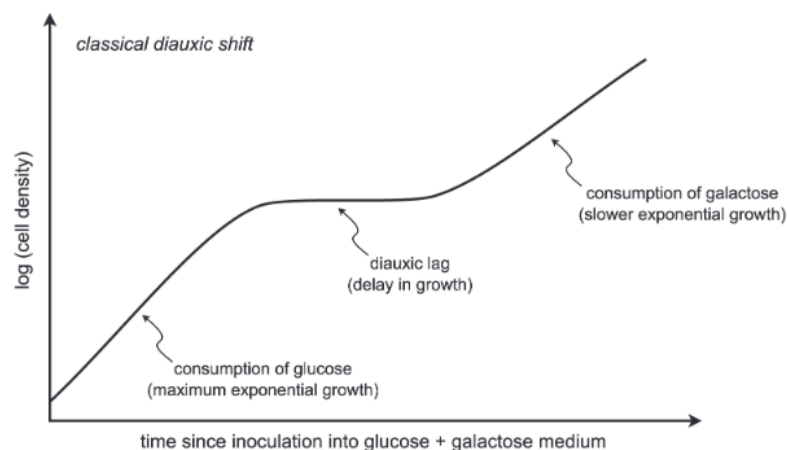


Figure 13 : La croissance diauxique sur des mélanges de substrats (d'après Monod, 1942 et Siegal, 2015)

Selon Jacques Monod, la phase de latence correspond au temps requis par la population pour réorganiser son métabolisme afin d'utiliser spécifiquement un autre substrat. Ainsi, la première phase de croissance exponentielle correspond à l'assimilation du premier substrat, puis la phase de latence (où le taux de croissance est nul) permettrait une transition métabolique homogène de l'ensemble de la population. La seconde phase de croissance serait liée à l'utilisation du second substrat. Le métabolisme carboné central est régulé pour faire du glucose la source de carbone préférentielle via la répression catabolique (voir I. 3. D) (Deutscher, 2008). Cette utilisation séquentielle a longtemps été regardée comme une capacité d'adaptation permettant de consommer le substrat le plus bénéfique en premier. En effet, il a été observé que plus la consommation d'un substrat

permet un taux de croissance élevé, plus ce dernier sera utilisé de manière préférentielle (Aidelberg *et al.*, 2014).

Des variations intracellulaires durant la phase de latence entre deux substrats lors de croissance en mélange, indiquent la levée de la répression catabolique, ainsi que la mise en place transitoire de la réponse stringente et de la réponse au stress, médiée par le facteur σ^S . Ainsi, il a pu être observé au niveau intracellulaire, lors de croissance sur un mélange glucose/lactose, que la concentration d'AMPc augmente rapidement à la disparition du glucose, puis diminue fortement quand le second substrat commence à être consommé (Inada *et al.*, 1996). Il en est de même pour l'expression du facteur σ^S (Fischer *et al.*, 1998). Une accumulation de (p)ppGpp a été observée à la disparition du glucose, puis une diminution à la reprise de croissance sur le second substrat lors de croissances en mélange de sucre glucose/succinate (Lazzarini *et al.*, 1971), ainsi que lors de croissances en mélange glucose/lactose avec en plus une augmentation de l'acétylation des protéines (Fernández-Coll and Cashel, 2018). De plus, lors de croissances en mélange glucose/acétate à la disparition du glucose, l'AMPc est fortement produit et les gènes de réponse aux stress ainsi que ceux d'utilisation des sucres de réserve sont induits (Enjalbert *et al.*, 2013).

II.3.B Motifs de régulation dans la consommation des sucres

La stratégie d'utilisation séquentielle des substrats provient d'une adaptation du métabolisme mettant en jeu la répression catabolique, médiée par le complexe CRP-AMPc, mais aussi d'autres facteurs de transcription régulant l'expression des opérons nécessaires à la consommation de substrats alternatifs au glucose (exemple XylR, AraC, LacI...). A la disparition du premier substrat, le métabolisme se réorganise pour s'adapter à la consommation du second substrat en mettant en jeu des boucles de régulation. La forme la plus présente de ce type de régulation est appelée « feedforward » ou boucle de rétrocontrôle, cette régulation pouvant être à la fois positive ou négative.

Dans ce type de régulation, un facteur de transcription général régule l'activation d'un second facteur de transcription spécifique et ces deux facteurs de transcription vont venir co-réguler l'expression d'un opéron en particulier (Shen-Orr *et al.*, 2002). C'est le cas de l'opéron arabinose (Figure 14), où le complexe CRP-AMPc va venir activer la transcription d'*araC*, puis AraC et CRP-AMPc vont conjointement activer la transcription des opérons responsables de la catalyse et du transport de l'arabinose.

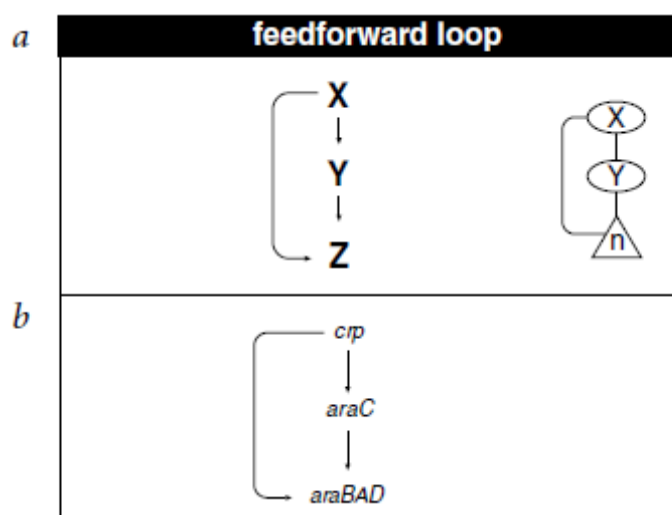


Figure 14 : Exemples de boucle de rétrocontrôle positive (Shen-Orr *et al.*, 2002)

L'exemple le plus étudié de mécanisme de régulation d'utilisation de substrat est le cas de l'opéron lactose. Il s'agit du premier opéron découvert (Jacob and Monod, 1961) et est constitué de *lacZ*, *lacY* et *lacA*. L'expression de cet opéron est sous le contrôle du répresseur LacI (Lewis, 2005). Le site de fixation de LacI est présent dans la région promotrice de l'opéron lactose. En présence de lactose, LacI se dissocie de son site de fixation, pour former un complexe avec l'allolactose permettant ainsi l'initiation de la transcription. Ces motifs de régulation permettent une réponse rapide, avec une augmentation rapide du niveau de transcription. L'opéron lactose présente une réponse d'induction graduelle en fonction de la concentration présente d'inducteur (Ozbudak *et al.*, 2004).

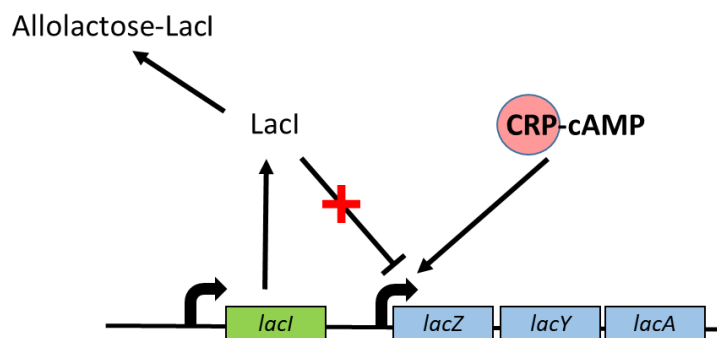


Figure 15 : Schéma de régulation de l'opéron lactose en présence de lactose

L'activation de l'opéron xylose semble lui aussi soumis à un schéma de « feedforward loop » positif comme décrit précédemment pour le lactose et l'arabinose. La présence de xylose induit l'activation de l'opéron *xylFGHR* via le complexe xylose-XylR. Ainsi, XylR induit sa propre expression, permettant une augmentation rapide du niveau de transcription des opérons. Cette boucle entraîne une régulation bistable du promoteur *P_{xylA}* où, à faible concentration de xylose, une partie des cellules est entièrement induite et l'autre partie entièrement non-induite (Afroz *et al.*, 2014). Ce type de motif pourrait mener à des niveaux d'expression différents entre les cellules quand l'observation est réalisée au niveau de la cellule unique.

Ainsi l'expression des voies métaboliques pourrait ne pas être homogène entre les cellules et la présence d'une seule molécule, de manière stochastique, pourrait mener à un changement de phénotype à l'échelle de la cellule (Choi *et al.*, 2008).

III. Hétérogénéité phénotypique au sein des populations

III. 1. Définition

Les phénomènes d'hétérogénéité au sein des cultures ont longtemps été regardées d'un point de vue génétique. Ainsi il a été montré que dans des environnements changeants, l'alternance entre des phases de croissance et de carence favorise l'apparition de mutation au sein de la population (Zambrano *et al.*, 1993; Finkel, 2006). Plus récemment, de nombreuses études qui seront abordées par la suite ont montré que les phénomènes d'hétérogénéité phénotypique n'étaient pas uniquement le fait de mutations, mais que des individus identiques génétiquement pouvaient présenter une variation de phénotype au sein de la même culture. L'hétérogénéité phénotypique est donc définie comme le fait que des cellules isogéniques, c'est-à-dire présentant le même patrimoine génétique, issus de la même cellule mère, n'ont pas un comportement phénotypique identique, homogène. Le phénotype est un caractère observable, il est défini en fonction de critères morphologiques, fonctionnels et métaboliques de la cellule.

L'hétérogénéité phénotypique est un aspect de la culture qui ne peut être observé qu'à l'échelle de la cellule sur un grand nombre d'individus (Ackermann, 2015). L'intérêt pour l'hétérogénéité phénotypique est grandissant sur les 15 dernières années. Jusqu'alors, les techniques analytiques mises en place étaient réalisées à l'échelle de la population entière, ne permettant pas d'observer ou de quantifier la variabilité entre les individus, résumant ainsi leur comportement à la moyenne de la population globale. L'émergence de nouvelles technologies basées sur l'analyse de cellules uniques, mais aussi l'augmentation de l'interdisciplinarité entre biologie, mathématique et physique, ont menées à un changement d'échelle d'analyse permettant l'essor des études sur l'hétérogénéité phénotypique (Sampaio and Dunlop, 2020). Ainsi nombre de ces études sont des études interdisciplinaires intégrant de la microscopie, de la microfluidique, de la modélisation de données et de la prédiction de comportement.

L'hétérogénéité phénotypique telle que définie aujourd'hui peut avoir plusieurs origines. Elle peut être due à la présence de microenvironnements au sein des cultures, menant des cellules génétiquement identiques à ne pas percevoir les mêmes stimuli environnementaux, ou bien provenir d'une variation de l'expression des gènes entre les cellules présentes dans un environnement homogène.

III. 2. Présence de microenvironnements au sein des cultures

Une des premières causes d'hétérogénéité phénotypique étudiée a été la présence de microenvironnements au sein des cultures. En effet les cultures liquides présentent un environnement qui n'est pas toujours complètement homogène. Il a été montré lors de cultures en réacteurs de gros volume (30 m³), que des gradients de glucose existaient entre le haut du réacteur (où est situé l'alimentation en glucose) et le bas du réacteur, la concentration en glucose pouvant être jusqu'à deux fois supérieure entre le haut du réacteur et le bas (Larsson *et al.*, 1996). Des gradients peuvent aussi être observés pour le pH et la dissolution de l'oxygène (Lara *et al.*, 2006). Cette différence dans la composition de l'environnement peut mener à ce que les cellules présentes dans un fermenteur ne soient pas exposées au même environnement (Delvigne and Goffin, 2014; Heins and Weuster-Botz, 2018). Cette différence environnementale peut conduire à une hétérogénéité phénotypique dans le métabolisme des cellules, avec une réduction des performances, en terme de rendement et de productivité à échelle industrielle (Delvigne *et al.*, 2014).

L'hétérogénéité environnementale n'est pas uniquement présente dans le cas de cultures liquides mais aussi dans le cas de cultures en milieux solides. En effet, au sein d'une même colonie, les cellules ne ressentent pas le même environnement direct. D'après la modélisation des colonies, la position de la cellule affecte son métabolisme. Lors de croissance sur milieu minimum contenant du glucose comme unique source de carbone, les cellules produisant de l'acétate se retrouveraient dans la partie intérieure de la colonie adjacente à la gélose, où du glucose est disponible mais peu d'oxygène est présent. Les cellules consommatrices d'acétate se retrouveraient au sommet du dôme de la colonie, où il n'y a pas de glucose mais l'oxygène n'est pas limité (Cole *et al.*, 2015). Ces observations sont soutenues par une étude des flux métaboliques qui prédit que 92% des cellules de la colonie consomment du glucose et secrètent de l'acétate et que les 8% situées sur le

dessus de la colonie, n'ayant pas accès au glucose, consomment cet acétate comme source de carbone, (Wolfsberg *et al.*, 2018).

Dans les conditions citées précédemment, la présence de microenvironnements induirait une hétérogénéité phénotypique qui permettrait la mise en place d'une stratégie de « cross-feeding », soit d'alimentation croisée. Cela consiste en un échange de métabolites entre les cellules. Ce type d'hétérogénéité permettrait une forme de dialogue entre les individus et serait bénéfique pour la population. D'après Evans *et al.*, (2020), le « cross-feeding » est bénéfique car il permettrait dans un premier temps d'augmenter l'efficacité de la population par l'allègement énergétique que permet la division de l'activation des voies métaboliques entre des cellules différentes. Il permet aussi la réutilisation d'un co-produit comme apport principal. De plus, il pourrait rendre compatible l'expression de voies métaboliques, une chez chaque individu, qui ne le serait pas dans une seule cellule.

III. 3. Généralité sur l'hétérogénéité phénotypique non environnementale

Les premiers phénomènes d'hétérogénéité phénotypique non environnementale ont été découverts au sein de populations de *Staphylococcus pyogenes* traitées à la pénicilline, où une fraction de la population avaient survécu à la présence d'antibiotique (Bigger, 1944). De la même manière des comportements non-homogènes ont été observés chez *Bacillus subtilis*, lors de l'activation de programmes génétiques spécifiques telles que la compétence et la sporulation (Nester and Stocker, 1963).

Les découvertes sur les 15 dernières années ont mené à se rendre compte que la variabilité intercellulaire lors de cultures de populations clonales n'est pas un phénomène ponctuel lié à certaines conditions très particulières, mais semble beaucoup plus général au sein du vivant.

Des études plus récentes ont montré que l'hétérogénéité phénotypique n'est pas un phénomène uniquement dû à la présence de microenvironnements au sein des cultures, mais que le phénomène peut aussi être observé dans des environnements considérés comme parfaitement homogènes, dans des cultures de petits volumes infiniment mélangées.

Il a ainsi été montré que l'expression génétique est soumise à un bruit, c'est-à-dire que toutes les cellules ne présentent pas exactement le même niveau d'expression pour un gène donné (Elowitz *et al.*, 2002). Dans le cas où l'expression d'un gène est soumise à des boucles de rétrocontrôles positives (voir II. 3. B), cette différence de niveau d'expression peut mener à une réponse binaire sur un caractère particulier, avec certaines cellules ayant un niveau d'expression très élevé pour un gène particulier et d'autres cellules avec un niveau d'expression très faible pour le même gène. Dans ce cas-là, l'expression est dite bistable. A l'échelle de la population, cela peut se traduire par de fortes différences phénotypiques entre les cellules : une population bactérienne clonale peut ainsi se diviser en deux sous-populations, ayant des comportements opposés sur un caractère particulier. Par exemple, une partie de la population sera fortement productrice d'une protéine, alors que l'autre ne la produira pas du tout (Veening *et al.*, 2008).

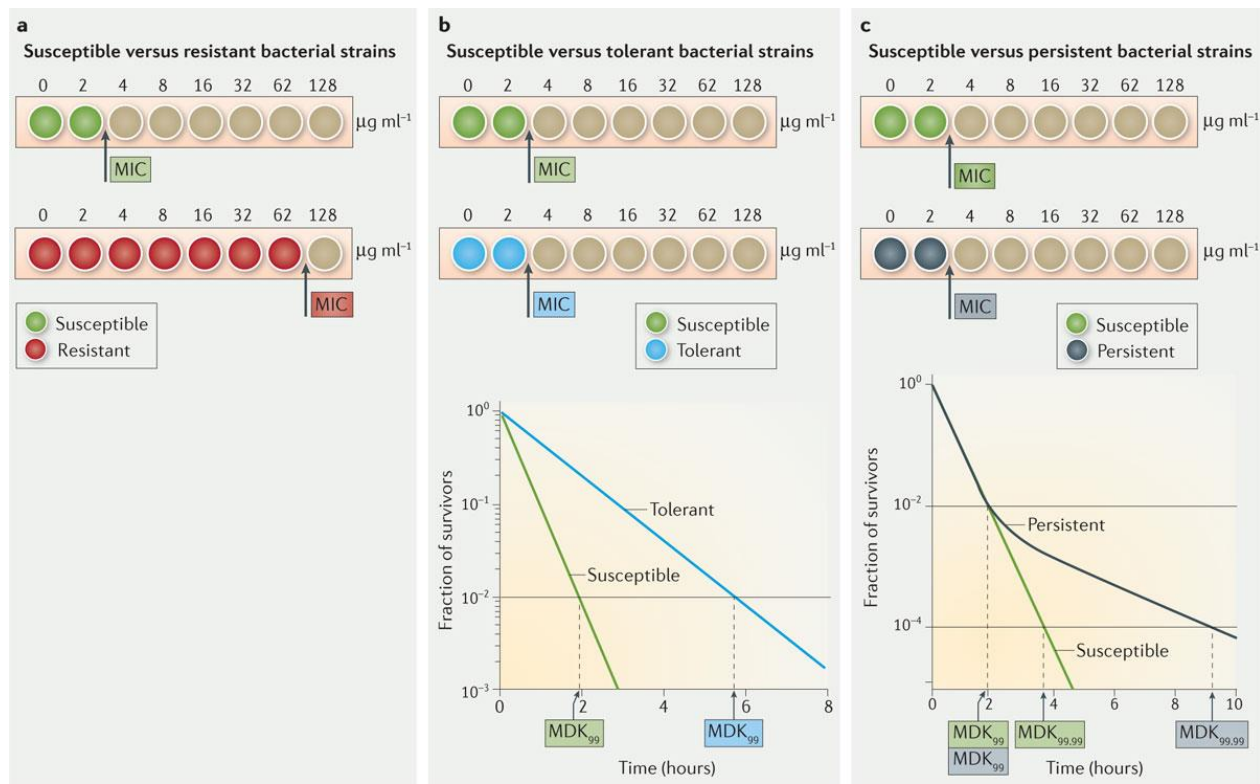
Le cas de la sporulation chez *Bacillus subtilis* en est un exemple. Le régulateur de la formation de spore chez *B. subtilis* est Spo0A, qui est induit lors de carence nutritive. Il a été montré que seule une partie de la population exprime le gène *spo0A*. Ainsi, l'expression de *spo0A* est bistable au sein de la population en raison de la boucle de rétrocontrôle positive activant l'expression de *spo0A*. Spo0A active une cascade de phospho-relais qui permet de renforcer l'expression du promoteur de *spo0A* (Veening *et al.*, 2005). Les cellules exprimant un niveau de *spo0A* plus fort vont activer toute la boucle de rétrocontrôle positive, permettant l'emballement du système, menant à la production de spores. Les cellules ayant un niveau d'expression de *spo0A* plus faible n'activeront pas la boucle. Ainsi le programme génétique permettant l'entrée en sporulation n'est pas exprimé chez toutes les cellules.

Des boucles de rétrocontrôle similaires impliquées dans l'expression hétérogène de programmes génétiques ont aussi été observées dans la mise en place de la compétence génétique chez *B. subtilis* (Dubnau and Losick, 2006; Veening *et al.*, 2008), pour la mise en place de la virulence chez *Salmonella typhimurium* (Arnoldini *et al.*, 2014), dans l'établissement du «quorum sensing» chez différentes espèces (Bettenworth *et al.*, 2019), ou encore, dans la réponse au stress osmotique chez *Saccharomyces cerevisiae* (Pelet *et al.*, 2011).

III. 4. La persistance bactérienne

III. 4. A Définition

La persistance est à l'heure actuelle, un des phénomènes d'hétérogénéité phénotypique non héréditaire les plus intensivement étudiés. Le phénomène de persistance est défini par la présence rare et transitoire de cellules non croissantes pouvant tolérer de fortes concentrations d'antibiotiques au sein d'une population génétiquement identique et pourtant sensible aux antibiotiques. Une population contenant des "persisters" est donc par définition hétérogène (Gefen and Balaban, 2009; Maisonneuve and Gerdes, 2014). Le phénomène de persistance n'est pas héritable et doit donc être dissocié des phénomènes de résistance aux antibiotiques, qui sont héréditaires et acquis par mutation génétique. Ainsi, des cellules persistantes donneront une descendance composée de cellules sensibles et de cellules persistantes (Bigger, 1944; Balaban *et al.*, 2004). La présence de cellules persistantes peut se démontrer en effectuant des courbes de mortalité après traitement aux antibiotiques. La persistance est caractérisée par une courbe bi-phasique résultant de la somme des courbes de la sous-population sensible et de la sous-population tolérante (Brauner *et al.*, 2016) (Figure 16).



Nature Reviews | Microbiology

Figure 16 : Différences entre résistance, tolérance et persistance (Brauner *et al.*, 2016). Le MIC (Minimum Inhibitory Concentration) est la concentration d'antibiotique minimum nécessaire pour empêcher la réplication de la bactérie et le MDK₉₉ (Minimum Duration for Killing) est la durée minimum d'exposition à l'antibiotique pour tuer 99% de la population. En vert sont représentées les tendances pour une population sensible à l'antibiotique, en rouge pour une population résistante à l'antibiotique, en bleu pour une population tolérante et en gris pour une population persistante. a) La CMI pour une population résistante à un antibiotique est plus élevée que pour une population sensible. b) La CMI pour une population tolérante est similaire à la CMI pour une population sensible, mais la MDK₉₉ pour une population tolérante est plus élevée que la MDK₉₉ pour une souche sensible. c) Les populations sensibles et persistantes ont une MIC et une MDK₉₉ similaire, cependant la MDK₉₉ est plus élevée pour une population persistante que pour une population sensible.

III. 4. B Mécanismes

Les mécanismes moléculaires menant à la présence de cellules persistantes ne sont, pour l'heure actuelle, pas entièrement compris. Ce phénomène, a longtemps été lié à un état métabolique particulier, dit de dormance, état dans lequel le métabolisme cellulaire est restreint et donc moins sensible aux effets bactéricides. Ainsi les cellules persistantes sont des cellules non croissantes, où la traduction et la transcription sont fortement réduites (Kwan *et al.*, 2013).

Cependant, les avis sont aujourd'hui partagés. Il a été montré que l'état de dormance n'est pas suffisant pour expliquer la persistance. Toutes les cellules en dormance ne sont pas des cellules persistantes et inversement, toutes les cellules persistantes ne sont pas non-réplicatives avant le traitement antibiotique (Orman and Brynildsen, 2013). Il a toutefois été prouvé que la fraction non-réplicative des cellules est enrichie en cellules persistantes (Helaine *et al.*, 2014). Il est ainsi admis que la persistance est un phénomène codé génétiquement.

Chez *E. coli* le gène *hipA* est le premier gène identifié comme ayant un impact sur la persistance d'une population. En effet les mutants *hipA* ont une proportion de cellules persistantes plus élevée que chez la souche sauvage (Balaban *et al.*, 2004). Chez les mutants *hipA*, le passage par la phase stationnaire permet de générer des cellules persistantes. HipA fait partie du système toxine/antitoxine HipBA. L'étude des mutants *hipA* a permis de mettre en évidence le lien entre persistance et les systèmes toxines/antitoxines présents chez les bactéries. Par la suite plusieurs études ont montré que des mutations dans les systèmes toxines/antitoxine (T/TA) avaient un impact sur la proportion de cellules persistantes au sein de la population (Kim and Wood, 2010; Dörr *et al.*, 2010; Tripathi *et al.*, 2014). Ces systèmes T/TA permettent la virulence des souches et sont induits en conditions de stress.

Cependant, le lien direct entre l'apparition de cellules persistantes et les systèmes T/TA n'a toujours pas été établi. Il semblerait que leur apparition soit liée à la réponse stringente, via le (p)ppGpp, qui est mise en place en cas de carence, notamment en acides aminés. L'induction du (p)ppGpp est impliqué dans l'augmentation du nombre de cellules persistantes (Nguyen *et al.*, 2011; Helaine *et al.*, 2014; Harms *et al.*, 2016; Strugeon *et al.*, 2016). Ainsi, la proportion de cellules persistantes diminue chez les mutants *relA* et *spoT*,

où la synthèse de (p)ppGpp est inhibée (Nguyen *et al.*, 2011; Brown, 2019). De plus, la persistance semble aussi diminuer chez les mutants *crp* incapable de mettre en place la répression catabolique (Mok *et al.*, 2015).

Il a aussi été montré que le taux de persistants augmentait dans une population avec la diminution de la force protomotrice et avec la dépolarisation de la membrane (Verstraeten *et al.*, 2015; Wang *et al.*, 2018). D'autres équipes avancent des hypothèses liées à l'âge des cellules : les cellules âgées auraient une croissance plus lente, ainsi le vieillissement des cellules serait un mécanisme possible pour générer des cellules persistantes (Klapper *et al.*, 2007).

Le mécanisme moléculaire régulant la génération de cellules persistantes n'est toujours pas clair et il est possible que deux types de persistants existent. Dans un premier temps, certaines cellules persistantes seraient générées de manière stochastique et spontanée durant la phase exponentielle de croissance (Balaban *et al.*, 2004; Shah *et al.*, 2006; Radzikowski *et al.*, 2016). Dans un deuxième temps, des cellules persistantes seraient générées en lien avec la réponse au stress. Ainsi, l'augmentation de la production de cellules persistantes a été observée en corrélation avec l'activation de la réponse stringente (Nguyen *et al.*, 2011; Strugeon *et al.*, 2016), mais aussi lors de transition diauxique (Amato *et al.*, 2013) et de carence en azote (Brown, 2019). Toutes ces voies sont en lien avec l'augmentation intracellulaire de (p)ppGpp. Il a été proposé que la formation de cellules persistantes est liée à la réponse au stress, médiée par le facteur sigma RpoS (Radzikowski *et al.*, 2016), activé aussi par l'augmentation intracellulaire de (p)ppGpp.

Aujourd'hui, l'interaction entre l'augmentation du (p)ppGpp et la production de cellules persistantes reste non élucidée. Une étude a montré que même en absence de (p)ppGpp, des cellules persistantes peuvent se former (Chowdhury *et al.*, 2016). Ainsi la persistance d'une population augmente avec l'arrêt de la croissance par une limitation environnementale, mais cela ne semble pas être uniquement dû à la présence de (p)ppGpp.

III. 5. Hétérogénéité métabolique liée à l'utilisation de source de carbone

III. 5. A Définition de l'hétérogénéité métabolique

Cette partie traitera plus particulièrement des phénomènes d'hétérogénéités phénotypiques touchants le métabolisme carboné de la souche, appelé ici hétérogénéité métabolique. Ainsi, cette partie aborde l'hétérogénéité métabolique de croissance, d'utilisation des substrats en co-consommation et lors de carences nutritives. L'hétérogénéité en croissance diauxique sera traitée plus spécifiquement dans une deuxième partie. L'évolution des techniques au niveau de la cellule unique a permis de mettre en évidence que des cellules clonales peuvent présenter d'importantes variabilités dans leur métabolisme, pouvant mener dans des cas extrêmes à des sous-populations distinctes ayant des activités intracellulaires différentes. Ce phénomène n'est pas dépendant de l'environnement : en condition de chemostat infiniment mélangé, une étude en spectrométrie de masse a permis de montrer que des cellules de *Staphylococcus aureus* présentaient un taux d'assimilation d'azote différent entre elles (Kopf *et al.*, 2015). Cette hétérogénéité de l'activité cellulaire apparaît même dans des populations cultivées en culture liquide continue bien mélangées, où il n'y a théoriquement pas de gradient qui aurait pu créer des différences environnementales.

Cette hétérogénéité métabolique semble apparaître à différents niveaux. Ainsi le taux de croissance lors de phase de croissance stable peut être variable au sein d'une population (Kiviet *et al.*, 2014; Hashimoto *et al.*, 2016). Chez *E. coli*, après une carence en carbone prolongée, des cellules ont été remises en croissance en milieu riche, certaines cellules sont capables de redemarrer leur croissance immédiatement et une sous-population de cellules ne redemarrant pas leurs croissances. Elles sont ici considérées comme dormantes (Şimşek and Kim, 2018). Des cultures en mélange glucose/acétate chez *E. coli* ont permis de mettre en évidence des cellules plus robustes au stress inhibiteur de l'acétate que d'autres (Heins *et al.*, 2019).

Cette hétérogénéité se manifeste aussi au niveau de l'assimilation des substrats. Chez *E. coli* une étude en spectrométrie de masse, en condition de chemostat limité en carbone avec un apport d'un mélange d'arabinose et de glucose, a montré que les deux

substrats sont consommés simultanément mais que la quantité relative absorbée de chaque substrat varie entre les cellules. Ainsi, au niveau de la population les cellules présentent différents niveaux de spécialisation dans l'assimilation et l'utilisation d'arabinose (Nikolic *et al.*, 2017). Des variations du taux de croissance ont aussi été observées entre les cellules.

Enfin, chez *Bacillus subtilis*, en co-consommation de glucose et de malate, une différenciation des voies métaboliques induites a pu être observée grâce à l'utilisation de marqueurs fluorescents (Rosenthal *et al.*, 2018). Cette hétérogénéité métabolique observée pourrait provenir de bruit dans le processus de transcription.

III. 5. B Le cas de l'hétérogénéité métabolique en croissance diauxique

La croissance en milieu diauxique est caractérisée par l'utilisation successive de deux substrats entrecoupés d'une phase de latence (voir II. 3. A). Cette phase a longtemps été décrite comme le temps nécessaire pour la réorganisation du métabolisme à l'échelle de la cellule (Monod, 1949; Loomis and Magasanik, 1967). Cependant, de plus récentes études ont montré que cette phase de latence diauxique pourrait être caractérisée par la présence d'une forte hétérogénéité entre les cellules.

Chez *E. coli* des transitions de glucose à d'autres substrats neoglucogéniques (acétate, fumarate, malate ou succinate) ont été réalisées (Kotte *et al.*, 2014). Au moment de la transition les membranes cellulaires ont été teintées avec un intercalant fluorescent, puis la mesure de la fluorescence a montré l'émergence de deux sous-populations. La première conservant un niveau de fluorescence élevé, car les cellules ne se diviseraient pas et la seconde en croissance où le nombre augmenterait au même rythme que la fluorescence déclinerait (lié à la dilution dans les cellules filles) (Kotte *et al.*, 2014). Dans leurs conditions les deux phénotypes différents semblent être générés au moment de la transition vers le second substrat. Cependant il semblerait que ce phénotype soit exacerbé par la présence d'importants lavages dans le protocole (Heinemann *et al.*, 2020).

Un phénomène similaire a été observé, chez la bactérie Gram+ *Lactococcus lactis* en mélange glucose/cellobiose (Solopova *et al.*, 2014). Pour cela une souche de *L. lactis* a été créée, contenant le promoteur *Pcel* (promoteur d'un gène du système de transport PTS du cellobiose) associé à une étiquette GFP et étudié en microscopie. Au moment de la

transition diauxique du glucose au cellobiose, *L. lactis* se différencie en deux sous-populations, une qui cesse de se diviser et qui est non fluorescente, tandis que l'autre continue de se diviser sur la deuxième source de carbone et présente une fluorescence verte.

Des phénotypes différents ont aussi pu être observés chez la levure *Saccharomyces cerevisiae* lors de transitions glucose/trehalose ou glucose/galactose (Heerden *et al.*, 2014; Siegal, 2015). Chez *Methylobacterium extroquens* sur mélange méthanol/succinate (McClelland *et al.*, 2020), c'est aussi le cas mais il semble ne pas y avoir de présence de sous populations distinctes, mais plutôt certaines cellules présentant un temps de latence de plusieurs heures tandis que d'autres ne recommencent pas à croître.

Il a aussi été reporté, lors de croissance diauxique sur mélange glucose-lactose, qu'environ 15% des cellules sont préalablement adaptées au second substrat et ne subissent pas de latence diauxique (Boulineau *et al.*, 2013). Le reste de la population présente une réelle phase de latence, dont la durée est variable entre les cellules. Une autre étude a montré que des changements successifs de substrat entraînent l'apparition d'une sous-population capable de démarrer sa croissance plus rapidement. Dans leur condition d'étude, 27% des cellules ont une latence faible, 68% une latence longue et 5% sont incapables de repartir en croissance (Kaiser *et al.*, 2018). Cependant, l'apparition de sous-populations bien distinctes ne semble pas être une généralité et dépend des conditions d'étude (Heins *et al.*, 2019). Le manque d'études actuelles ne permet pas de généraliser le phénomène et son mécanisme d'apparition. Le devenir de ces cellules, incapables de transiter sur le second substrat reste aussi à élucider. Plusieurs pistes sont à l'étude, notamment sur l'avantage adaptatif (décrit en partie III. 6) que peut conférer un réservoir de cellules non croissantes. Il a été montré chez *E. coli*, que la formation de cellules persistantes augmente lors de transitions diauxiques. En effet, lors de croissance en mélange diauxique, le pourcentage de cellules persistantes augmente à la disparition du premier substrat (glucose), et est significativement supérieur au taux de cellules persistantes après la disparition du glucose lors de croissance uniquement sur glucose (Amato *et al.*, 2013).

Ces différents exemples montrent qu'en présence de différentes sources de carbone, l'ensemble d'une population isogénique n'aurait pas le même comportement

métabolique lors de la disparition du premier substrat. Le temps de latence semble variable entre les cellules et certaines pourraient être incapables de transiter sur le second substrat.

III. 5. C Mécanisme d'apparition de l'hétérogénéité métabolique

Le mécanisme sous-tendant ce phénomène reste, pour le moment, toujours incertain. La communauté scientifique s'accorde à dire qu'il s'agit d'une régulation épigénétique, c'est-à-dire que le phénomène n'est pas lié à des mutations génétiques entre les cellules. Actuellement, le peu d'études ne permettent pas de mettre en évidence un mécanisme consensus.

Cependant, l'hypothèse majeure est que l'hétérogénéité de l'utilisation des sources de carbone provient de l'expression stochastique des gènes. Cette variation d'expression des gènes, appelée bruit d'expression, pourrait mener à une variation des contenus intracellulaires (Figure 17). En effet, le niveau d'expression d'un gène dans une population n'est pas identique entre cellules, c'est-à-dire que certaines cellules peuvent avoir une expression beaucoup plus élevée ou beaucoup plus basse que la moyenne pour un même gène (Elowitz *et al.*, 2002; Afroz *et al.*, 2014). Il est aussi possible que la variation du taux d'expression des gènes soit liée à la structure physique de l'ADN (Casadesús and Low, 2013; Ackermann, 2015).

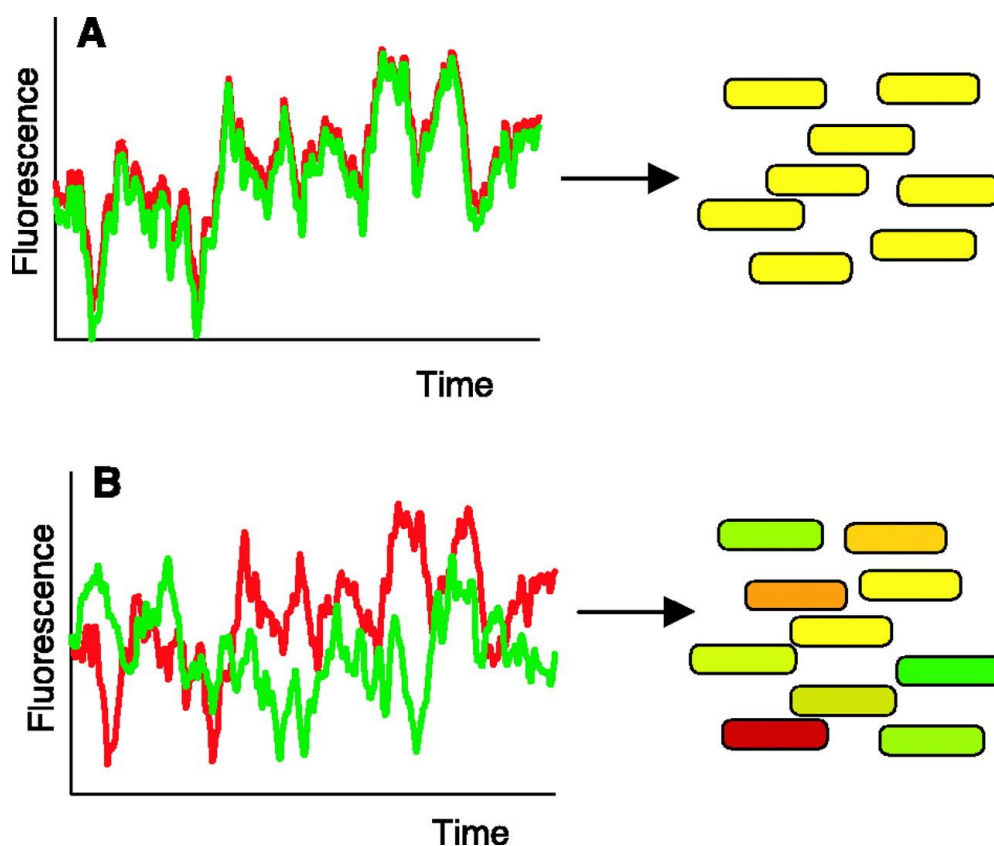


Figure 17 : L'expression stochastique des gènes peut mener à une variation de contenu intra cellulaire (Elowitz *et al.*, 2002). Les bruits intrinsèques (inhérents aux processus biochimiques d'expression des gènes) et extrinsèques (liés à des fluctuations des autres composantes cellulaires) peuvent être mesurés et distingués avec deux gènes (cfp, représenté en vert; yfp, représenté en rouge), contrôlés par des séquences régulatrices identiques. Les cellules contenant la même quantité de chaque protéine apparaissent en jaune, tandis que les cellules exprimant plus une protéine fluorescente que l'autre apparaissent en rouge ou en vert. (A) En l'absence de bruit intrinsèque, les deux protéines fluorescentes fluctuent de manière corrélée au fil du temps dans une seule cellule. Ainsi, dans une population, chaque cellule aura la même quantité des deux protéines, bien que cette quantité puisse différer d'une cellule à l'autre en raison du bruit extrinsèque. (B) L'expression des deux gènes peut se décorrélérer à l'échelle des cellules individuelles en raison du bruit intrinsèque, donnant lieu à une population dans laquelle certaines cellules expriment plus une protéine fluorescente que l'autre.

La variation de ce niveau d'expression pourrait entraîner des boucles de régulations menant à un comportement métabolique différent (Choi *et al.*, 2008; Labhsetwar *et al.*, 2013; Ackermann, 2015). Ce phénomène pourrait plus précisément être dû à la variation de la concentration des facteurs de transcription soumis à la répression catabolique ou à l'activation de la réponse stringente. En effet, dans leur étude, Kotte *et al.*, (2014) ont montré que leur population d'*E. coli* posséderait une distribution bimodale des taux de croissance sur acétate, certaines cellules étant capables de croître sur acétate, d'autre non. Ils attribuent cela à l'expression stochastique de *cra*, un facteur de transcription global et interrogent le rôle de métabolites intervenant dans la régulation comme le fructose-1,6-bisphosphate (FBP) inhibiteur de *cra* lors de croissance sur glucose.

Lors de croissance en mélange glucose/lactose l'expression basale des enzymes impliquées dans le métabolisme du lactose serait, ici aussi, stochastique et permettrait aux cellules avec une plus forte concentration de lactose perméase d'induire le métabolisme du lactose rapidement (Boulineau *et al.*, 2013). Cependant dans leurs conditions, une petite proportion de cellules semble avoir au contraire une phase de latence beaucoup plus longue que la moyenne qui pourrait, selon eux être corrélée à l'activation de la réponse stringente.

Une autre étude émet l'hypothèse de l'implication de la réponse stringente (Solopova *et al.*, 2014). Ils proposent que l'hétérogénéité métabolique qu'ils observent provient d'une capacité différente de chaque cellule à gérer la diminution de la répression catabolique nécessaire pour croître sur le deuxième sucre et l'activation de la réponse stringente, empêchant certaines cellules de consommer le second substrat. Selon eux, la fraction de cellules capable de consommer le second substrat dépend de phénomènes épigénétiques et des conditions de préculture.

Chez la levure *Saccharomyces cerevisiae*, lors de transition artificielle de glucose à maltose, certaines cellules démarrent leur croissance directement, d'autre au bout de 20h. L'hypothèse de Cerulus *et al.*, 2018 est que pendant la croissance sur glucose, certaines cellules présentent une répression de la respiration plus forte que d'autres. Les cellules dont la respiration est moins verrouillée peuvent présenter des phases de latence plus

courtes. De plus, il a aussi été observé que les cellules de levure pouvaient être préparées au second substrat par l'expression antérieure des opérons nécessaires à sa consommation, ce phénomène est lié à une répression catabolique plus faible chez certaines cellules (New *et al.*, 2014; Wang *et al.*, 2015). Cela n'a pour l'instant jamais été observé chez les procaryotes.

Outre les phénomènes d'expression stochastique des gènes, d'autres hypothèses existent quant aux mécanismes créant l'hétérogénéité métabolique. Il est possible que cette variabilité soit due à un héritage protéique, aussi appelé mémoire phénotypique, liée à la transmission de protéines cytoplasmiques aux cellules filles (Lambert and Kussell, 2014). Cette transmission protéique pourrait se faire de manière asymétrique entraînant une variabilité intracellulaire dans la descendance (Huh and Paulsson, 2011). De plus il semblerait qu'indépendamment du niveau d'expression des gènes, il existe aussi une stochasticité au niveau des réactions métaboliques pouvant affecter le taux de production et de consommation de métabolites ainsi que la réserve énergétique des cellules (Delvigne *et al.*, 2014).

L'amplitude du phénomène, sa régulation et sa généralité restent encore à caractériser. Aujourd'hui le peu d'études semblent mettre en évidence l'effet du bruit dans l'expression des gènes. Cependant la quantification directe n'a pas été réalisée et la manière dont la variation d'expression des gènes d'une cellule à l'autre mène à différents phénotypes reste peu approfondie. Toutefois, il semble clair que l'organisation cellulaire ne soit pas aussi simple que celle décrite par le concept d'une réorganisation complète et homogène du métabolisme pour l'ensemble de la population tel qu'historiquement admis.

III. 6. Le rôle de l'hétérogénéité phénotypique

Le rôle de ces phénomènes d'hétérogénéité est aujourd'hui questionné. D'un premier abord, la variabilité intercellulaire pourrait mener à ce que toutes les cellules ne soient pas adaptées de manière optimale à leur condition de croissance et donc semblerait réduire l'efficacité sur un environnement stable. Cependant, les micro-organismes en conditions environnementales réelles sont soumis à des changements fréquents dans la quantité et la nature des ressources disponibles. Ainsi, l'adaptation à ces changements est un facteur crucial pour la survie de la population.

Une des hypothèses du rôle de l'hétérogénéité métabolique est la mise en place d'une stratégie de minimisation des risques, dite de « bet-hedging » (Veening *et al.*, 2008; Grimbergen *et al.*, 2015), selon laquelle la multiplication aléatoire de différents phénotypes au sein d'une population augmenterait la probabilité qu'au moins une d'elles soit plus adaptée aux changements environnementaux, que ce soit l'apparition d'antibiotiques (Amato *et al.*, 2013) ou un changement de substrat (Solopova *et al.*, 2014). Ainsi, lors de variations environnementales (épuisement d'un substrat ou autre type de stress), cette hétérogénéité constituerait un avantage adaptatif. En effet, elle permettrait une meilleure compétitivité de la population, grâce à un redémarrage plus rapide d'une des sous-populations après l'épuisement du substrat préférentiel via la réorientation sans latence du métabolisme carboné central (Boulineau *et al.*, 2013). Cela est soutenu par des études *in silico* qui ont montré que des individus avec un plus grand nombre de facteur de transcription Cra étaient capable de transiter plus vite sur acétate lors de transition glucose/acétate (Schmutzer and Wagner, 2020). De la même manière, en modélisation *in silico*, une population de *L. lactis* hétérogène contenant des cellules capables de métaboliser le cellobiose et d'autres non, semblerait présenter un avantage sur une population où toutes les cellules sont capables de le métaboliser, lorsque les conditions environnementales futures sont imprévisibles (Solopova *et al.*, 2014).

D'autre émettent l'hypothèse que cette différenciation ne serait pas aléatoire, et que l'hétérogénéité phénotypique pourrait être une stratégie alternative pour l'adaptation aux environnements changeants par des cellules qui s'adaptent en jouant sur leur état physiologique. Ainsi l'hétérogénéité serait influencée par l'environnement et les antécédents de croissance (Rosenthal *et al.*, 2018; Heins *et al.*, 2019).

Un autre des rôles de l'hétérogénéité phénotypique serait la division du travail entre les individus appelé « division of labour ». Dans ce cas-là, la population est hétérogène avec un phénotype de cellule qui est bénéfique pour le second phénotype existant dans la population. Cette hypothèse induit une interaction bénéfique entre les sous-populations. Par exemple, chez *Salmonella typhimurium*, seule une sous-population exprime les gènes de virulence permettant le développement de la population entière. De plus, l'expression de ces gènes de virulence est corrélée avec l'augmentation de la persistance (Arnoldini *et al.*, 2014). L'hétérogénéité métabolique est un facteur important à intégrer dans les modèles d'infection car elle influence la virulence et la sensibilité aux antibiotiques (Ruiz *et al.*, 2020). Un comportement similaire a été observé chez *Bacillus subtilis* avec une sous-population capable de produire de l'acétoïne, qui permet de stabiliser le pH extracellulaire, en faveur d'une sous-population qui ne produit pas d'acétoïne mais qui croît plus rapidement (Elowitz *et al.*, 2002).

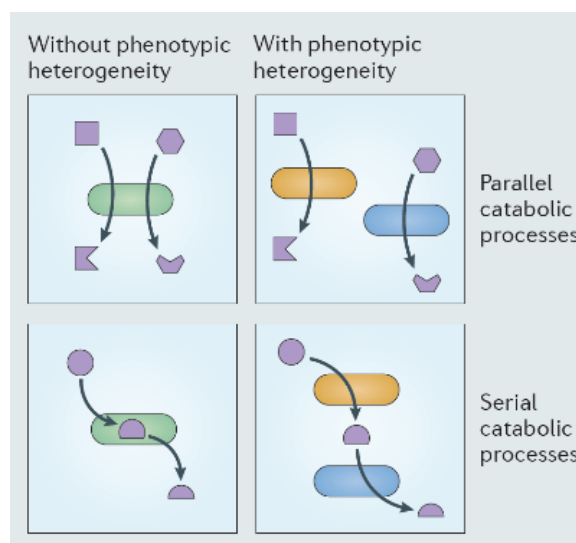


Figure 18 : L'importance de l'hétérogénéité phénotypique dans la division du travail (Ackermann, 2015)

Il est aujourd'hui admis que l'hétérogénéité métabolique pourrait contribuer à une meilleure résilience de la population face à des variations environnementales (Ackermann, 2015). La variation phénotypique issue du bruit d'expression génique peut améliorer les capacités en réduisant le temps de latence d'une population avant qu'elle ne reprenne sa croissance dans un nouvel environnement. Ce phénomène d'hétérogénéité est donc un nouveau facteur à prendre en compte dans l'étude de la capacité adaptative des populations bactériennes.

IV. Objectifs de la thèse :

L'introduction bibliographique expose l'importance pour la bactérie *Escherichia coli* de s'adapter aux changements d'environnement. Naturellement présente dans la flore intestinale, cette espèce bactérienne est soumise à des variations fréquentes de la quantité et de la nature des nutriments disponibles. S'adapter efficacement à ces fluctuations est nécessaire pour la survie de l'espèce en milieu compétitif. L'apport variable de nutriments entraîne une alternance entre phase de croissance et phase de latence. Dans le cas d'un apport de sources carbonées multiples, la cellule utilise ces ressources séquentiellement ou simultanément, en fonction de leur nature. Dans le cas de la croissance diauxique, le substrat préférentiel est celui permettant le plus haut taux de croissance. A l'épuisement de ce dernier, la population entre en phase de latence avant d'entamer la consommation du substrat suivant. Cette phase de latence est considérée comme une phase d'adaptation au second substrat. Néanmoins, l'arrêt de croissance au niveau de la population globale peut entraîner un ralentissement du développement et donc de la compétitivité pour la consommation des ressources. Ces transitions ne sont pour l'heure actuelle pas entièrement comprises.

Les comportements phénotypiques bactériens ont longtemps été étudiés à l'échelle de la population, regardée comme homogène. Ainsi, il a été classiquement admis que la phase de latence correspond au temps requis par la population pour effectuer une réorganisation métabolique homogène permettant la consommation du second substrat. Plus récemment, le développement des techniques d'étude à l'échelle des cellules individuelles a permis de mettre en évidence des variations observables entre individus issus de la même population. La notion d'hétérogénéité phénotypique a ainsi émergé.

Nous avons présenté dans l'introduction bibliographique des études révélant qu'un changement de substrat pouvait favoriser l'émergence de sous-populations. Cette organisation différentielle du métabolisme au sein des populations pourrait avoir un impact sur la capacité d'adaptation des bactéries. Ce phénomène a été récemment observé chez *Escherichia coli* en mélange glucose/lactose (Kaiser *et al.*, 2018), glucose/acétate et glucose/furamate (Kotte *et al.*, 2014), chez *Lactococcus lactis* en mélange glucose/cellobiose (Solopova *et al.*, 2014), ainsi que chez *Saccharomyces cerevisiae* en mélange glucose/galactose (Heerden *et al.*, 2014). Ainsi, la phase de latence observée au

niveau de la population pourrait être due au temps nécessaire pour que les cellules en croissance constituent une fraction importante de la population (Kotte *et al.*, 2014). A partir de ces études, notre hypothèse principale de travail est que l'adaptation pourrait être le fait d'une sous-population, toutes les cellules ne s'adaptant pas au changement de substrat. Les cellules non adaptées pourraient permettre de fournir un réservoir non croissant, mais résistant en cas de stress majeur subit par la population.

Les mécanismes qui pourraient expliquer ce phénomène commencent à être étudiés mais restent encore très mal compris. D'une étude à l'autre, les explications mécanistiques avancées sont différentes (voir III. 5. C). A l'heure actuelle, le peu de conditions étudiées ne permettent pas de mettre en évidence un mécanisme consensus. Dans un grand nombre de cas, le mécanisme menant à cette hétérogénéité de la population reste à élucider.

Les micro-organismes sont aujourd'hui des outils de transformation de la matière au même titre que la synthèse chimique de molécule. L'étude des phénomènes d'hétérogénéité métabolique présente donc un intérêt majeur dans le domaine des bioprocédés. En effet, dans un contexte industriel, cette hétérogénéité au sein de la population pourrait avoir un impact sur le rendement des bioprocédés. L'optimisation des procédés de fermentation (maximisation des rendements, diminution des durées de fermentation) ainsi que l'optimisation d'utilisation de substrats complexes passent par la compréhension de ces mécanismes. Dans ce contexte-là, mon projet de thèse a pour but l'étude approfondie de la transition glucose/xylose chez la bactérie *Escherichia coli*, qui a peu été étudié et n'a jamais été regardé du point de vue de l'hétérogénéité de la population. Ce mélange de substrats carbonés a un intérêt dans la valorisation de la biomasse végétale, car le xylose et le glucose sont les principaux monomères issus de la dégradation de la lignocellulose (Okeke and Obi, 1994). De plus le glucose est le substrat préférentiel chez *Escherichia coli* et l'utilisation du xylose implique une voie métabolique, décrite précédemment (I. 4), bien connue et comportant peu d'intermédiaires. Ainsi nous avons utilisé cette transition comme modèle d'étude.

Le premier objectif de cette thèse a consisté à mettre en place une méthodologie permettant d'observer la variabilité phénotypique d'une population clonale d'*Escherichia coli*, lors de croissance en mélange glucose/xylose. Pour cela nous avons choisi d'utiliser des protéines fluorescentes comme rapporteurs du métabolisme des cellules. Ce travail

de mise au point a été conséquent au début de la thèse et l'optimisation de la méthodologie a été poursuivie tout au long de l'étude. Cette partie est présentée dans le premier chapitre de la thèse.

Dans un deuxième temps, nous avons procédé à l'étude approfondie de la croissance diauxique glucose/xylose chez la bactérie *Escherichia coli*, avec comme objectif de tester la potentielle présence de sous-populations dans ces conditions d'étude. Nous avons dans cette partie, étudié les mécanismes impliqués dans la transition et dans la régulation du phénomène d'hétérogénéité métabolique. La dernière partie sera focalisée sur l'analyse du phénotype des sous-populations présentes afin d'identifier leurs caractéristiques.

Chapitre 1 :

Mise en place de la méthodologie pour l'étude des sous populations

Dans cette thèse nous nous intéressons à la transition glucose xylose dans les populations d'*Escherichia coli*. Afin de pouvoir étudier le comportement d'*E. coli* à l'échelle de la cellule, il est nécessaire de construire des marqueurs biologiques permettant de rendre compte de l'état physiologique de la cellule sur le critère de l'utilisation de substrat. Ces marqueurs nous permettront d'observer des possibles phénomènes d'hétérogénéité au sein de la population. Cette première partie de la thèse porte ainsi sur la mise au point de notre méthodologie d'étude pour la quantification de l'hétérogénéité métabolique en mélange glucose/xylose chez *E. coli*. Le travail sur cette méthodologie a été important au début de l'étude par la mise en place, mais aussi tout au long de la thèse par son optimisation progressive. Ce chapitre traite des choix expérimentaux réalisés pour étudier l'hétérogénéité lors de croissance en mélange glucose/xylose, ainsi que des tests réalisés pour valider la méthodologie mise en place.

I. Choix expérimentaux pour la construction des marqueurs de l'hétérogénéité

La première partie porte sur le choix de la condition modèle et des méthodologies d'étude des sous-populations ainsi que sur la réalisation des premières constructions et les premiers tests permettant de vérifier la fiabilité de notre méthodologie.

I. 1. Choix expérimentaux

I. 1. A Le modèle glucose/xylose

Le glucose et le xylose sont deux sucres issus de la dégradation de la lignocellulose, molécule qui constitue la paroi des végétaux (Okeke and Obi, 1994). L'utilisation de ce mélange présente ainsi un intérêt économique fort car il permettrait de revaloriser les déchets de biomasse végétale en apport de carbone pour la production d'autres molécules d'intérêt (Dien *et al.*, 2002; Cirino *et al.*, 2006; Saini *et al.*, 2017). Comme montré précédemment, l'utilisation du glucose et du xylose chez *Escherichia coli* implique des voies métaboliques connues et décrites (Bibliographie I. 3 et I.4). Ainsi nous avons utilisé cette transition comme modèle d'étude, nous nous sommes plus particulièrement

intéressés à la croissance diauxique, car il s'agit du modèle le plus proche des conditions naturelles.

I.1.B Zones promotrices

Pour cette étude, il est nécessaire de mettre au point un système permettant de distinguer les cellules qui consomment le glucose de celles qui consomment le xylose. Pour cela, nous avons décidé d'utiliser des zones promotrices de gènes spécifiquement exprimés sur l'un des deux substrats et permettant l'expression d'un rapporteur quantifiable. Nous avons recherché des marqueurs de la consommation du xylose, du glucose mais aussi de l'acétate qui est produit lors de la consommation du glucose comme du xylose. Ainsi, une quinzaine de promoteurs ont été sélectionnés (Tableau 1). Les zones promotrices utilisées sont celles de gènes impliqués dans le métabolisme de chacun des substrats. Ces gènes codent pour des transporteurs ou des enzymes cataboliques sélectionnés sur la base de leur fonction et sur leur profil d'expression. Ces données d'expression ont été obtenues à partir de plusieurs études de transcriptome (Liu *et al.*, 2005; Enjalbert *et al.*, 2013; Sandberg *et al.*, 2017) (Figure 19). Seul les gènes, *ompC*, *ompX* et *glgS*, ne codent pas pour des protéines qui interviennent directement dans le métabolisme du substrat dont elles sont théoriquement spécifiques. Ces trois gènes ont été sélectionnés car leur expression est significativement supérieure lors de la croissance sur l'un des trois substrats de l'étude, soit le glucose pour *ompC* et *ompX*, soit l'acétate pour *glgS*. Les voies métaboliques d'utilisation des trois substrats de l'étude ont été détaillées dans la partie précédente (Bibliographie partie I. 3 ; I. 4 et I. 5).

Promoteur du gène :	Fonction métabolique du gène	Spécificité théorique
<i>xylA</i>	Dégradation du xylose en xylulose	Xylose
<i>xylF</i>	Transporteur du xylose	Xylose
<i>xylE</i>	Transporteur du xylose	Xylose
<i>ptsG</i>	Transporteur du glucose	Glucose
<i>ptsH</i>	Impliqué dans le phosphorelais du transporteur PtsG (transport du glucose)	Glucose
<i>pfkA</i>	Enzyme de la glycolyse	Glucose
<i>pfkB</i>	Enzyme de la glycolyse	Glucose
<i>gapA</i>	Enzyme de la glycolyse	Glucose
<i>icd</i>	Enzyme du cycle du krebs	Glucose
<i>ompC</i>	Protéine de la membrane externe	Glucose
<i>ompX</i>	Protéine de la membrane externe	Glucose
<i>aceB</i>	Shunt glyoxylique	Acétate
<i>acs</i>	Dégradation acétate en acétyl-CoA	Acétate
<i>pck</i>	Transformation de l'oxaloacetate en phosphoenolpyruvate (PEP)	Acétate
<i>glgS</i>	Régulateur de la composition de la surface cellulaire	Acétate
<i>ihfB</i>	Sous unité beta d'un facteur d'intégration	Constitutive
<i>cysG</i>	Sirohème synthetase	Constitutive
<i>hcaT</i>	Transporteur avec un domaine transmembranaire	Constitutive
<i>idnT</i>	Transporteur du L-idonate et du D-gluconate	Constitutive

Tableau 1 : Liste des gènes dont les zones promotrices ont été testées dans le cadre de l'étude. La fonction du gène est spécifiée, ainsi que l'expression spécifique théorique du gène.

En plus des zones promotrices spécifiques de l'utilisation d'un substrat, il est intéressant de disposer de promoteurs constitutifs comme contrôles. Ainsi en parallèle de la sélection des zones promotrices spécifiques de l'utilisation d'un substrat, nous avons sélectionné les promoteurs des gènes *idnT*, *ihfB*, *cysG* ou *hcaT*, décrits pour avoir une expression constitutive (Zhou *et al.*, 2011) (Figure 19).

	Study	E-MTAB-3392	Liu et al., 2005	E-MTAB-3392	Sandberg 2017	Sandberg 2017	Liu et al., 2005	Liu et al., 2005	Liu et al., 2005
	Condition	Glucose	Acetate	Acetate	Acetate	Xylose	Succinate	Alanine	Proline
Identifiant	Gene name								
b1101	PtsG	1.0	0.2	0.2	0.1	0.3	0.1	0.2	0.2
b2415	PtsH	1.0	0.4	0.3	0.7	0.6	0.5	0.5	0.4
b3916	PfkA	1.0	0.7	0.4	0.6	1.3	0.6	0.6	0.6
b1723	PfkB	1.0	0.6	0.7	0.8	0.4	0.6	0.7	0.5
b1779	GapA	1.0	0.3	0.5	0.6	1.5	0.5	0.4	0.3
b1136	Icd	1.0	0.7	0.6	0.4	0.1	0.8	0.8	0.6
b2215	OmpC	1.0	1.3	0.7	1.0	0.5	0.5	1.0	1.4
b0814	OmpX	1.0	0.3	0.3	0.3	0.5	0.4	0.3	0.4
b4014	AceB	1.0	4.0	1.7	2.7	0.1	1.7	2.5	4.4
b4015	AceA	1.0	6.3	2.1	3.6	0.1	2.8	4.0	7.3
b4069	Acs	1.0	19.3	15.1	7.4	0.1	6.4	12.5	20.2
b3403	Pck	1.0	4.6	6.6	5.1	5.6	3.5	3.8	3.7
b3049	GlgS	1.0	3.6	2.3	4.5	1.5	1.0	2.5	2.7
b0912	IhfB	1.0	1.3	1.2	1.2	1.2	1.1	1.2	1.2
b3368	CysG	1.0	0.9	0.4	1.0	0.9	1.0	1.0	1.1
b2536	HcaT	1.0	1.6	0.7	1.0	0.7	1.4	1.6	1.7
b4265	IdnT	1.0	0.9	0.8	0.6	1.2	1.1	1.0	1.1
b3564	XylB	1.0	3.0	0.7	1.1	87.4	2.1	3.5	3.5
b3565	XylA	1.0	5.1	1.0	3.2	720.1	1.6	3.2	4.3
b3566	XylF	1.0	9.5	20.6	3.8	370.2	2.3	6.2	8.1
b3567	XylG	1.0	2.8	8.4	1.3	185.3	1.1	1.6	2.4
b3568	XylH	1.0	1.8	1.7	3.0	198.0	1.5	1.4	2.1
b3569	XylR	1.0	1.1	0.5	1.0	9.2	1.4	1.3	1.4

Figure 19 : Représentation du niveau d'expression des gènes par rapport au niveau d'expression sur glucose, lors de croissances sur différents substrats. Les données sont issues de trois études (Liu *et al.*, 2005; Enjalbert *et al.*, 2013; Sandberg *et al.*, 2017) et ont servi à la sélection des promoteurs d'intérêts de cette thèse.

I.1.C Rapporteur de l'expression: fluorescence

Nous avons choisi d'utiliser la fluorescence comme marqueur biologique. Le principe de la fluorescence repose sur l'état d'excitation des molécules. Une molécule fluorescente est capable d'absorber un photon issu d'un rayonnement lumineux. Elle se trouve alors dans un état excité. L'état excité de la molécule étant instable, l'énergie absorbée est restituée sous forme d'émission de photon, donc d'émission de lumière, à une certaine longueur d'onde, différente de celle de l'excitation. Durant l'excitation de la molécule, une partie de l'énergie est dissipée par vibration des électrons, ainsi l'énergie émise est plus faible que l'énergie excitatrice et la longueur d'onde d'émission est plus élevée que la longueur d'onde d'excitation (Shashkova and Leake, 2017). Une molécule capable d'émettre de la fluorescence après excitation est appelé fluorophore (ou fluorochrome).

Nous avons utilisé des protéines fluorescentes comme rapporteurs de l'utilisation d'un substrat. La fluorescence présente l'avantage d'être non toxique pour l'expérimentaliste et peu toxique pour la cellule (Shaner *et al.*, 2005). L'expression de cette fluorescence nous permettra de pouvoir analyser les cellules par cytométrie en flux. Cette technologie traite l'information des cellules individuellement et ne donne pas qu'une moyenne du signal contrairement à d'autres techniques de fluorimétrie. De plus, elle présente la possibilité de traiter un grand nombre de cellules en peu de temps, jusqu'à 15 000 évènements/s pour le MACSQuant utilisé dans cette étude (Miltenyi Biotec, Allemagne), ce qui est beaucoup plus que la capacité de traitement en microscopie. Ainsi nous avons choisi d'utiliser des molécules fluorescentes avec des longueurs d'ondes d'émissions éloignées afin de ne pas avoir d'interférences si nous choisissons d'en utiliser plusieurs ensemble (Tableau 2). Les gènes codants pour des protéines fluorescentes ont été placés sous contrôle des zones promotrices de gènes natifs de *E. coli*, présentés précédemment. Cette zone promotrice correspond aux 800 bp en amont de l'ATG. Il s'agit donc d'une fusion transcriptionnelle, car elle intervient dans la région transcrite mais non traduite.

Rapporteur	Excitation (nm)	Emission (nm)	Couleur	FPbaseID
mRFP1	584	607	Rouge	5YCFA
sYFP2	515	527	Jaune/Vert	GB9EO
mTagBFP	399	511	Bleu	BFJKS

Tableau 2 : Tableau des différents rapporteurs fluorescents utilisés dans cette étude, avec leurs longueurs d'ondes d'excitation et d'émission, la couleur de la fluorescence émise et le numéro d'identification de la protéine dans la « fluorescence protein data base » (Lambert, 2019).

I.1.D Construction plasmidique

La méthode que nous avons mise en œuvre pour mesurer l'hétérogénéité métabolique chez *E. coli* est basée sur l'utilisation de vecteur plasmidique. Les plasmides construits permettent l'expression d'un marqueur fluorescent en réponse à l'utilisation d'un substrat particulier. La construction du rapporteur sur châssis plasmidique présente l'avantage d'augmenter significativement le signal, en comparaison à une construction intégrée dans le génome. De plus, le vecteur plasmidique offre une maniabilité intéressante puisqu'il est possible d'insérer facilement le plasmide dans différentes souches d'*Escherichia coli*. Cependant l'utilisation de plasmide peut entraîner une surcharge métabolique chez la cellule, et une possible variabilité du nombre de copies de plasmide entre cellules, ces paramètres devant être contrôlés.

Les plasmides ont été construits selon le modèle présenté en figure 20, dans le vecteur pSB1C3. Il s'agit d'un plasmide possédant une origine de réplication ColE1 dérégulée. Il se retrouve donc à haut nombre de copies dans la cellule (environ 200 copies). Le plasmide pSB1C3 porte un gène de résistance au chloramphénicol, permettant la sélection des cellules possédant le plasmide. Le chloramphénicol inhibe la synthèse protéique en se liant à la sous-unité 23S du ribosome. L'avantage de cet antibiotique est que la résistance se met en place au niveau intracellulaire, limitant le risque de développement de cellules ne possédant pas la résistance par effet satellite. Le plasmide pSB1C3 est le châssis de la collection iGEM, ainsi, nous avons à disposition un jeu de plusieurs plasmides pSB1C3 possédant des gènes codants pour des protéines fluorescentes différentes, facilement inter-changeables (voir I. 2).

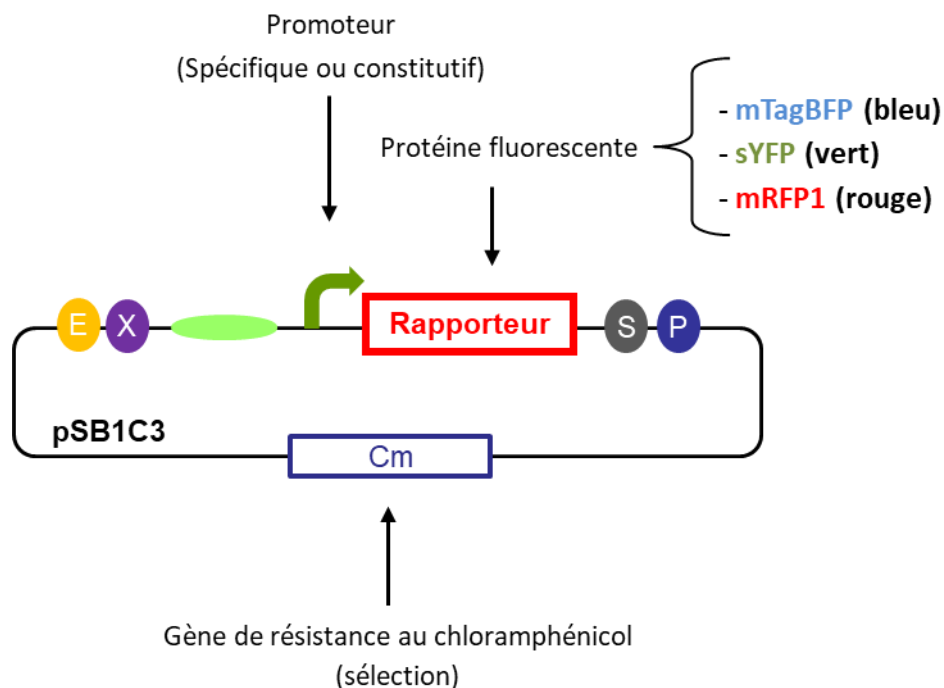


Figure 20 : Schéma modèle des constructions plasmidiques de l'étude. Le vecteur de base est pSB1C3 dans lequel un promoteur et un gène codant pour une protéine fluorescente ont été insérés entre les site EcoRI, XbaI et les sites SpeI et PstI. Les sites EcoRI, XbaI, SpeI et PstI sont respectivement représentés par les lettres E, X, S et P.

I.1.E Souches

Dans un premier temps, les plasmides ont été insérés par transformation de la souche *E. coli* BW25113, qui a servi de modèle principal pour cette étude. Cette souche est dérivée de la K12. Le séquençage de son génome a permis de montrer que les différences majoritaires entre les souches MG1655 (classiquement utilisée comme référence des *E. coli*) et BW25113 sont la perte de l'antigène O, du plasmide F, du bactériophage λ , ainsi que la délétion des opérons *araBAD* et *rhaDAB* (Grenier *et al.*, 2014). La souche BW25113 sert de support à la collection KEIO. Il s'agit d'une collection de 3985 souches mutantes d'*E. coli*, chacune inactivées pour un gène non essentiel (Baba *et al.*, 2006).

I. 2. Méthodologie de construction des premiers plasmides

Afin de produire tous les plasmides de l'étude sur le même format et de manière générique, la méthode de clonage utilisée pour réaliser les premiers vecteurs a été celle de l'iGEM, qui repose sur l'utilisation de 4 enzymes de restrictions : EcoRI, XbaI, SpeI et PstI (Figure 21). L'utilisation de cette méthode de clonage permet une combinaison simple des différentes constructions entre elles en utilisant la complémentarité des sites SpeI et XbaI.

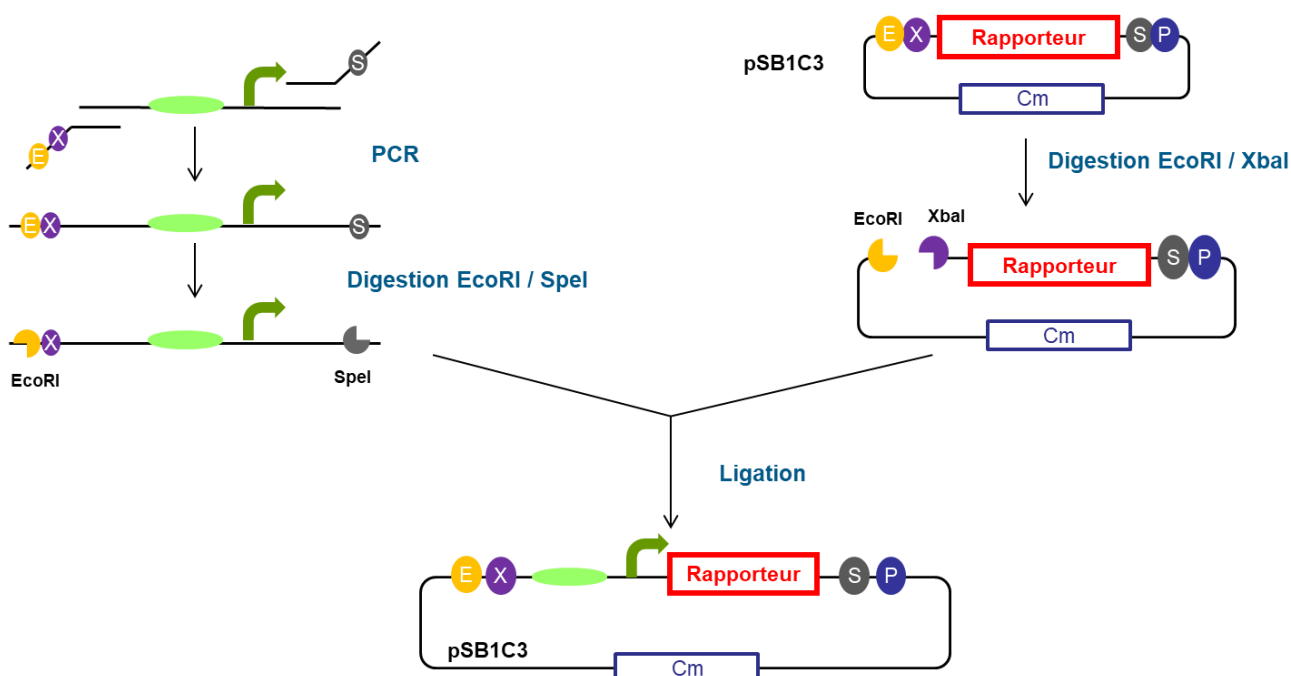


Figure 21: Méthode de clonage utilisée durant cette étude. Les sites EcoRI, XbaI, SpeI et PstI sont respectivement représentés par les lettres E, X, S et P. La stratégie consiste à amplifier les promoteurs par PCR à partir du génome d'*E. coli*. Lors de cette amplification, les sites E, X et S sont ajoutés grâce à des queues flottantes sur les amorces. Le produit PCR est digéré avec les enzymes EcoRI et SpeI et le vecteur pSB1C3, contenant le gène codant pour une protéine fluorescente, est digéré avec les enzymes EcoRI et XbaI. Le nouveau plasmide est assemblé par ligation grâce à la complémentarité des sites SpeI et XbaI. Cette ligation conserve le site EcoRI mais supprime les sites XbaI et SpeI. Ce système permet ainsi de retrouver, après ligation, les mêmes sites de part et d'autre que ceux présents avant l'insertion.

Les zones promotrices ont été amplifiées par PCR sur le génome de *E. coli* BW25113, avant d'être clonées dans le plasmide pSB1C3 en amont du gène codant pour une protéine fluorescente. Les amorces utilisées permettent une amplification de toute la zone régulatrice de l'expression du gène, soit 800 pb en amont du codon d'initiation de la traduction. Elles possèdent une queue flottante permettant l'insertion des sites de restriction EcoRI et XbaI en 5' et du site de restriction SpeI en 3'. Les amorces ont été dessinées à partir des séquences obtenues sur la base de données Ecocyc (Karp *et al.*, 2018). Elles ont été vérifiées par Blast sur le génome d'*E. coli* K12 BW25113 puis commandées (Eurofins, France). L'amplification des inserts a été réalisée par PCR dans un volume final de 50µL, contenant 0,5 µL de polymérase Q5 à 2000 unités/ml (NEB, USA), 1µL de dNTP à 2,5mM, 2µL de matrice d'ADN génomique à 25ng/µL (*E. coli*, BW25113), 10 µL de Tampon Q5 5X (NEB, USA), 31,5 µL d'eau milliQ filtrée et 2,5 µL de chaque amorce (forward et reverse) à 10 mM. Elle a été réalisée selon le programme suivant : 5 minutes à 98°C, puis 30 cycles de 98°C pendant 30 secondes, 58°C pendant 30 secondes et 72°C pendant 30 secondes, puis une étape d'élongation finale à 72°C pendant 5 minutes (MJ mini Personal Thermal Cycler, Biorad, USA). Le vecteur pSB1C3 a été digéré par les enzymes de restriction EcoRI et XbaI et chaque insert ont été digérés par les enzymes de restriction EcoRI et SpeI. Le vecteur et les inserts digérés ont été purifiés grâce au kit gel extraction (Qiagen, Allemagne), puis leurs concentrations ont été quantifiées au Nanodrop ® (ND-1000- spectrophotometer). Ils ont été assemblés par ligation avec la Quick Ligase (NEB, USA) avec un rapport molaire de 1:3 (vecteur : insert). La réaction a été effectuée pendant 1h à température ambiante. Le produit de la ligation est ensuite transformé dans la souche *E. coli* NEB 5-alpha (NEB, USA) et les colonies sont sélectionnées sur LB agar contenant 40 µg/mL de chloramphénicol. L'ADN plasmidique de 5 clones a ensuite été extrait en utilisant le kit QIAprep Spin Miniprep Kit (Qiagen) et vérifié par digestion enzymatique et par séquençage (Eurofins, France) afin de valider la construction.

Au total une quinzaine de plasmides ont été construits, certains promoteurs ont été clonés en amont de différentes protéines fluorescentes afin de réaliser des contrôles.

I. 3. Test d'expression et de spécificité des promoteurs

I. 3. A Test d'expression

Dans un premier temps les promoteurs théoriquement spécifiques de l'utilisation d'un substrat ont été ajoutés dans les châssis pSB1C3 contenant le gène codant pour la mRFP1, donc permettant l'expression de fluorescence rouge. Les promoteurs constitutifs ont été clonés en amont du gène codant pour la sYFP, permettant l'expression de fluorescence verte. Une fois les plasmides construits avec la méthodologie présentée sur la figure 21, la production de la protéine fluorescente a été analysée de manière qualitative, par étalement sur boîte des différentes constructions et visualisation de la fluorescence sur plaque UV (Figure 22).

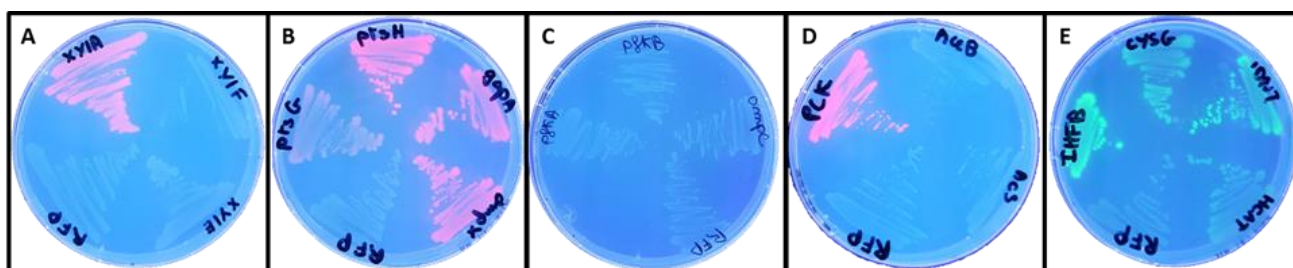


Figure 22 : Test d'expression des constructions sur boîtes. A) Etalement des souches *PxylAmRFP1*, *PxylF mRFP1*, *PxylEmRFP1* et *mRFP1* (Témoin négatif : souche contenant le plasmide avec le gène codant pour la protéine fluorescente mais sans promoteur) sur boîtes M9 (milieu minimum) avec du xylose comme seule source carbonée. B) Etalement des souches *PptsG mRFP1*, *PptsH mRFP1*, *PgapAmRFP1*, *ompX mRFP1* et *mRFP1* sur M9 glucose. C) Etalement des souches *PpfkA mRFP1*, *PpfkB mRFP1*, *PompC mRFP1* et *mRFP1* sur M9 glucose. D) *Ppck mRFP1*, *PaceB mRFP1*, *Pacs mRFP1* et *mRFP1* sur M9 acétate. E) *PihfB-sYFP*, *PcysG-sYFP*, *PidnT-sYFP*, *PhcaT-sYFP* et *mRFP1* sur M9 glucose. Photos réalisées sur table à UV.

Ce premier test préliminaire a permis de s'assurer que notre choix expérimental de construction permet bien de produire les protéines fluorescentes d'un point de vue qualitatif. Cependant, il est déjà possible d'apercevoir des différences de niveau d'expression entre les constructions, provenant très probablement des différences de force d'expression des promoteurs. Cette méthodologie n'étant pas quantitative, les souches ont été par la suite analysées en cytométrie en flux afin de quantifier leur niveau d'expression de fluorescence.

I.3.B Test de spécificité sur xylose

Les différentes souches ont été analysées en cytométrie en flux dans le but d'obtenir des données quantitatives sur le niveau d'expression des différents promoteurs, mais aussi de tester la spécificité d'expression lors de la consommation de différentes sources de carbone.

Les analyses de cytométrie ont été effectuées sur un cytomètre MASCQuant VYB (Miltenyi Biotec, Allemagne). Ce cytomètre est équipé d'un laser jaune (561 nm) permettant l'excitation de la mRFP1, d'un laser violet (405 nm) permettant l'excitation de la mTagBFP et d'un laser bleu (488 nm) permettant l'excitation de la sYFP2. L'acquisition des données a été fixée à 40 000 événements par échantillon. Le traitement des données a été réalisé avec le logiciel FlowJo X. Une fenêtre (« gate ») de sélection a été créée dans le diagramme de point du forward scatter channel (FSC-H) par rapport au side scatter channel (SSC-H) pour distinguer les cellules du bruit technique de l'appareil. Une deuxième sélection a été créée dans le diagramme de points du SSC-H par rapport au SSC-A pour sélectionner des cellules individuelles. Les émissions de fluorescence mTagBFP ont été analysées avec le canal V1-H (BP450 / 25 nm), les émissions mRFP1 avec le canal Y2-H (BP615 / 10 nm) et les émissions sYFP2 avec le canal B1-H (B1: 525/50 nm).

Pour les tests de spécificité, les différentes souches ont été cultivées sur milieu minimum M9 xylose, glucose ou acétate. Les cultures ont été réalisées dans un volume de 25 mL de milieu M9, en erlenmeyer de 100 mL, à 37°C avec agitation (150 rpm). Elles ont étéensemencées à $DO_{600nm} = 0,05$ à partir d'une pré-culture sur la nuit également réalisée en milieu M9. Les échantillons sont prélevés à $DO_{600nm} \approx 1$ en phase exponentielle de croissance. Les échantillons sont alors immédiatement congelés à -20°C. Après décongélation, ils sont dilués dans du NaCl à 0,9% filtré pour obtenir une concentration cellulaire de 1×10^6 cellules par mL. Les échantillons de cultures ont été analysés en cytométrie de flux. Les résultats pour les souches supposées spécifiques du xylose sont présentés en figure 23.

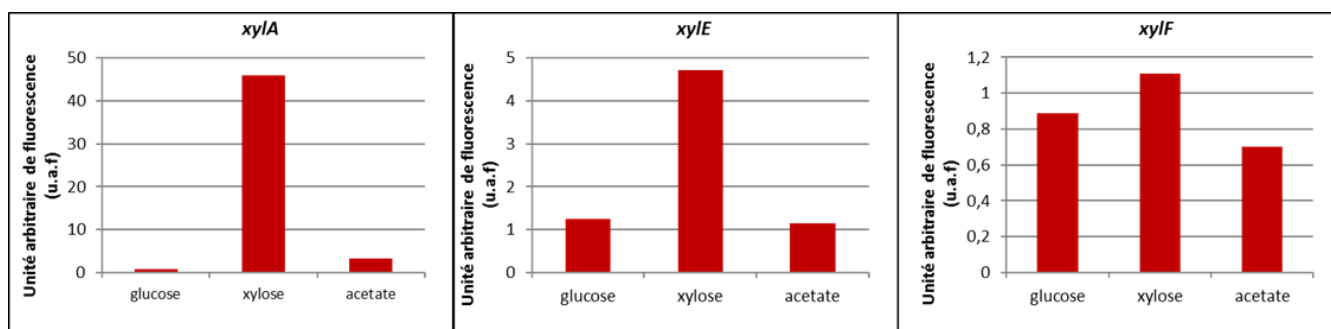


Figure 23 : Tests de spécificité pour les rapporteurs du xylose. Les souches ont été cultivées sur milieu minimum additionné d'un des trois substrats, puis prélevées en phase exponentielle de croissance et analysées au cytomètre de flux. Les éléments détectés ont été analysés en fonction de leur taille et de leur granulométrie, afin de n'étudier que les cellules de *E. coli*. Une analyse des singulets a été réalisée pour ne pas prendre en compte les agrégats de cellules. Les figures présentent l'intensité de fluorescence médiane des cellules sur le canal Y2-A (n=1).

Les trois souches *PxylAmRFP1*, *PxylEmRFP1* et *PxylFmRFP1* présentent respectivement une expression sur xylose de 45,91, de 4,71 et de 1,11 unité arbitraire de fluorescence (u.a.f). Ainsi les souches *PxylEmRFP1* et *PxylFmRFP1* expriment la protéine mRFP1, mais avec un niveau d'expression au moins dix fois plus faible que la souche *PxylAmRFP1*. Ces résultats n'étaient pas visibles sur boîte de pétri. De plus, il est possible de remarquer qu'il y a bien une meilleure expression par ces promoteurs lors de la consommation de xylose que sur glucose ou acétate, surtout dans le cas des promoteurs *xylA* et *xylE*. En effet, pour le promoteur *xylA*, les expressions sur glucose, xylose et acétate sont respectivement de 0,79, de 45,91 et de 3,19 u.a.f. L'expression de la fluorescence est donc 40 fois supérieure sur xylose que sur glucose et plus de 10 fois supérieure sur xylose que sur acétate. Pour la souche *PxylEmRFP1*, la fluorescence est de 1,25 u.a.f. sur glucose, de 4,71 sur xylose et de 1,15 sur acétate. Elle est donc 4 fois supérieure sur xylose que sur les autres substrats. La souche *PxylFmRFP1* présente un niveau d'expression trop faible et la différence de niveau d'expression entre les substrats n'est pas suffisante, elle n'a donc pas été gardée pour la suite de l'étude. Nous avons gardé le promoteur *P_{xylA}* pour marquer l'utilisation du xylose.

I. 4. C Test de spécificité sur glucose

Les plasmides construits pour avoir une expression fluorescente dépendante de l'utilisation du glucose ont été testés de la même manière que les constructions présumées spécifiques du xylose.

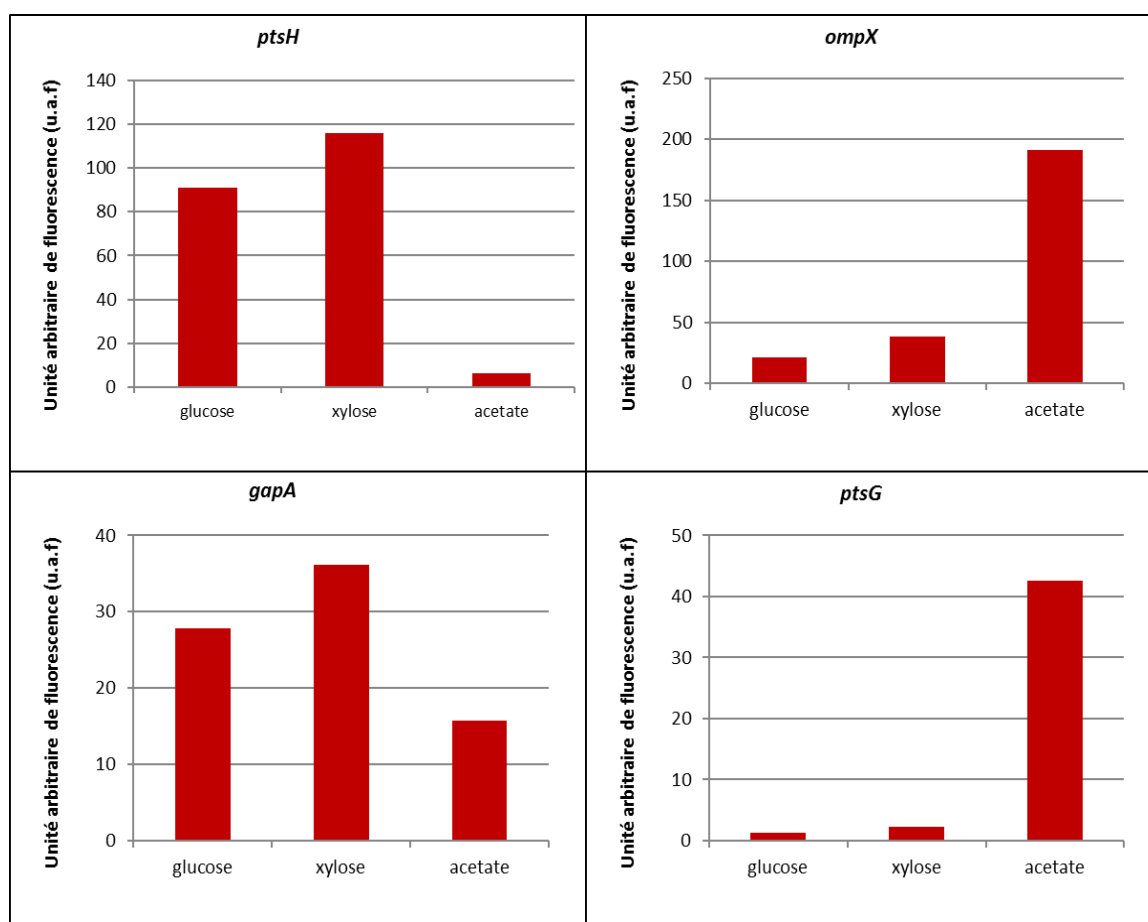


Figure 24 : Tests de spécificité pour les rapporteurs du glucose Les souches ont été cultivées sur milieu minimum additionné d'un des trois substrats, puis prélevées en phase exponentielle de croissance et analysées au cytomètre de flux. Les éléments détectés ont été analysés en fonction de leur taille et de leur granulométrie afin de n'étudier que les cellules de *E. coli*. Une analyse des singulets a été réalisée pour ne pas prendre en compte les agrégats de cellules. Les figures présentent l'intensité de fluorescence médiane des cellules sur le canal Y2-A (n=1).

Aucune des souches construites pour marquer le métabolisme du glucose ne présentent de spécificité d'expression réelle pour ce substrat. Ainsi de manière surprenante les souches *PompX*-mRFP1 et *PptsG*-mRFP1 fluorescent plus fortement sur milieu M9-acétate. La souche *PgapA*-mRFP1 exprime une fluorescence sur les trois

substrats et la souche *PptsH*-mRFP1 a un fort niveau de fluorescence sur xylose et glucose. Cependant, cette dernière souche possède un potentiel d'utilisation dans le cas d'étude de mélange glucose-acétate car son niveau de fluorescence sur acétate est très faible, soit 6,2 u.a.f par rapport à 90,9 u.a.f sur glucose, donc un niveau d'expression sur glucose environ 15 fois supérieur à celui sur acétate.

I. 4. D Test de spécificité sur acétate

Comme attendu, les trois souches *Pacs*-mRFP1, *PaceB*-mRFP1, *Ppck*-mRFP1 présentent une expression de fluorescence plus forte sur acétate que sur les deux autres substrats. La souche *PglgS*-mRFP1 ne présente pas de spécificité pour l'acétate, de plus son niveau d'expression est très faible, d'environ 1 u.a de fluorescence pour les trois substrats. Cette construction n'a donc pas été retenue pour le reste de l'étude. Le promoteur du gène *pck* présente un intérêt par rapport aux deux autres promoteurs *acs* et *ace*, en raison de son fort niveau d'expression, potentiellement avantageux lors d'insertions chromosomiques par exemple. Les constructions contenant les promoteurs *acs* et *aceB* expriment un niveau de fluorescence faible, mais cependant plus élevé sur acétate.

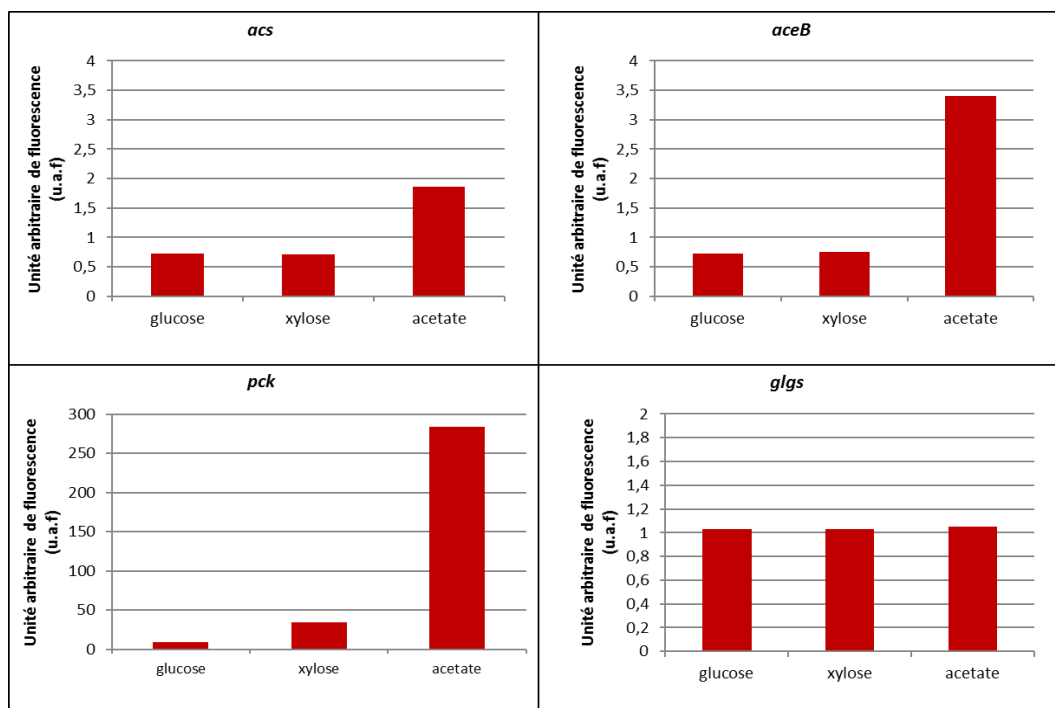


Figure 25 : Tests de spécificité pour les rapporteurs de l'acétate. Les souches ont été cultivées sur milieu minimum additionné d'un des trois substrats, puis prélevées en phase exponentielle de croissance et analysées au cytomètre de flux. Les éléments détectés ont été analysés en fonction de leur taille et de leur granulométrie, afin de n'étudier que les cellules de *E. coli*. Une analyse des singulets a été réalisée pour ne pas prendre en compte les agrégats de cellules. Les figures présentent l'intensité de fluorescence médiane des cellules sur le canal Y2-A.

En conclusion les expériences de cytométrie de flux ont permis de valider ou non la spécificité de différents promoteurs lors de croissance sur glucose, xylose et acétate. Ainsi, les promoteurs *xyIA* et *xyIE* se révèlent être de bon outils pour étudier l'utilisation du xylose. Nous avons privilégié l'utilisation du promoteur *xyIA* de par son niveau d'expression plus élevé. Les promoteurs *pck*, *aceB* et *acs* pourront être utilisés pour la croissance sur acétate. De plus, le promoteur *ptsH* peut être intéressant pour suivre l'utilisation du glucose dans le cas d'une croissance en mélange glucose et acétate. Ils ne seront pas utilisés dans le cadre de cette thèse mais présentent un intérêt pour les perspectives du projet.

I. 5. Test d'impact des marqueurs fluorescents lors de la croissance

L'expression de protéines fluorescentes pourrait impacter le métabolisme de la cellule et notamment induire potentiellement un stress oxydatif (Ganini *et al.*, 2017). Afin de déterminer si l'expression de protéines fluorescentes a un impact sur la croissance et si cet impact est dépendant de la nature de la protéine fluorescente, des tests préliminaires ont été réalisés. Nous avons choisi de comparer des croissances avec l'expression de différentes protéines fluorescentes (mRFP1, mTagBFP et sYFP), sous contrôle du même promoteur constitutif *ihfB*. Nous avons aussi utilisé comme contrôle la souche contenant le plasmide vide, c'est à dire le plasmide pSB1C3_mRFP1 qui contient le gène codant pour la mRFP1, mais sans promoteur en amont du gène. Il est utilisé afin de déterminer l'impact du plasmide sans production de protéine fluorescente sur le taux de croissance. Trois plasmides construits sur le même modèle ont été utilisés, ils sont issus du châssis pSB1C3 et contiennent le promoteur *ihfB* natif d'*E. coli* BW25113 en amont des gènes codants pour la mRFP1, mTagBFP ou la sYFP (Figure 26). Les conditions de culture sont identiques entre le contrôle et les plasmides exprimant la fluorescence.

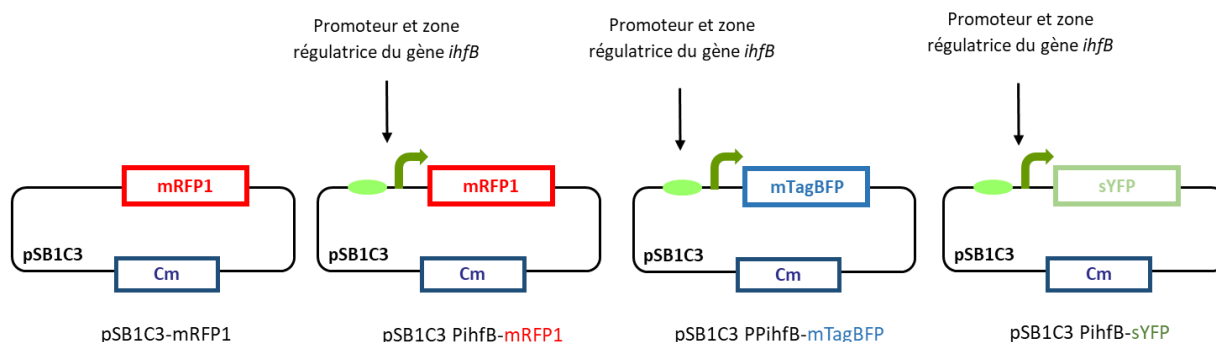


Figure 26 : Schéma des plasmides utilisés pour mesurer l'impact des protéines fluorescentes sur la croissance.

Les 4 plasmides ont été insérés dans la souche *E. coli* BW25113. Les cellules ont été pré-cultivées pendant une nuit dans des tubes de 5 ml avec du milieu M9 glucose à 37°C, avec agitation à 150 rpm. Les pré-cultures ont été centrifugées à 5000 rpm pendant 3 min et les culots ont été remis en suspension dans du M9 glucose ou du M9 xylose, à une $DO_{600\text{ nm}}$ initiale de 0,01. Les microplaques 96 puits ont été inoculées en triplicat avec 200

μL de chaque re-suspension, ou de milieu non ensemencé en contrôle négatif. La cinétique est réalisée dans le lecteur de microplaques Versamax (Molecular Devices, USA), à 37°C avec agitation. La $\text{DO}_{600\text{nm}}$ a été lue toute les 17 min pendant 24 h. Les taux de croissance maximum (μ_{max}) ont été déterminés pendant la phase exponentielle de croissance en utilisant le coefficient de régression linéaire de la courbe représentant $\ln(\text{DO}_{600\text{nm}})$ en fonction du temps.

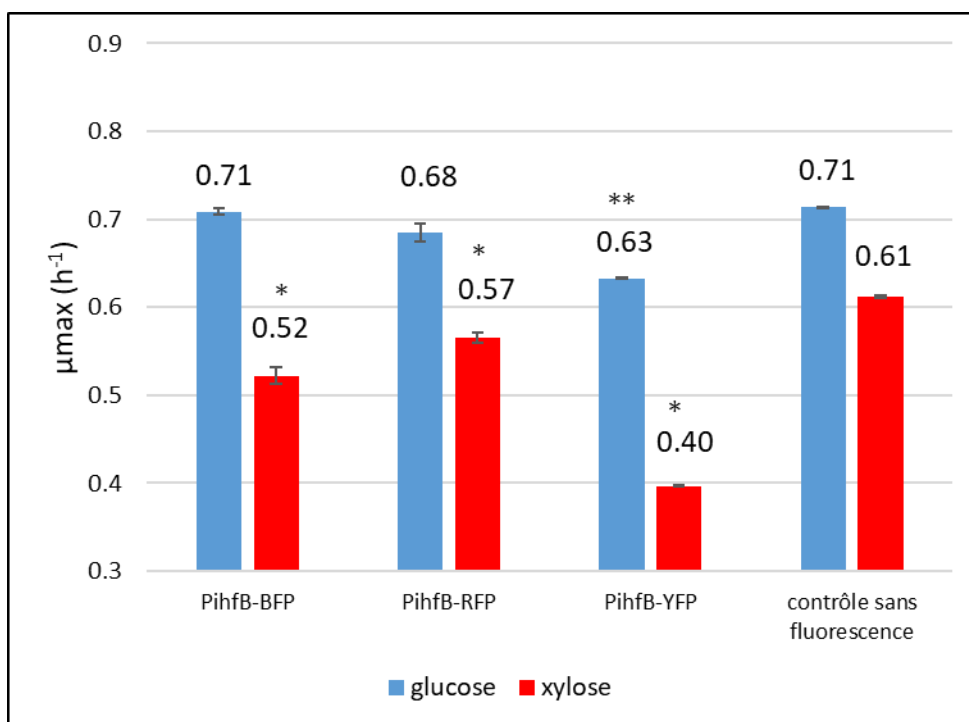


Figure 27 : Influence du marqueur fluorescent sur le taux de croissance, le μ_{max} moyen calculé lors de croissance en microplaque est représenté, lors de croissance sur glucose en bleu et lors de croissance sur xylose en rouge ($n=3$). Les différences statistiques ont été analysées par test de Student. Pour les souches contenant les constructions $P_{ihfB}\text{mTagBFP}$ et $P_{ihfB}\text{mRFP1}$ en croissance sur glucose (en bleu), il n'y a pas de différence significative avec la référence pSB1C3 ($p\text{-value}>0.05$), à la différence de la souche $P_{ihfB}\text{YFP}$ ($p\text{-value}<0.05$) indiqué par **. Lors de croissance sur xylose (en rouge), les taux de croissances sont significativement différents pour les trois souches exprimant des protéines fluorescentes en comparaison à la référence pSB1C3 ($p\text{-value}<0,05$) indiqué par *.

Les taux de croissance maximum (μ_{max}) ont été calculés en phase exponentielle de croissance pour chaque souche sur glucose et sur xylose, les résultats sont présentés en figure 27. L'expression des protéines mTagBFP (BFP) et mRFP1 (RFP) n'impacte pas la

croissance sur glucose (p-value > 0,05). Le taux de croissance semble légèrement plus bas sur xylose quand ces protéines sont constitutivement produites. La production de la protéine sYFP2 (YFP) semble avoir un fort impact sur le taux de croissance de la souche sur xylose (p-value < 0.05). Ainsi, le taux de croissance du contrôle sur xylose est de 0.61 h⁻¹ alors qu'il est de 0.40 h⁻¹ chez la souche contenant le plasmide pSB1C3-P_{ihfB}-YFP. Le marqueur fluorescent sYFP n'a donc pas été retenu pour la suite de l'étude. L'expression des protéines mRFP1 et mTagBFP ne semble pas avoir d'impact majeur sur la croissance de la souche, cependant le taux de croissance est légèrement ralenti mais de manière significative lorsque les deux protéines sont produites lors de croissance sur xylose, la surcharge métabolique et le coût énergétique nécessaire à la production des deux molécules semble donc impacter légèrement la production de biomasse.

II. Création du plasmide d'étude portant une double fluorescence

Les tests précédemment réalisés nous ont permis de quantifier l'impact de la production de fluorescence, et de définir les promoteurs et rapporteurs fluorescents les plus appropriés permettant l'expression d'une fluorescence constitutive et spécifique à l'utilisation du xylose. Les cellules bactériennes sont de petites tailles (entre 1 et 2 μm chez *E. coli*). De fait, les signaux FSC et SSC en cytométrie permettant respectivement la détection de la taille et de la granulométrie ne permettent pas de distinguer de façon absolue la population du bruit de fond de l'appareil. Ainsi, le but d'avoir cette double fluorescence est de réaliser un contrôle interne grâce au rapporteur constitutif. L'expression de la seconde fluorescence constitutive permet de distinguer les cellules du bruit de fond lors de l'analyse en cytométrie. De plus, le nombre de copie du plasmide peut varier d'une cellule à l'autre au sein de la population (Jahn *et al.*, 2014). La présence de la fluorescence bleue pourrait servir de contrôle permettant de normaliser le nombre de copies du plasmide au sein de la cellule, en plus de servir de témoin du maintien du plasmide dans la cellule.

Ainsi, la seconde phase a été de construire un plasmide présentant une double fluorescence constitutive et spécifique. Nous avons choisi d'utiliser les protéines mTagBFP et mRFP1 car elles impactent peu la croissance (cf partie I.5). Ces deux protéines présentent de plus des longueurs d'ondes d'excitation et d'émission éloignées, limitant ainsi les risques d'interférences. Ce plasmide porte la mTagBFP sous contrôle du promoteur constitutif *ihfB* et la mRFP1 sous contrôle du promoteur *xylA* inductible au xylose, tel que présenté dans la figure 28. Il a été construit par assemblage des plasmides pSB1C3 *PxylA*-mRFP1 et pSB1C3-*PihfB*-mTagBFP, selon la méthodologie iGEM, présentée en figure 21. Ce plasmide portant la double fluorescence a été inséré dans la souche *E. coli* BW25113.

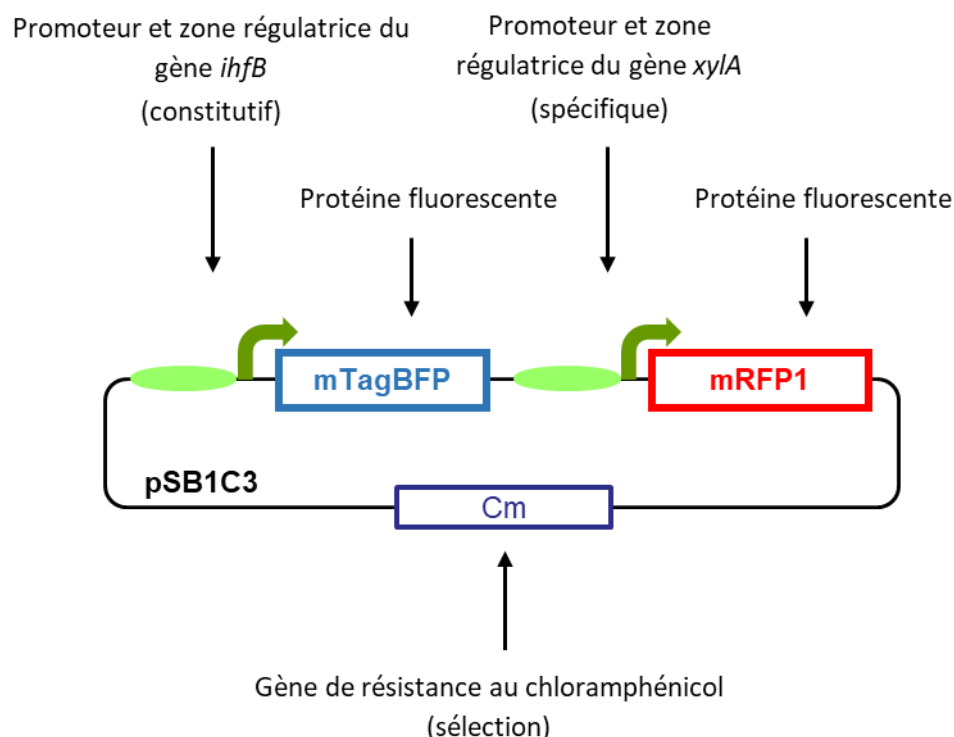


Figure 28 : Schéma de la construction plasmidique portant une double fluorescence. L'expression de la fluorescence rouge est soumise à un promoteur spécifique de l'utilisation du xylose, l'expression de la fluorescence bleue est soumise à un promoteur constitutif.

La mesure du taux de croissance de la construction présentant la double fluorescence a été réalisée sur microplaque (de la même manière que précédemment), afin de mesurer si l'expression conjointe de la fluorescence bleue et de la fluorescence rouge impacte la croissance. On peut remarquer que le taux de croissance sur glucose est impacté plus fortement que lorsqu'une seule fluorescence est produite (Figure 29). Le taux de croissance sur glucose pour la souche n'exprimant pas de fluorescence est de $0.71 \text{ h}^{-1} \pm 0.01$ et pour la souche avec la fluorescence rouge inductible au xylose de $0.77 \text{ h}^{-1} \pm 0.01$, alors que le taux de croissance de la souche exprimant la double fluorescence est de $0.65 \text{ h}^{-1} \pm 0.01$ sur glucose. Cette baisse du taux de croissance est possiblement liée au coût énergétique de la production des deux fluorophores. Cependant le taux de croissance sur xylose ne semble pas impacté entre la souche produisant une seule fluorescence et la double fluorescence (respectivement $\mu_{\text{max}} = 0.53 \text{ h}^{-1} \pm 0.02$ et $\mu_{\text{max}} = 0.52 \text{ h}^{-1} \pm 0.01$). Le taux de croissance étant plus faible sur xylose, il est peut-être moins impacté par la charge métabolique supplémentaire due à la production des protéines fluorescentes.

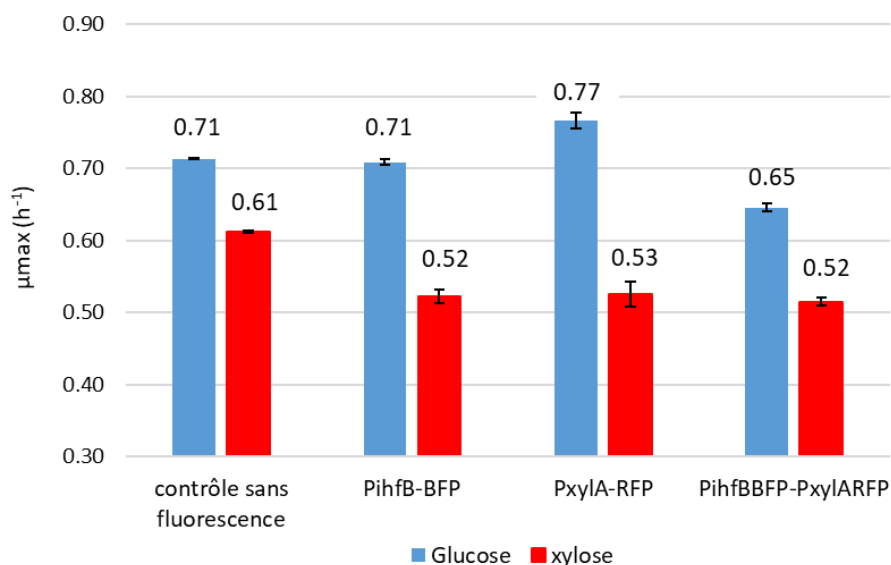


Figure 29 : Impact de la double construction sur le taux de croissance. Le $\mu_{\max}$ moyen calculé lors de croissance en microplaque est représenté, lors de croissance sur glucose en bleu et lors de croissance sur xylose en rouge (n=3).

Nous avons par la suite cherché à tester les propriétés de la souche en terme de production de fluorescence. Pour cela, la souche a été cultivée sur milieu minimum M9 avec glucose ou xylose comme source de carbone et des échantillons ont été prélevés en phase exponentielle de croissance à $DO_{600nm} = 1$. Les échantillons de cultures ont été analysés en cytométrie de flux avec un cytomètre MASCQuant VYB (Miltenyi, Allemagne). Les données ont été traitées à l'aide du logiciel FlowJo (BectonDickinson, USA). Elles sont présentées en figure 30.

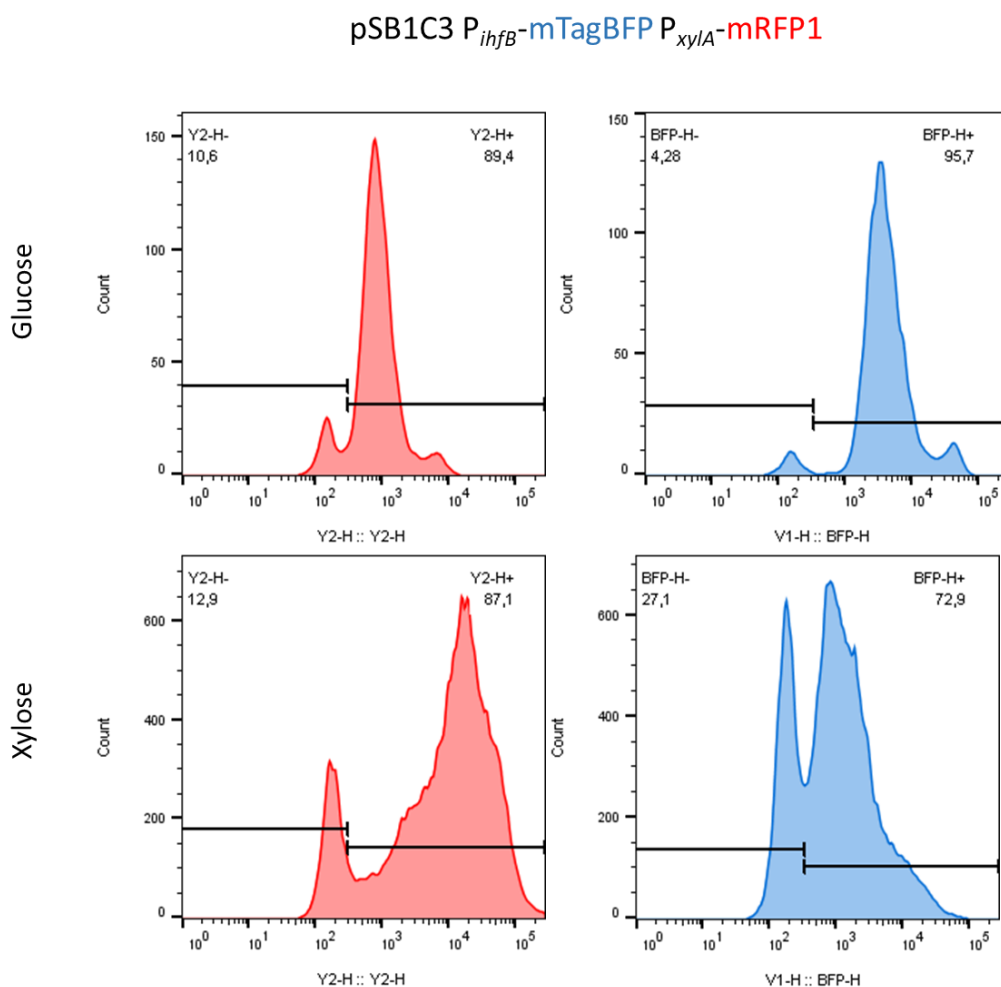


Figure 30 : Caractéristiques du signal fluorescent de la souche contenant le plasmide pSB1C3 P_{ihfB} -mTagBFP P_{xylA} -mRFP1 sur les canaux Y2-H (rouge) et V1-H (bleu) en cytométrie en flux, lors de croissance en glucose ou en xylose. Les barres noires présentent le niveau négatif de fluorescence (Y2-H -) contre le niveau positif de fluorescence (Y2-H +) fixé à l'aide de cellules d'*Escherichia coli* non fluorescentes.

La souche présente bien un niveau de fluorescence dans les deux canaux Y2 (rouge) et V1 (bleu), ainsi les deux protéines fluorescentes sont exprimées. De plus, nous avons vérifié qu'il n'y a pas d'interférence entre l'expression des deux protéines fluorescentes. Il est impossible en effet d'observer une fluorescence sur le canal Y2 lorsque seule la protéine mTagBFP est exprimée. De même, il est impossible d'observer une fluorescence sur le canal V1 lorsque la mRFP1 est exprimée seule (données non montrées). Il n'y a donc pas de fuite de fluorescence entre les différents canaux d'analyse. L'analyse du niveau de fluorescence en rouge de la souche portant ce plasmide montre que l'intensité médiane de fluorescence lors de croissance sur xylose est de 2×10^4 u.a.f soit 20 fois supérieure au niveau de fluorescence lors de la croissance sur glucose (1×10^3

u.a.f) (Figure 30). Cependant, le niveau de fluorescence rouge sur glucose n'est pas égal au niveau d'auto-fluorescence de la souche sauvage, qui a servi à fixer la population non fluorescente. De plus ce niveau de fluorescence rouge sur glucose est supérieur au niveau observé sur glucose avec le plasmide portant uniquement la fluorescence rouge sous contrôle du P_{xyIA} . Ce phénomène semble être dû à l'expression de la fluorescence bleue : en effet il est possible que le promoteur *ihfB* induise une fuite d'expression de la protéine mRFP1, les deux gènes étant positionnés dans le même sens d'expression (Figure 28). En ce qui concerne le signal fluorescent bleu, lors de la croissance sur glucose, le niveau médian de fluorescence est de $3,5 \times 10^3$ u.a.f, alors que le niveau de fluorescence bleu sur xylose est de $1,5 \times 10^3$ u.a.f. Il est aussi possible de remarquer, pour les deux fluorescences que les pics sont plus étalés lors de croissance sur xylose, ainsi la variabilité du signal est plus grande (Figure 30). Ces variations des niveaux de fluorescence pourraient être liée à une diminution du nombre de copie du plasmide sur xylose.

En regardant en parallèle le signal de fluorescence rouge (Y2-H) en abscisse et le signal de fluorescence bleue (V1-H) en ordonné, nous pouvons voir que les profils de fluorescences, représentés par les lassos noirs, des cellules en croissance sur glucose et celles en croissances sur xylose sont différents et ne s'entrecoupent pas (Figure 31).

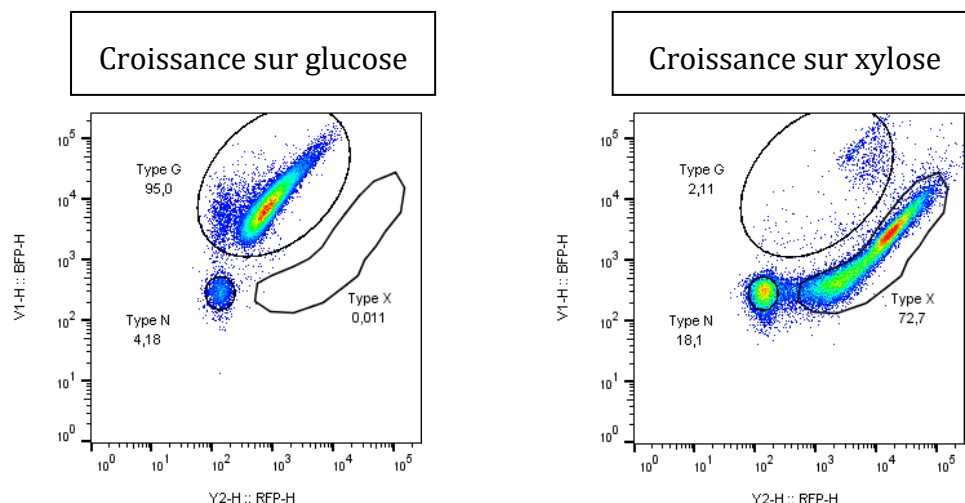


Figure 31 : Profils en nuage de point de la population lors de croissance sur M9 glucose ou M9 xylose avec le signal fluorescent V1-H en ordonné et le signal fluorescent Y2-H en abscisse. Le lasso noir type G représente le profil de fluorescence des cellules en croissance sur glucose, le lasso noir type X celui des cellules en croissances sur xylose et le lasso noir type N des cellules non fluorescentes.

Il est possible de voir des cellules présentant un signal non fluorescent. La difficulté ici tient à la nature de cette « sous-population ». Nous avons essayé plusieurs expériences et avons découvert qu'elles avaient en fait des origines différentes. Tout d'abord, certains événements sont dus à des bruits électroniques et de tampons. En effet, la solution de NaCl filtrée et l'eau filtrée fournissent des signaux (FSC, SSC) similaires aux signaux émis par des cellules d'*Escherichia coli*, il n'est pas possible de supprimer totalement le bruit de fond. Deuxièmement, la granulométrie (signaux SSC) et la taille (signaux FSC) de la « population » non fluorescente évoluent d'une condition à l'autre. La « population » non fluorescente sur le glucose semble fournir un signal SSC et FSC proche des bruits électroniques et tampons. Les signaux sur xylose semblent être un mélange de cellules réelles et de bruits puisqu'ils augmentent avec le temps, contrairement aux signaux sur glucose. Ce comportement pourrait être lié à des débris cellulaires cumulés pendant la croissance. Cela a réconforté l'utilisation du marqueur constitutif bleu pour distinguer efficacement les cellules des artefacts. En conclusion, nous pensons que cette population non fluorescente est principalement constituée de débris cellulaires et non cellulaires, expliquant ainsi leur non-fluorescence. Il aurait été beaucoup plus facile de les supprimer de l'étude, mais cela aurait été inexact, étant donné qu'ils peuvent compter dans la DO600nm.

Ainsi cette souche présente un fort potentiel pour l'étude de l'utilisation différentielle des substrats. En effet la signature fluorescente des cellules lors de croissance sur glucose ou sur xylose est très différente et permet de les différencier sans ambiguïté. De plus, le signal fluorescent bleu peut être utilisé en normalisateur pour palier à la variation possible du nombre de copie du plasmide. Malgré des améliorations possibles du système, le plasmide pSB1C3 P_{ihfB} -mTagBFP P_{xylA} -mRFP1 a été utilisé pour produire une grande partie des résultats de cette étude car il possède l'avantage de discriminer efficacement les cellules consommant le glucose de celle consommant le xylose (Figure 31).

Chapitre 2 :

La disponibilité du facteur de transcription XylR influe sur l'hétérogénéité phénotypique et le temps de transition d'*Escherichia coli* lors de l'adaptation du glucose au xylose

Dans le chapitre précédent nous avons pu valider la construction permettant l'expression d'une protéine fluorescente en réponse à l'utilisation du xylose et d'une protéine fluorescente bleue de manière constitutive. Dans cette partie, nous allons aborder l'étude approfondie de la transition glucose/xylose en utilisant la souche précédemment construite. L'objectif initial est d'analyser si le comportement des cellules est de nature homogène lors de la transition du glucose au xylose, ou s'il y a présence de sous-populations, avec par exemple des cellules capables de consommer le xylose et d'autres non. Nos résultats nous ont conduit à étudier l'importance du facteur de transcription XylR lors de cette transition.

Très peu d'informations sont disponibles sur la transition glucose/xylose dans la littérature. Nous avons donc cherché à éclaircir certaines de ces caractéristiques. Dans un premier temps, nous nous sommes intéressés à la durée de cette transition et à sa variabilité entre les souches d'*Escherichia coli* transférées d'un milieu contenant du glucose, à un milieu contenant du xylose. Nous avons ainsi pu constater que ce temps de latence est fortement variable d'une souche à une autre.

Cette variabilité nous a conduit à étudier si cette durée peut être modulée (allongée ou réduite) et quels sont les régulateurs qui peuvent l'impacter. À l'aide de constructions, nous avons fait varier la disponibilité en XylR à l'intérieur de la cellule par titration du facteur de transcription XylR, grâce à l'augmentation du nombre de sites de fixations ou à la surexpression de XylR. Nous avons pu constater que la disponibilité de XylR impacte ce temps de latence, sans avoir d'autre effet sur la vitesse de croissance sur xylose ou glucose.

Au niveau de la cellule unique, chez la souche portant le plasmide pSB1C3 P_{xylA} -mRFP1 P_{ihfB} -mTagBFP, une forte hétérogénéité métabolique a été observée lors du passage du glucose au xylose, avec une sous-population incapable de reprendre une croissance exponentielle alors que l'autre sous-population augmentait de manière exponentielle sur le xylose.

Toutefois, cette forte hétérogénéité résulte de la présence du plasmide. Pour extrapoler les résultats vers le comportement de la souche sauvage, nous avons construit un modèle stochastique, basé sur l'hypothèse que la disponibilité en XylR influence la probabilité que des cellules individuelles s'adaptent à la croissance du xylose. Le modèle a été utilisé pour comprendre comment XylR peut jouer le rôle d'interrupteur moléculaire

et induit un état bistable dans la population en présence de notre construction. Le modèle a aussi permis de prédire le comportement de la souche sauvage, illustrant la présence théorique de cellules non-adaptées, bien que rares.

Ce travail montre que la durée des phases de latence chez *E. coli* est contrôlable et renforce le rôle des mécanismes stochastiques dans l'adaptation cellulaire. Il a donné lieu à une publication dans la revue mBio. L'article a été publié le 22 décembre 2020.

Le texte intégral est retranscrit ici accompagné des figures supplémentaires.

Availability of the molecular switch XylR controls phenotypic heterogeneity and lag duration during *Escherichia coli* adaptation from glucose to xylose

Manon BARTHE^a, Josué TCHOUANTI^b, Pedro Henrique GOMES^a, Carine BIDEAUX^a, Delphine LESTRADE^c, Carl GRAHAM^{b-d}, Jean-Philippe STEYER^e, Sylvie MELEARD^{b-f}, Jérôme HARMAND^e, Nathalie GORRET^a, Muriel COCAIGN-BOUSQUET^{a#†}, and Brice ENJALBERT^{a†}

- a. TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France
- b. CMAP, CNRS, Ecole polytechnique, IP Paris, Palaiseau France.
- c. TWB, Université de Toulouse, CNRS, INRA, INSA, Ramonville-Saint-Agne, France.
- d. Inria
- e. INRAE, Univ Montpellier, LBE, Narbonne, France.
- f. Institut Universitaire de France

Author to whom correspondence should be addressed: Muriel Coccain-Bousquet. E-Mail: cocaign@insa-toulouse.fr

† These authors contributed equally to this work.

Abstract:

The glucose-xylose metabolic transition is of growing interest as a model to explore cellular adaption since these molecules are the main substrates resulting from the deconstruction of lignocellulosic biomass. Here, we investigated the role of the XylR transcription factor in the length of the lag phases when the bacterium *Escherichia coli* needs to adapt from glucose to xylose-based growth. First, a variety of lag times were observed when different strains of *E. coli* were switched from glucose to xylose. These lag times were shown to be controlled by XylR availability in the cells with no further effect on the growth rate on xylose. XylR titration provoked long lag times demonstrated to result from phenotypic heterogeneity during the switch from glucose to xylose, with a sub-population unable to resume exponential growth whereas the other sub-population grew exponentially on xylose. A stochastic model was then constructed based on the assumption that XylR availability influences the probability of individual cells to switch to xylose growth. The model was used to understand how XylR behaves as a molecular switch determining the bistability set-up. This work shows that the length of lag phases in *E. coli* is controllable and reinforces the role of stochastic mechanism in cellular adaptation, paving the way for new strategies for the better use of sustainable carbon sources in bioeconomy.

Importance:

For decennia, it was thought that the lags observed when microorganisms switch from one substrate to another are inherent to the time required to adapt the molecular machinery to the new substrate. Here, the lag duration was found out to be the time necessary for a subpopulation of adapted cells to emerge and to become the main population. By identifying the molecular mechanism controlling the subpopulation emergence, we were able to extend or reduce the duration of the lags. This work is of special importance since it demonstrates the unexpected complexity of monoclonal populations during growth on mix substrates and provides novel mechanistic insights with regards to bacterial cellular adaptation.

Keywords: Adaptation / heterogeneity / metabolic transition / *Escherichia coli* / Sub-populations.

Introduction

Microorganisms live in a competitive and fluctuating environment, with frequent changes in the nature and availability of nutrients, thus requiring efficient adaptation if they are to continue to grow and survive. Such changes trigger regrowth lag phases defined as the time needed to resume growth, either after starvation when a substrate again becomes available (hereby named regrowth lag), or when the cells switch from one substrate to another (hereby named diauxic lag) (Monod, 1949; Madar *et al.*, 2013; Jöers and Tenson, 2016; Chu and Barnes, 2016). The latter occurs when bacteria are faced with mixed carbon sources. In such case, they generally use the carbon sources sequentially, starting from their preferred source (Monod, 1949; Aidelberg *et al.*, 2014; Ammar *et al.*, 2018). Use of the secondary substrate is repressed by the preferred one via a mechanism called carbon catabolite repression (Görke and Stülke, 2008). More particularly, when *Escherichia coli* cells are cultivated on glucose plus another substrate, their growth is described as diauxic, with exponential growth phases on each substrate interrupted by a diauxic lag phase of minimum or no growth, as first described by Jacques Monod (Monod, 1949). Sequential use is widely considered as an adaptation to rapidly consume the most beneficial substrate. This is supported by the fact that carbon sources allowing the higher growth rates are the first to be consumed (Aidelberg *et al.*, 2014).

Unravelling the mechanisms behind metabolic adaptation is a long-term effort of the microbiology community. For decades, the lag phases before resuming growth have been considered as an inevitable temporal constraint to produce the machinery required for growth on this substrate (Monod, 1949; Wolfe, 2005; Traxler *et al.*, 2006). The recent emergence of single cell technologies has undermined this assumption as it has revealed differential behaviors in cells belonging to the same population when they switch from one substrate to another (Boulineau *et al.*, 2013; Heerden *et al.*, 2014; Solopova *et al.*, 2014; Kotte *et al.*, 2014; Nikolic *et al.*, 2017). Not all the cells are subject to growth lag: bimodality could be observed, with a fraction of the population not able to grow on the new substrate. This was proposed as a possible explanation for lag phases. However, the mechanisms causing microbial populations to end up in metabolically different phenotypes remain mostly unknown (Kaiser *et al.*, 2018; Takhaviev and Heinemann, 2018; Bertrand, 2019). The molecular explanations are so far very different from one study to another (Takhaviev and Heinemann, 2018). A role of stringent response, cell-

dependent variations of intracellular contents, low concentration transcription factors, feedback regulatory circuits, catabolite repression, epigenetic-related regulatory mechanisms have been reported so far, depending on microorganisms and conditions (Ito *et al.*, 2009; Solopova *et al.*, 2014; Kotte *et al.*, 2014; Lambert and Kussell, 2014; Yaginuma *et al.*, 2014; Martins and Locke, 2015; Amato and Brynildsen, 2015; Axelrod *et al.*, 2015; Kaiser *et al.*, 2018; Dal Co *et al.*, 2018). Due to this profusion of explanations, laying the foundations for metabolic adaptation will require further studies to better understand lag phases and adaptation mechanisms.

The glucose-to-xylose transition is a highly interesting model to explore lag phases in *E. coli*. This transition is of economic interest since glucose and xylose are the two main compounds released during degradation of lignocellulose (one of the main renewable alternatives to the use of petrol) (Eiteman *et al.*, 2008; Kim *et al.*, 2015; Sievert *et al.*, 2017; Rossoni *et al.*, 2018). Xylose metabolism is simple, comprises only a few metabolic steps and is well described, but the mechanisms involved in the switch from glucose to xylose remain to be elucidated. Xylose uptake is mediated through an ABC transporter (encoded by *xylFGH*) or a symporter (XylE) (Sumiya *et al.*, 1995). Once in the cytoplasm, xylose is converted into xylulose by the xylose isomerase XylA (Wovcha *et al.*, 1983; Briggs *et al.*, 1984), and phosphorylated into xylulose-5-phosphate by the xylulokinase XylB (Lawlis *et al.*, 1984) before entering the pentose phosphate pathway. Xylose metabolism involves two operons (catabolism operon *xylAB* and transporter/regulatory operon *xylFGHR*) that share the same promoter region and are both positively controlled by the XylR transcription factor. When *E. coli* cells are grown in batches on a mixture of glucose and xylose, they use glucose first and catabolite repression prevents the expression of the two *xyl*-operons (Song and Park, 1997). Consequently, during growth on glucose, the concentration of XylR can be as low as one molecule per cell (Schmidt *et al.*, 2016). Once the glucose is exhausted, an increase in the concentration of intracellular cAMP leads to the formation of the CRP-cAMP complex, thereby alleviating the glucose-mediated carbon catabolite repression of the promoters (Deutscher, 2008; Görke and Stülke, 2008; Shimizu, 2013). Xylose then binds to and activates the transcription factor XylR, which, conjointly with the CRP-cAMP complex, triggers the transcription of the *xylAB* and *xylFGHR* operons (Song and Park, 1997; Laikova *et al.*, 2001). The binding of XylR-xylose to its own promoter triggers a positive loop that in turn, amplifies its own concentration up to 57 molecules per cell (Schmidt *et*

al., 2016). This triggers the production of the enzymatic machinery required for xylose consumption (Song and Park, 1997). Interestingly, a bimodal induction response to xylose was reported for the *xylAB* promoter with an all or nothing expression (Afroz *et al.*, 2014).

In this study, we used both experimental and modeling approaches to investigate the role of XylR as a molecular switch, defined here as a molecule able to orientate the cellular metabolic fate in a "on/off" manner. We show that XylR promote the cells' capacity to resume growth on xylose after the glucose is exhausted. XylR availability controls the length of the lag and influences the emergence of a subpopulation able to resume exponential growth on xylose. This work highlights the role of molecular switches in the establishment of phenotypic heterogeneity and proposes the glucose-xylose transition as a paradigm to understand lag phases and metabolic adaption.

Results

The lag time required to adapt to xylose depends on the genetic background of E. coli

To check the consistency of the glucose-xylose transition, we first explored the variability of the lag duration in 13 *E. coli* strains selected for their origin, lifestyle, and phylogenetic group (see Table 1 in Materials and Methods). We quantified their capacity to resume growth on M9 minimal medium supplemented with either xylose or glucose as a control after overnight growth on glucose by measuring the regrowth lag time as previously described (Enjalbert *et al.*, 2015). On glucose, the regrowth lag times were homogeneous and all the strains resumed growth with an average lag of 0.80 ± 0.24 h, except for the BL21(DE3) strain, which immediately started to grow again (lag time = 0.11 ± 0.19 h)(Figure 1A). The maximum growth rates (μ_{max}) of each strain on glucose were also quantified and ranged from 0.62 to 0.95 h^{-1} (Figure 1A). The regrowth lag times required to adapt to xylose proved to be significantly longer and varied more between strains: the lag times ranged from 1.8 ± 0.14 h lag for *E. coli* BW25113 to as high as 5.8 ± 0.72 h for strain E2348/69, while the BL21 (DE3) strain, for unknown reason, resumed growth even faster than its maximum growth rate after the switch (Figure 1A). Maximum growth rates on xylose also proved to be more variable between strains than on glucose and ranged from 0.22 to 0.91 h^{-1} (Figure 1A). No correlation was found between the maximum growth rates and regrowth lag times, either during the resumption of growth on glucose or on xylose (Figure 1B). This indicates that growth rates and lag times are not linked. We also

compared the regrowth lags resulting from resumption of growth on xylose after overnight growth on glucose with the diauxic lag observed during a diauxic growth on a glucose-xylose mix (Figure 1C). The diauxic lags were about 6 times shorter, but the two types of lags were correlated. In conclusion, lag duration appears to be a specific trait of the strains and can last for more than five hours.

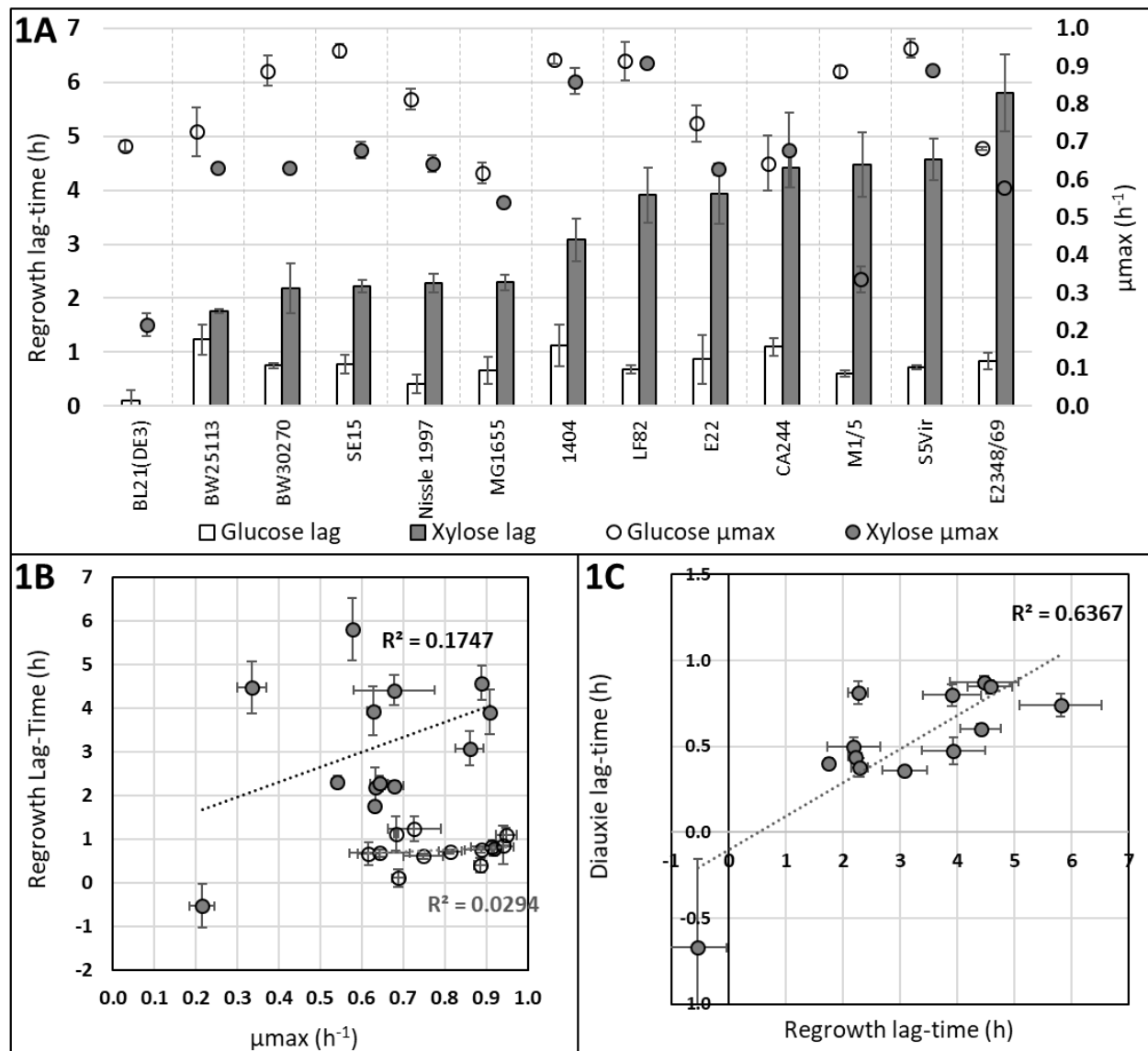


Figure 1: The length of the glucose-xylose lags in *E. coli* is background dependent. 13 different strains of *E. coli* were cultivated by switching strains previously grown on M9 glucose medium to 96-well microplates filled with 200 μ L of fresh M9 xylose or M9 glucose media. **1A)** Lag times (bars) and maximum growth rates (circles) after the switch to glucose (white) or xylose (gray) for the 13 *E. coli* strains (n=9 with 3 technical replicates for each of the 3 biological replicates). **1B)** Scatterplot of the regrowth lag time on

xylose or glucose as a function of the maximum growth rate on xylose or glucose for each strain (gray circles for xylose, white circles for glucose). **1C)** Scatterplot of glucose-xylose diauxic lag times versus regrowth lag times (cells grown overnight on glucose switched to xylose). For diauxic growth, strains previously grown on glucose were transferred to a glucose-xylose M9 medium (12.5% glucose and 87.5% xylose). Diauxic lags were calculated as explained in Materials and Methods. The negative regrowth lag time calculated for the BL21 strain was due to a higher growth rate when this strain resumed growth on xylose than its maximum growth rate during its exponential growth on xylose. For diauxic lags, $n=9$ for BW25113, 1404, LF82 and S5vir, and $n=6$ for MG1655, E2348/69, E22, BL21(DE3), BW3070, CA244, M1/5 Nissle and SP15 strains.

XylR overexpression reduces the duration of the regrowth lag

XylR is a crucial transcription factor to grow on xylose. To study the relationship between XylR expression and the duration of the regrowth lag, we overexpressed it in four of the *E. coli* wild type strains (BW25113, E22, E2348 and S5Vir). These strains were transformed with the plasmid pMET219_xylR that contains or not the *xylR* gene under the control of the constitutive *E. coli* *cysG* promoter (Zhou *et al.*, 2011). Expression of *xylR* in cells harboring this plasmid will therefore be constant and not coupled with carbon catabolite repression nor with the presence of xylose. As shown in Figure 1, the ability of the strains to resume growth on glucose or on xylose was assessed after overnight growth on glucose (Figure 2). No significant growth changes was observed with the pMET219_xylR plasmid compared to with the plasmid pMET219, but for a slight decrease in strains E22 and E2348/69 on xylose only (Figure 2). All the strains containing the pMET219_xylR plasmid showed a significantly reduced regrowth lag time when resuming growth on xylose compared to the strain containing the empty control plasmid pMET219. This reduction was 13% in the BW25113 strain (which naturally presents a short lag time), 44% in S5Vir, 55% in E2348/69, and up to 60% in E22. No significant reduction in the lag duration was observed in the controls when switched to glucose. Our results thus show that *xylR* overexpression is sufficient to reduce the regrowth lag time of *E. coli* strains resuming growth on xylose, without necessarily affecting the growth rate.

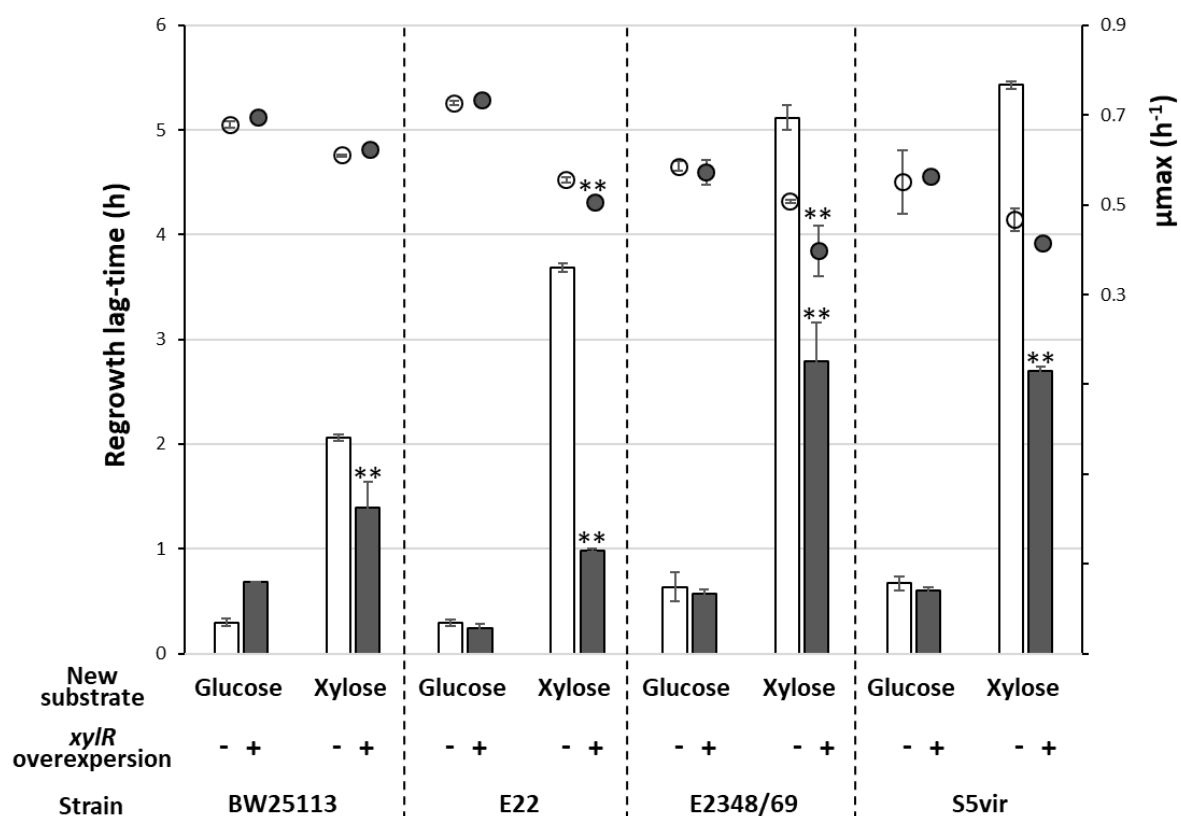


Figure 2: XylR overexpression reduces regrowth lag time during the glucose-xylose shift of different *E. coli* strains. Strains carrying the pMET219 plasmid as control (white symbols), or the same backbone plasmid with *xyIR* expression under the control of the *cysG* constitutive promoter (pMET219_*xyIR*; gray symbols), were grown like those shown in figure 1A (9 biological replicates but for strain E22 without *xyIR* overexpression where $n=6$; ** indicates a significant p -value < 0.01 compared to the control strain not overexpressing *xyIR*). The bar chart represents the lag time (in hours) before growth resumed. The circles represent the growth rates.

XylR titration extends the regrowth lag

Since increasing XylR concentration is sufficient to reduce the regrowth lag, reducing its probability to attach to its own promoter by decreasing its availability may lengthen the lag by preventing its positive retro-control. To test this hypothesis, we increased the number of XylR targets in the cells by multiplying the *xylA* promoter in order to create a titration of the transcription factor and hence reduce its capacity to promote its own transcription through the positive feedback regulatory loop. To this end, the *xylA* promoter was placed on the pSB1C3 high copy number plasmid and introduced into the BW25113 strain. The capacity of this strain to resume growth on xylose was assessed after overnight growth on glucose (Figure 3A). As expected, there was a marked increase in the lag duration from 1.7 h for the pSB1C3 plasmid without the construction to more than 19 h for the plasmid harboring the *xylA* promoter with no impact on the maximum growth rate. The lag increased to 23 h in the *xylA* promoter driving the expression of a mRFP1 red fluorescent reporter. To confirm that the increase in the regrowth lag originated specifically from the interaction between XylR and its targets on the plasmidic *xylA* promoter, we mutated the two XylR binding sites in the promoter sequence. As expected, these mutations completely stopped the increase in the lag. None of these constructions affected the resumption of growth on glucose (Figure 3B). These results show that the regrowth lag duration can be extended by titration of XylR. This result and the previous one showing that increasing XylR concentration reduces the regrowth lag duration (Figure 2) demonstrate that the lag in resuming growth on xylose depends on XylR availability.

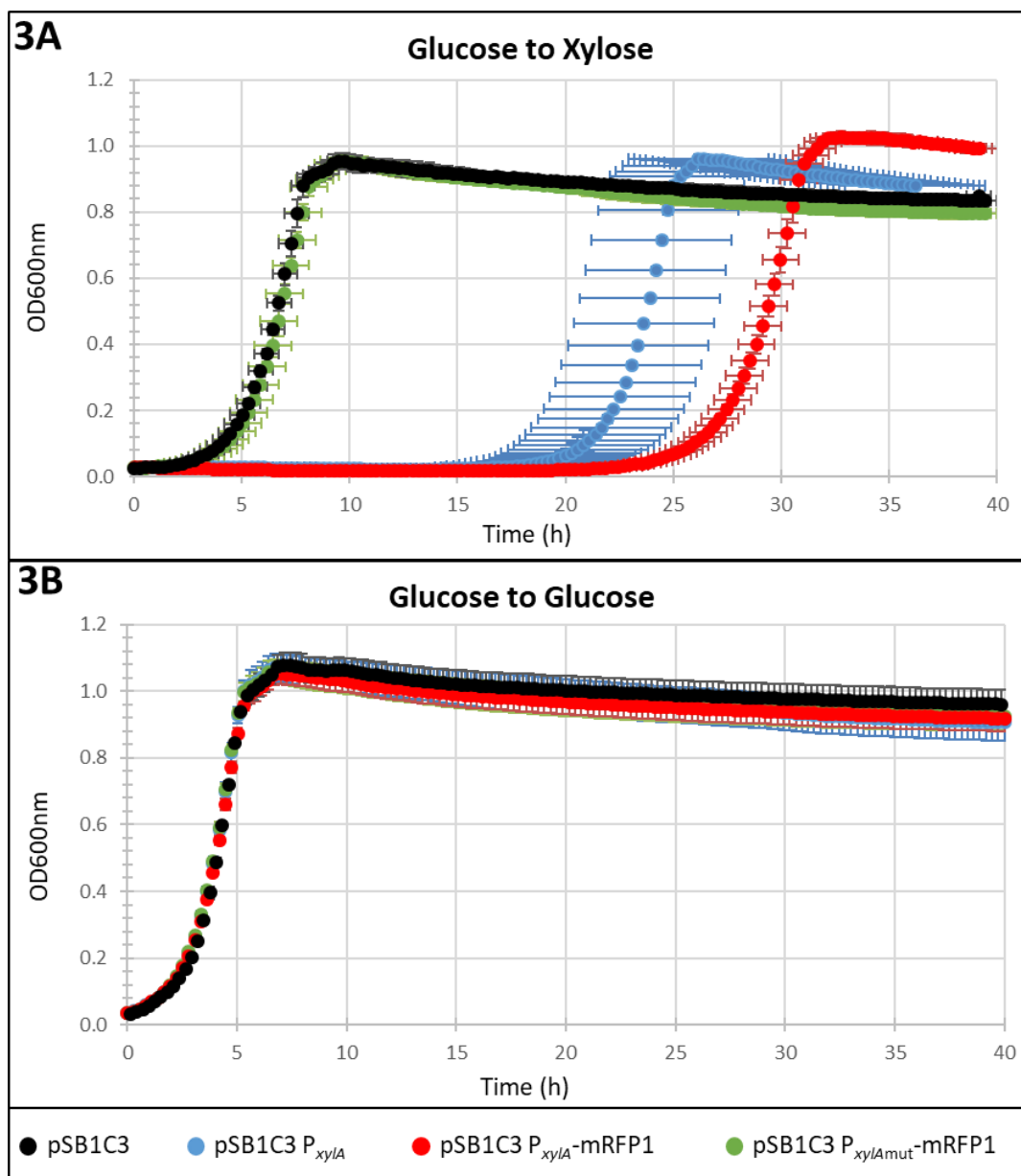


Figure 3: XylR titration modifies regrowth lag time. The ability of the *E. coli* BW25113 strain to resume growth with different plasmid constructions was assessed by switching cells after overnight culture on glucose to a 96-well microplate filled with 200 μ L of fresh M9 xylose (**3A**) or M9 glucose (**3B**) media. The BW25113 strain (in black) with the empty pSB1C3 plasmid as control, the BW25113 strain (in blue) with a pSB1C3 plasmid carrying the *xyIA* promoter, the BW25113 strain (in red) with a pSB1C3 plasmid expressing the red fluorescent protein mRFP1 under the control of the *xyIA* promoter, and the BW25113 strain (in green) with a pSB1C3 expressing mRFP1 under the control of a *xyIA* promoter deprived of its XylR binding sites. $n=6$ for all conditions.

Emergence of sub-populations during the glucose to xylose regrowth and diauxic lags

In the previous experiment, we managed to dramatically extend the duration of the regrowth lag to more than 19 h without affecting the maximum growth rate on xylose. We wondered why the culture took so long to resume growth. In one hand, this behavior is artefactual in the BW25113 strain since it is due to the plasmid presence. In the other hand, long lags up to 6 h were observed in other genetic background (Fig. 1). Since the plasmid presence allows to investigate the capacity of individual cells to express the *xylAB* operon by flow cytometry, and since bimodality has been reported for the *xylAB* promoter (Afroz *et al.*, 2014), we investigated the putative existence of subpopulations during glucose to xylose adaptation in the modified BW25113 strain. To this end, we used the high copy number plasmid described above that expresses the red fluorescent protein mRFP1 under the control of the *xylA* promoter (pSB1C3 P_{xylA} -mRFP1) so that the reporter is expressed when the cells metabolize xylose (Supplementary data file S1). We also added a P_{ihfB} -mTagBFP construction to constitutively express the blue fluorescent protein under the control of the *ihfB* promoter (Zhou *et al.*, 2011). This facilitated the detection of cells by flow cytometry. We also integrated the construction in the bacterial chromosome but the fluorescence levels were insufficient to track subpopulations (data not shown). The pSB1C3 P_{xylA} -mRFP1 P_{ihfB} -mTagBFP plasmid was inserted into the BW25113 strain and grown in controlled aerobic batch conditions. Cultures with different glucose-xylose ratios totalling up to 90 mM_{eqC} were performed, as well as 100% glucose or 100% xylose controls in M9 media (see Figure 4A for the 40% glucose and 60% xylose mix; Supplementary data file S2 for other mix ratios). With the 40% glucose and 60% xylose mix (Figure 4A), a first exponential growth phase was observed associated with glucose consumption and acetate production. After the glucose was exhausted, acetate catabolism was observed, with no detectable biomass production, as described in (Enjalbert *et al.*, 2013). A transition phase lasting more than 10 hours was then observed before the second exponential growth associated with xylose consumption, indicating that the plasmidic construction extends the diauxic lag (Fig 4A). This was also observed with the 100% xylose condition, indicating that the construction also increases the regrowth lag from glucose to xylose (Fig 4E).

	Median signal intensity mRFP1 (u.a.f)	Median signal intensity mTagBFP (u.a.f)
Glucose medium	3.55	47.04
Xylose medium	120.00	44.47

Figure S1: Median fluorescence intensity of *E. coli* BW25113 strain carrying the dual fluorescence plasmid pSB1C3 P_{xyIA} -mRFP1 P_{ihfBA} -mTagBFP in M9 glucose medium versus in M9 xylose medium. Fluorescence was measured by cytometry during the exponential growth phase. The Y2-A canal was used to quantify mRFP1 fluorescence intensity (inducible on xylose) and the V1-A canal to quantify mTagBFP fluorescence intensity (constitutive).

Flow cytometry analyses were performed with two-dimensional representation of the red inducible fluorescence versus the blue constitutive fluorescence (Figure 4B). The cytometric profiles are displayed for growth on glucose or on xylose (the first and last graph, respectively), every hour during growth on the 40% glucose and 60% xylose mix. A constant patch of non-fluorescent cells and debris (in the black circles) representing less than 10% of the signals was always observed and hence no longer taken into consideration. The profiles obtained with growth on single substrates revealed that the strain profiles differed on glucose (in the blue gate) and on xylose (in the red gate). Although these profiles were not highly focused, probably due to cell to cell variations in the plasmid copy number, combined analyses of the two reporters provided a discriminating flow cytometry signature of cells growing on glucose or on xylose. These specific cell signatures were then used for population identification and quantification in the different sugar mixtures.

During growth on glucose and until its exhaustion after 7 h, the cytometric profile of the population was homogenous and the population exhibited the same fluorescence pattern as the 100% glucose control, (Figure 4B). After glucose exhaustion (7 h), a new subpopulation gradually emerged with a cytometric pattern that matched the one obtained on the 100% xylose control. Very interestingly, the new xylose-specific

population did not appear to derive from a global shift of the whole population but to emerge from individual cells following glucose exhaustion. Figure 4C shows the theoretical biomass of each subpopulation extrapolated from the percentages resulting from the flow cytometry data. This reveals the exponential growth of the new “xylose” subpopulation, with a growth rate of 0.52 h^{-1} equivalent to what was observed with the single population growing in the 100% xylose control (0.55 h^{-1}) (see Figure 4E). Consistent results were obtained with other mixes (Supplementary data file S3), showing that glucose/xylose ratio are not impacting the phenomenon. The original population that grew on glucose grew weakly or not at all during the xylose consumption phase (0.04 h^{-1}). The same representations for the 100% glucose control and the 100% xylose controls are shown in Figure 4D and 4E with a homogenous population profile on glucose, and a long regrowth lag before the xylose subpopulation emerges on xylose. In conclusion, by titrating XylR availability, we observed the emergence of a subpopulation able to grow on xylose during the observed regrowth and diauxic phases while the initial subpopulation which developed on glucose appeared to remain stable. The new emerging subpopulation thus starts from a small number of exponentially growing individuals, and not from the adaptation of the whole population. This behavior explains why the lags can last for many hours in the modified BW25113, i.e. until the new population is big enough to be detected.

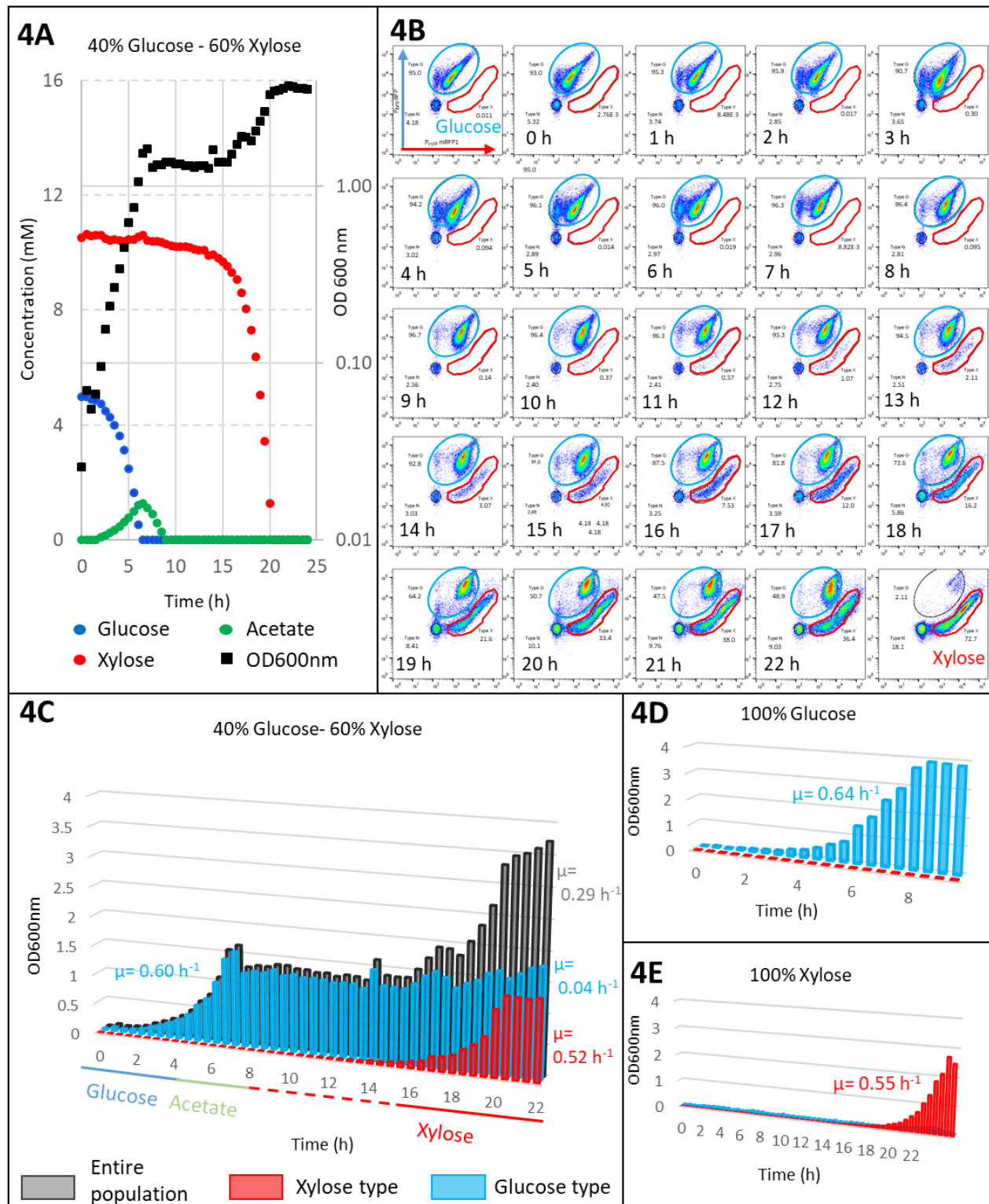


Figure 4: The length of the transition in the modified BW25113 strain is related to the emergence of a new subpopulation. Batch fermentations of the BW25113 strain transformed with the pSB1C3 plasmid carrying the P_{ihfB} -BFP P_{xylA} -mRFP1 construction were carried out in minimum medium of 90 mMeqC and sampled at 30-min intervals. **4A)** Kinetics of substrate consumption and production and growth of the strain in M9 medium supplemented with a 40% glucose and 60% xylose mix. **4B)** Cytometric profiles of the 40% glucose 60% xylose batch over time. The y-axis displays the blue fluorescence levels in arbitrary fluorescence units (a.f.u.) the x-axis the red fluorescence levels (a.f.u.). The blue gate represents the glucose population type as seen with the 100% glucose control. The red gate represents the xylose population type as seen with the 100% xylose control. **4C)** The theoretical biomass of each subpopulation was extrapolated from the percentages resulting from the flow cytometry analyses during growth on the 40% glucose and 60% xylose mix (glucose type cells in blue, xylose type cells in red, and whole population in grey). **4D)** Same representation with the 100% glucose growth. **4E)** same representation with the 100% xylose growth. Others substrate ratio are presented in the supplementary figure S2.

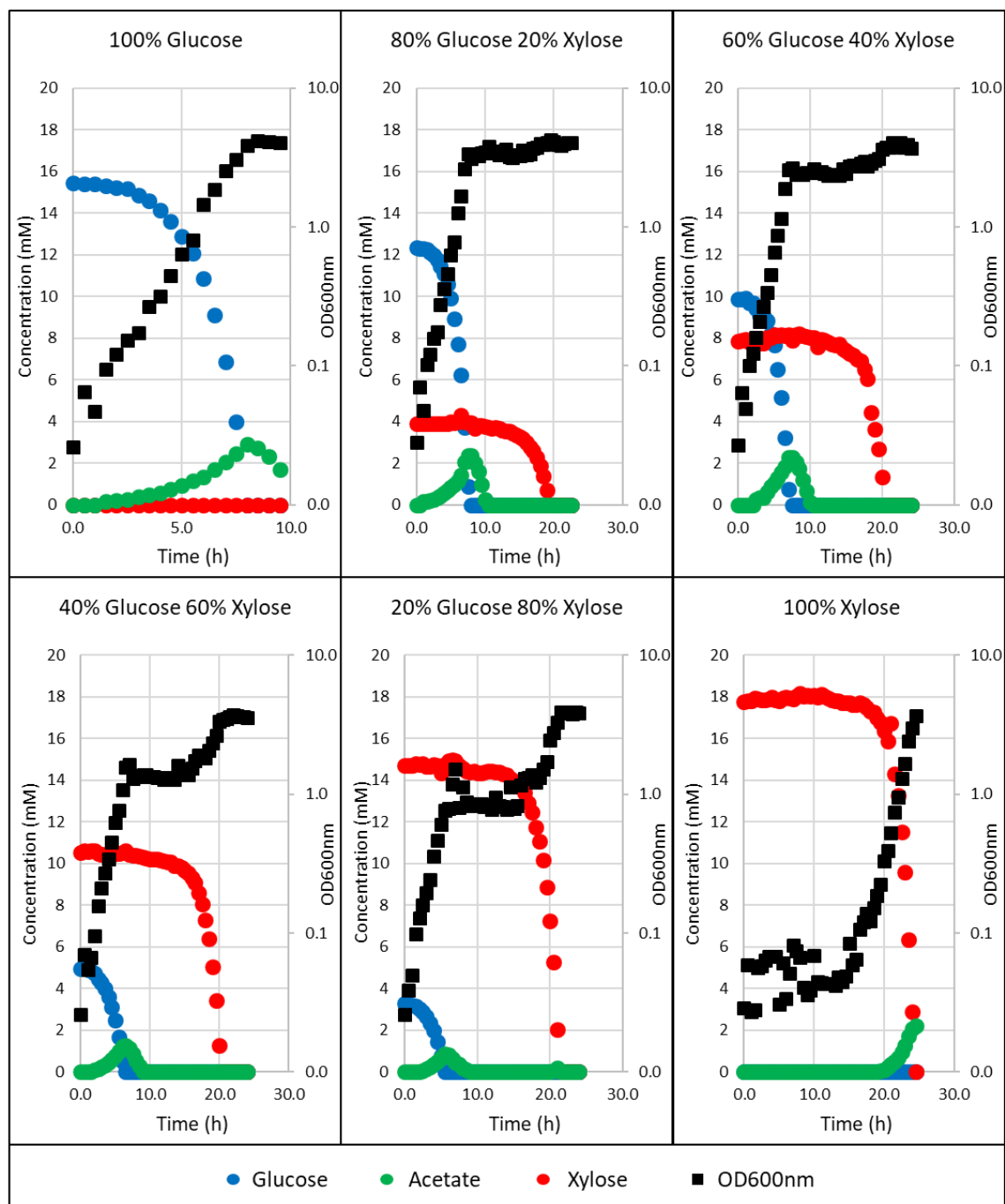


Figure S2: Kinetics of substrate consumption and production, and growth of the strain in M9 medium supplemented with a 100% glucose, 80% glucose and 20% xylose mix, 60% glucose and 40% xylose mix, 40% glucose and 60% xylose mix, 80% glucose and 20% xylose mix or 100% xylose. The marked variations in biomass observed under OD 0.1 are due to the technical limits of the spectrophotometer

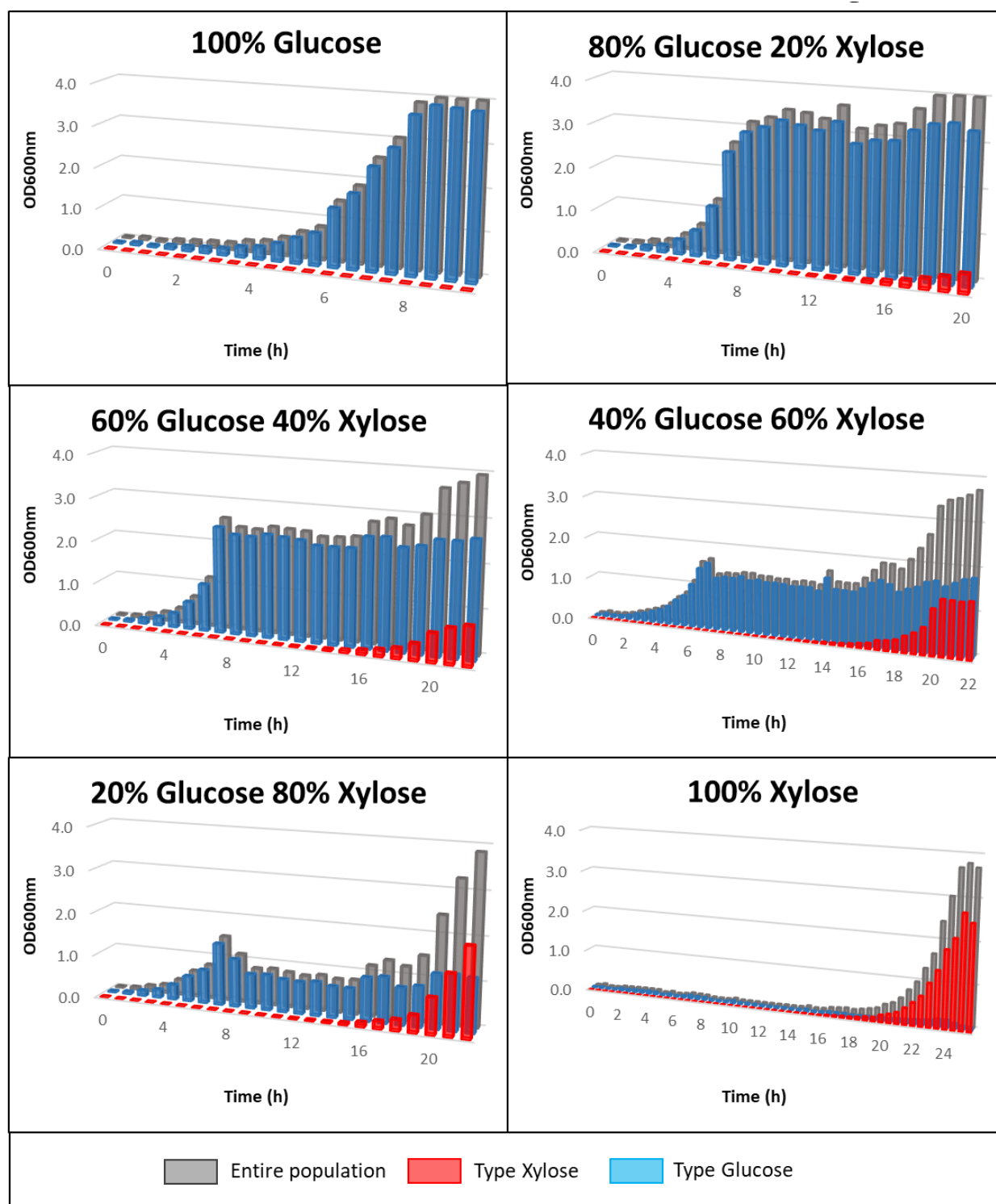


Figure S3: The biomass of each subpopulation extrapolated from the percentages resulting from the flow cytometry analyses during the growth on 100% glucose, 80% glucose and 20% xylose mix, 60% glucose and 40% xylose mix, 40% glucose and 60% xylose mix, 80% glucose and 20% xylose mix or 100% xylose (glucose type cells in blue, xylose type cells in red and the whole population in grey).

Xylose exposure creates a memory effect through XylR accumulation

Since the glucose to xylose adaptation is controlled by the XylR availability in cells, growth on xylose should produce cells with a high XylR copy number which, in turn, should allow growth to resume faster when xylose is again usable, i.e., through a memory effect. This hypothesis was tested by using the titrated strain BW25113 with the pSB1C3 P_{xylA} -mRFP1 P_{ihfB} -mTagBFP plasmid and its extended lags. The strain was grown overnight on xylose (or glucose as control) before switching to fresh medium with a mix of 12.5% glucose 87.5% xylose in microplates with inoculation concentrations ranging from OD_{600nm} 0.1 to 0.00001. These different concentrations of inoculant allowed the cells to perform a range of generation numbers from 1.8 to 14.8 on glucose before the final growth phase on xylose. We hypothesized that a more recent exposure to xylose would enable faster resumption of growth when xylose again became usable (i.e., once the glucose was exhausted). In this way, we measured the diauxic time to resume growth on xylose following the glucose growth phase (Figure 5). The diauxic lag appears way shorter using microplate than in bioreactors (as in Fig. 4A). For the five cultures pre-grown on xylose and switched at different inoculum rates to the glucose-xylose mix, we found a clear positive correlation between the diauxic lag durations and the number of generations on glucose since the last exposure to xylose, indicating that even a couple of generations on glucose seems efficient. No impact on diauxic time was observed in the control with pre-growth on glucose before switching to the carbon mix (data not shown). The proportionality between the diauxic lag times and the generation number on glucose is likely related to the XylR dilution in the daughter cells. This clearly identifies a "memory effect" when cells have been pre-exposed to xylose: XylR abundance in the cells that have been grown on xylose decreases when cells are no more producing XylR, like when grown on glucose, but remains higher than in cells that have never seen xylose, hence facilitating the switch toward xylose consumption when this sugar is again available.

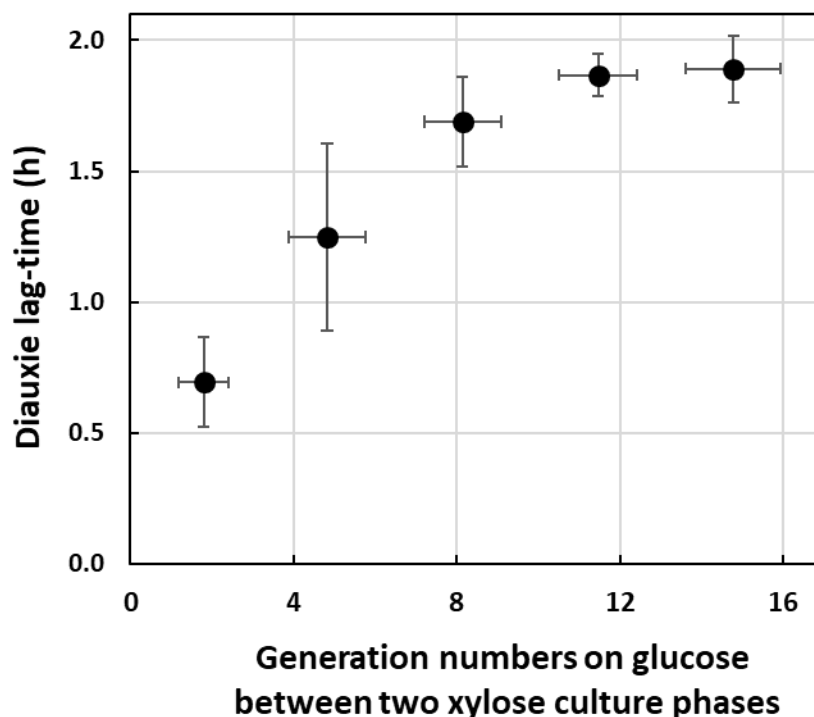


Figure 5: Previous exposure to xylose creates a memory effect. The strain with the titration plasmid was grown on xylose and switched to a glucose-xylose mix on microplates with different concentrations of inoculant so that the number of generations on glucose differs between two growth phases on xylose. The duration of the diauxic lags are represented as a function of the number of generations on glucose. For each condition, four replicates were performed. $n=4$ for all generation numbers but the highest ($n=2$).

Mathematical modeling of the transition highlights the effect of XylR quantity on the duration of the diauxic lag

To formalize the mechanism linking XylR concentration, subpopulation and the diauxic lag, a modeling approach was developed based on the results of cytometry obtained with the fluorescent reporter genes. The objective was to be able to predict the behavior of the non-fluorescent wild type for which modeling of the system was mandatory. The major phenomenon to be modeled is the appearance of the “xylose” subpopulation cells. From our experimental observations, they are clearly present 1.5 hours after exhaustion of the glucose (point 8 h in figure 4B), and are very likely present in very small numbers soon after glucose exhaustion when the carbon catabolite repression is switched off. Based on our previous finding concerning the major role played by XylR concentration in the lag

duration, our main hypothesis is that low XylR availability results in a low probability to trigger the XylR positive feedback loop and hence, to switch to xylose consumption. We decided to opt for a stochastic model to account for the probability that a very limited number of cells is able to resume growth on xylose at the end of glucose consumption. This scenario is backed up by proteomic data showing that XylR is present at a rate of 1 ± 0.4 copy per cell during growth on glucose and at 57 ± 8 copies per cell on xylose (Supplementary data file in (Schmidt *et al.*, 2016)). Thus, after glucose is exhausted, the single copy of XylR needs to find its target on the *xylAB* promoter to enable growth on xylose. Knowing that the pUC19-derived pSB1C3 is a high copy plasmid, it seemed obvious that target multiplication reduces the probability of triggering the loop. Starting from these hypotheses, we built the model presented in figure 6A and fully described in Supplementary file S8. Basically, the model allows a population X_1 with low level of XylR to grow on glucose and to switch to growth on xylose only if glucose is exhausted and xylose is present. In this situation, cells from population X_1 will turn at rate λ into cells with Y population features (i.e., xylose consumers with a high level of available XylR). Return to the initial X_1 state is possible in the model if glucose again becomes available: the cells will transit through an X_2 state characterized by a decrease in XylR content by dilution through cell division until the cells reach the XylR level of the X_1 population. For a population close to a billion cells, the stochastic model is approximated by a deterministic dynamic system, making it easier to calibrate the parameters (Supplementary material S9). In other words, the core of the model remains stochastic to capture the apparition of the new sub-population, but it could be approximated to a deterministic model once the number of individuals is high enough. This model was calibrated with data from culture with 100% glucose, 100% xylose, a 60% glucose-40% xylose mix and a 40% glucose-60% xylose mix. It accurately predicted these four conditions in terms of subpopulation distribution and substrate consumption as well as the data not used for the calibration from cultures with 80% glucose-20% xylose and 20% glucose-80% xylose mixes (Figure 6B). Parameter calibrations made it possible to quantify the rate of appearance of xylose consumers (Y) in the original population (X_1) after glucose exhaustion to $2.02 \times 10^{-4} \text{ h}^{-1}$, i.e., about 2 cells over 10,000 per hour.

Next, the model was used to simulate the behavior of the wild type strain throughout the diauxie; the structure of the model was retained but the values of growth and substrate consumption rates specific to the strain were recalibrated. Very interestingly, the

subpopulation balance appeared to be the inverse of what we observed with the strain carrying the pSB1C3 P_{xylA} -mRFP1 P_{ihfB} -mTagBFP plasmid (Supplementary file S4). Indeed, without the plasmid, the individual rate of appearance of xylose consumers from the original population after glucose exhaustion was 2.09 h^{-1} , i.e., about 8 760 cells to over 10,000 per hour. This result is of particular interest since it makes it possible to predict the behavior of the wild type strain without the bias created by markers. The model allowed us to estimate that, by the time xylose was exhausted, 99.46% of the X_1 population have shifted to xylose consumption. In a population of close to a billion cells, this predicts that about 5.4 million cells will remain in a dormant-like state.

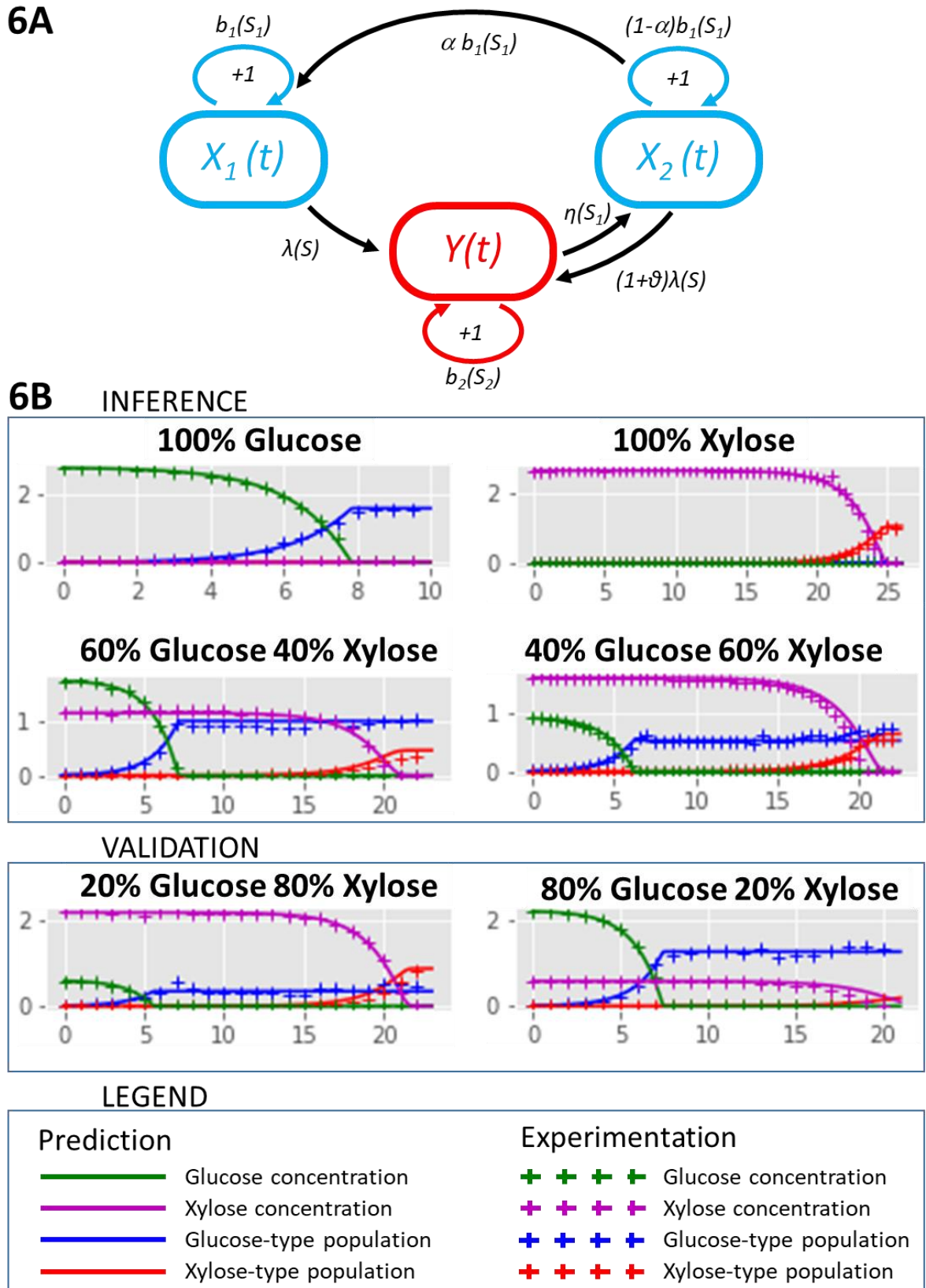


Figure 6: Modeling subpopulation behavior. 6A) Scheme of the stochastic model. The diagram shows individual growth and transitions between cell types in the population. X_1 corresponds to glucose consumers with the basal level of XylR, Y to xylose consumers with a high level of XylR, and X_2 to glucose consumers that previously grew on xylose with an initially high but decreasing level of XylR. We assume

these mechanisms occur after a random time distributed according to an exponential law with the corresponding rates. Indeed, an individual that grows on glucose (class X_1 or X_2) divides at rate $b_1(S_1)$ or switches to xylose consumption at rate $\lambda(S)$ and $(1+\theta)\lambda(S)$, respectively. Specifically, in the X_2 compartment, where cells have many *xylR* copies, each individual can give birth to a cell of the compartment X_1 with probability $0 < \alpha < 1$ because of *xylR* dilution. In addition, an individual growing on xylose (class Y) divides at rate $b_2(S_2)$ or switches to glucose consumption (class X_2) at rate $\eta(S_1)$, if glucose is abundant. **6B)** Validation and predictions of the model (solid lines) compared to the experimental data (crosses) during growth on 6 glucose/xylose mixes (glucose cell type in blue, xylose cell type in red, glucose in green, xylose in violet). The conditions 80% glucose 20% xylose and 20% glucose 80% xylose were not used to estimate the model parameters. The lines represent predictions whereas in the other conditions, they represent validation.

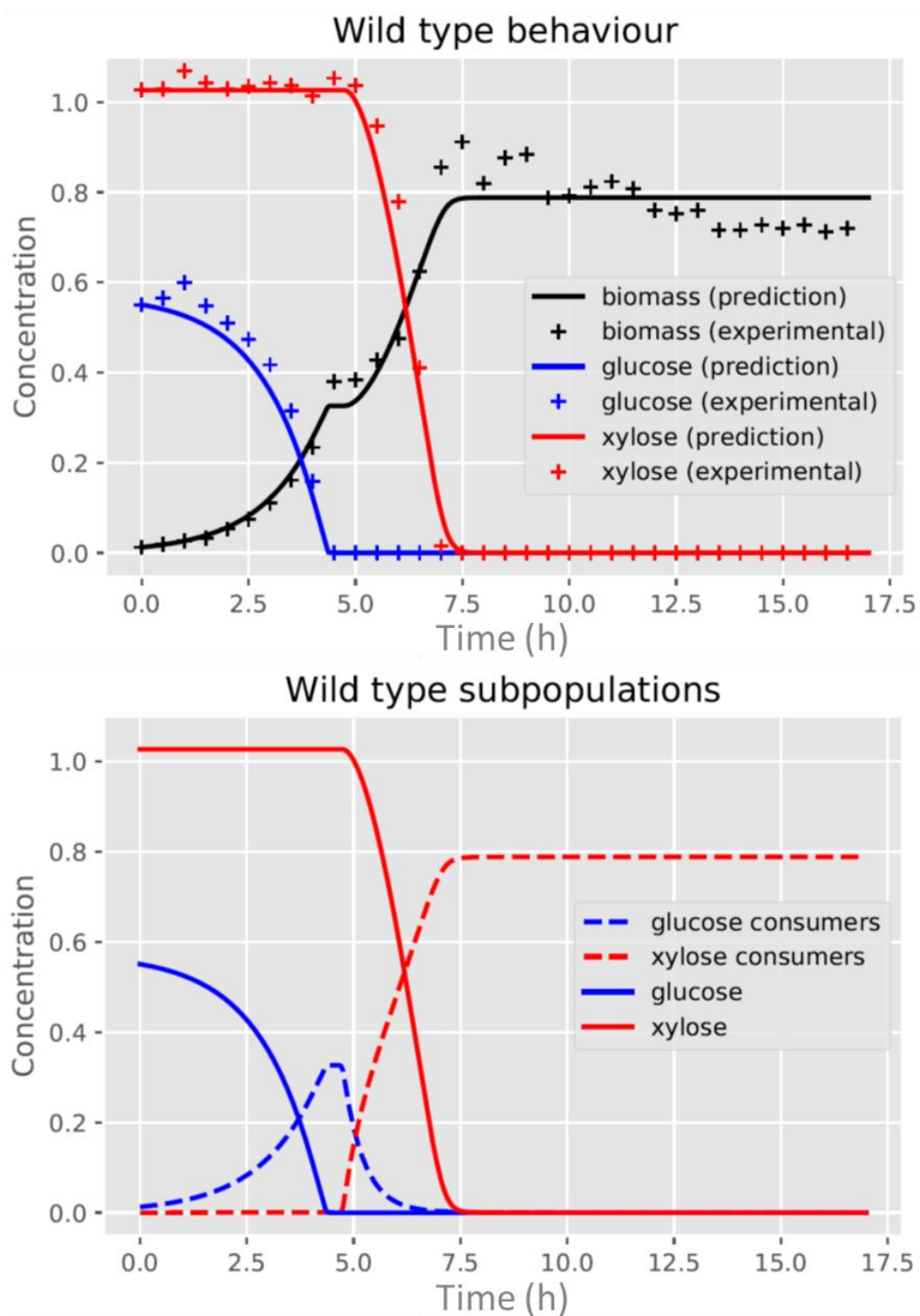


Figure S4: Behavior of wild type strain BW25113 on a glucose-xylose mix. Experimental and prediction data are presented for biomass concentration (g.L^{-1}) and substrates (g.L^{-1}) in the top row and prediction data for subpopulations in the bottom row of graphs.

Discussion

Lag times have long been considered as adaptation periods for a microbial population to use a new resource. The recent interest of the microbiologist community in single cell analysis revealed a more complex reality concerning these lag phases. Here, we demonstrated that the main parameter governing the length of the regrowth or diauxic lag phases is not a temporal requirement to produce new machinery for growth on the following substrate, but the availability of a specific transcription factor. In the glucose-xylose model we studied here, XylR accumulation is the master molecular switch governing the metabolic fate of the cells. Consequently, titration of XylR extends lag time while overexpression of XylR reduces it. These effects are limited to lag durations and do not affect the maximum growth rates in general, very likely because XylR is not needed on glucose while its regulatory loop ensures that it is not limiting on xylose.

This positive XylR feedback loop on its own expression is the centerpiece of the system. It enables the creation of the bimodality in the cell population reported by Afroz *et al.* (Afroz *et al.*, 2014), with some of the cells able to trigger the loop and thus shift to full growth on xylose while the others fail to do so and remain in a dormant-like state. Such a regulatory design has already been described in microorganisms faced with other substrates (Ozbudak *et al.*, 2004; Afroz *et al.*, 2014; Bhogale *et al.*, 2014; Koirala *et al.*, 2016; Guantes *et al.*, 2016). The question that remains is what determines the metabolic fate of individual cells. Our results suggest that XylR availability is sufficient to explain all the phenotypes we observed. Yet, it would be remarkable if a single protein accumulation could explain so much, especially in the light of all the different molecular explanations behind phenotypic heterogeneity in the substrate usage reported in the literature (Ito *et al.*, 2009; Solopova *et al.*, 2014; Kotte *et al.*, 2014; Lambert and Kussell, 2014; Yaginuma *et al.*, 2014; Martins and Locke, 2015; Amato and Brynildsen, 2015; Axelrod *et al.*, 2015; Kaiser *et al.*, 2018; Dal Co *et al.*, 2018). It would be interesting to study the noise in the expression of XylR and properties of the resulting distribution even if experimental XylR quantification at the single cell level remains a great experimental challenge. This could reveal the presence of cells with high levels of XylR even before glucose exhaustion, but likely not large enough to trigger the loop since we never observed cells with a xylose signature during growth on glucose. Another source of heterogeneity could be the unequal

distribution of transcription factor between the mother and daughter cells since the concentration of XylR is close to one copy per cell on glucose (Schmidt *et al.*, 2016). Indeed, this would create cells depleted of XylR that would be unable to trigger the loop in the wild type strain (Robert *et al.*, 2010). Our results do not prove that other parameters reported by others such as cell size, cell cycle phase at the moment of glucose exhaustion, cell age, ribosome availability, or ppGpp concentrations have no effect on the adaptation from glucose to xylose. The fact that XylR overexpression does not produce regrowth lags with the same optimal duration in different strains, or the different lag durations observed between microplate or cultures in bioreactor supports the hypothesis that other parameters are at stake here. It is for example rational to hypothesize a role for carbon catabolite repression in the process. The presence of CRP-AMPc on the promoter could be required for XylR to fix on its recognition sites, or to create a slight increase in *xylR* promoter expression ensuring that XylR is over a limiting threshold to trigger the loop that most cells with the titration plasmid cannot overcome.

The model parameter estimation showed that most of the cells are not able to switch to xylose consumption in the mutant strain because of XylR titration, while the inverse situation was observed in the wild-type strain with only a small fraction of cells unable to switch. The cells that are unable to use the available substrate could be dormant cells. Such behavior has already been reported in other conditions when a population was faced with a new substrate (Amato *et al.*, 2013; Solopova *et al.*, 2014; Jöers and Tenson, 2016). This was interpreted as an adaptive feature to minimize risk since the dormant cells are more resistant to environmental insults including stresses or antibiotics (Veening *et al.*, 2008; Amato *et al.*, 2013; Radzikowski *et al.*, 2016). These cells are metabolically active with very slow buildup of biomass and reduced metabolite pools (Radzikowski *et al.*, 2016), which could match the very small growth rate observed for this subpopulation in our study. Even if the fraction of such cells is low in the wild type strain, the model predicted that it nevertheless results in millions of individuals in the bacterial population as a whole. This reinforces the hypothesis that metabolic switches are related to antibiotic resistance (Amato *et al.*, 2013).

To our knowledge, this is the first work exploring phenotypic heterogeneity in the use of substrate during diauxic batch fermentation of a monoclonal population in infinitely stirred bioreactor. Most studies has indeed used microfluidic devices and/or manual switching of the cells from one condition to another, both being not ideal for analyzing responses to environmental changes that happen as a consequence of growth (Rosenthal *et al.*, 2018). This allows confirming that phenotypic heterogeneity in the use of substrate observed at the scale of few individual cells is also true in large populations growing naturally. Likewise, phenotypic heterogeneity has been reported to have an incidence on metabolic adaptation of a population from glycolytic to gluconeogenic substrates (Kotte *et al.*, 2014). Here we report that the same incidence could be true for metabolic adaptation from glycolytic to glycolytic substrates.

Several applications of this work can be envisaged. It now appears possible to optimize lags to ensure the profitability or sustainability of biotechnological processes, for example, by reducing the length of lag phases. Our results point to two levers that could achieve this. The first is by pre-exposing the cells to the substrates they will use during the process to take advantage of the memory effect. The effect has indeed been observed with a variety of substrates (Lambert and Kussell, 2014; Kaiser *et al.*, 2018) and this solution appears to be both cheap and effective. The second solution is modifying the positive feedback loop through overexpression of the limiting transcription factor as we did here, or by modifying the regulatory network. Indeed, such loops are not limited to the use of xylose but are reported with increasing frequency in adaptive phenomena (Choi *et al.*, 2008; Choudhury *et al.*, 2018; Kaiser *et al.*, 2018; Takhaveev and Heinemann, 2018). Other levers very likely exist. For example, we observed here that the transition is faster during microplate culture than during well controlled bioreactor culture even though we have no explanation for this as yet. Conversely, it would be interesting to extend the diauxic lag phase since this could promote the formation of dormant cells, which could be a key factor in ensuring the robustness of the strain during challenging bioprocesses, or for strain storage. Finally, the extended duration of the glucose-xylose transition could be of fundamental interest to study the properties of dormant cells.

Lastly, it is now clear that lag phases (as well as so-called “stationary phases”) are far more complex than originally thought by Jacques Monod in 1942 (Monod, 1942). Until the last decade, most interest focused on the exponential phases and we are only now beginning to become aware of the subtle behavior of cells that are not growing, from the organization of subpopulations to persistence and maintenance costs.

ACKNOWLEDGEMENTS

We thank Pierre Reitzer, Emmanuelle Marchaud, Sandrine Paise, Valérie Laroute, Julie Lesage, Manon Costa, Julie Taba and Sandrine Laguerre for their contribution to this work. This work was supported by the 3BCAR Carnot (project HME) and the ANR (project JANUS; ANR-19-CE43-0004-01). TWB (Toulouse White Biotechnology) and its staff members are gratefully acknowledged for allowing us access to their cytometry facilities. C.Graham, S. Méléard and J. Tchouanti were supported by the Chair « Modélisation mathématique et Biodiversité » of Veolia-Ecole polytechnique-Muséum national d'Histoire naturelle-Fondation X.

Materials and Methods

All data are accessible on the INRAE platform (<https://doi.org/10.15454/XWCMZE>).

Strains and plasmids

All the strains used in this study are listed in Table 1.

Table 1: *Escherichia coli* strains used in the study

Strain	Phylo Group	Description	Source or Reference
BW25113	A	K-12 derivative laboratory strain	(Baba <i>et al.</i> , 2006)
MG1655	A	K-12 derivative laboratory strain	(Baba <i>et al.</i> , 2006)
BL21(DE3)	A	Commercial strain for heterologous expression	NEB strain catalog no. C2527
BW30270	A	K-12 derivative laboratory strain	(Datsenko and Wanner, 2000)
E22	B1	Rabbit enteropathogenic strain	(Marchès <i>et al.</i> , 2000)
S5Vir	B1	Sheep enteropathogenic strain	(Smith <i>et al.</i> , 2018)
SP15	B2	Commensal strain from human feces	(Johnson <i>et al.</i> , 2008)
M1/5	B2	Commensal strain from human feces	(Martin <i>et al.</i> , 2013)
Nissle	B2	Human non-pathogenic probiotic strain	(Revelles <i>et al.</i> , 2013)
E2348/69	B2	Human prototype strain to study enteropathogenic <i>E. coli</i> biology (HG2)	(Iguchi <i>et al.</i> , 2009)
LF82	B2	Human strain isolated from a patient with Crohn's disease (HG2)	(Miquel <i>et al.</i> , 2010)
CA244	B2	K-12 derivative laboratory strain	(Reuven and Deutscher, 1993)
1404	C	Isolated from septicemic calf	(Pérès <i>et al.</i> , 1997)

All the plasmids used in this study are listed in Table 2.

Table 2: Plasmids used in this study

Plasmids	Relevant characteristics	Reference
pSB1C3 mRFP1	Cloning vector of this study. pUC19-derived pMB1 ori, chloramphenicol resistance, with mRFP1(FPbase ID:5YCFA) coding sequence	iGEM part BBa_E1010
pSB1C3 mTagBFP	Cloning vector. pUC19-derived pMB1 ori, chloramphenicol resistance, with mTagBFP (FPbase ID:BFJKS) coding sequence	iGEM part BBa_K592100
pSB1C3 P_{ihfB} mTagBFP	pSB1C3 plasmid containing mTagBFP under the control of <i>ihfB</i> promoter	This study
pSB1C3 P_{xylA} mRFP1	pSB1C3 plasmid containing mRFP1 under the control of <i>xylA</i> promoter	This study
pSB1C3 $P_{xylA-mut}$ -mRFP1	pSB1C3 plasmid containing mRFP1 under the control of a mutated <i>xylA</i> promoter, mutations were introduced in <i>xylR</i> binding sites	This study
pSB1C3 P_{xylA} -mRFP1 P_{ihfB} -mTagBFP	pSB1C3 plasmid containing mTagBFP under the control of <i>ihfB</i> promoter and mRFP1 under the control of <i>xylA</i> promoter	This study
pSB1C3 P_{xylA}	pSB1C3 plasmid containing <i>xylA</i> promoter	This study
pMET219	Plasmid pBR322 with tetR but ampR replaced by cmR. Negative control of the plasmid pMET219	Gift from Sébastien Nouaille
pMET219 $P_{cysG-xylR}$	Plasmid pMET219 with tetR replaced by <i>xylR</i> under control of <i>cysG</i> promoter	This study

All the plasmids based on the pSB1C3 backbone were constructed through digestion and ligation. The *xylA* promoter was obtained by amplification of 800 bp of the regulatory area of the gene upstream of its ATG from genomic DNA of the BW25113 strain. The *xylA* mutated promoter was designed as described in Supplementary data file S5 and obtained as a synthetic gene (Eurofins). *EcoRI* and *XbaI* restriction sites at 5' and *SpeI* restriction site at 3' were used for cloning on the chassis plasmid. Primers (Supplementary data file S6) were designed from the sequences in the Ecocyc database obtained from Eurofins (France). For the construction of the double fluorescence plasmid, the pSB1C3 P_{ihfB} -mTagBFP was digested with *EcoRI* and *SpeI* to be ligated in the pSB1C3 P_{xylA} -mRFP1 linearized with *EcoRI* and *XbaI* digestion. The plasmids based on pBR322 were constructed by DNA HiFi assembly (New England Biolab). Fragments were amplified from

E. coli DNA or plasmid purification with 20 bp of homology between each fragment. For all constructions, transformants were selected on LB plates containing 40 mg/L of chloramphenicol. The plasmids were checked by sequencing (Eurofins). Plasmids were then transformed into *E. coli* BW25113 strain.

CGCGGCCGCTTCTAGAATTTTGTAGCAACTAAACAGGGGAAAACTAATACAGAAATTTATCTTTTCG
ATTACGATATATGGTTTATTTCTTGATTATGACCGAGATCTTACTTTTGTGCGCAATTGTACTTAT
TGCATTTTCTCTTCGAGGAATTACCCAGTTTCATCATTCCATTTATTTTGTGAGCGAGCGCACAC
TTGTGTAAATCTCAATAGCAGTGTGATTAACATTAATGAGCAACTGAAAGGGAGTGCCCAATAT
TACGACATCATCCATCACCCGCGGCATTACCTGATTATGGAGTTCAATA

Figure S5: Sequence of the xylA mutated promoter used in this study. The xylR fixation sites are in red and mutation in these sites are in green.

Name	Sequence	Utilisation
REITZER_xylA3	ATGTACTAGTATTGAACTCCATAATCAGG	xylA promoter amplification
REITZER_xylA5	ATGTGAATTCGCGGCCGCTTCTAGAATTTTGTAGCAACTAAAC	xylA promoter amplification
New-ihfB3	ATGTACTAGTGATTCTCCGGTTCCTTAAAGG	ihfB promoter amplification
ihfB5-800	ATGTGAATTCGCGGCCGCTTCTAGAAGTTGACGCTAAAGGCGCAAC	ihfB promoter amplification
cysG3-800	ATGTACTAGTGTAACAAACCCCTTAGTAATTAACC	cysG promoter amplification
cysG5-801	ATGTGAATTCGCGGCCGCTTCTAGACTATTAATAAGTGTGCGGCTAAC	cysG promoter amplification
gib3_xylR_for	TTACACACTAGAATGTTTACTAAACGTCACCGCATC	xylR amplification
gib3_xylR_rev	TTCTAGTGTCTACAACATGACCTCGCTATTACATC	xylR amplification
XylR_CysG_for	CATCTACCTGCCTGGACAGCACTATTAATAAGTGTGC	pMET219-XylR construction
XylR_CysG_rev	TCACTATGGCGTGCTGCTAGTGTCTACAACATGACCTC	pMET219-XylR construction
pBR322_rev_gib2	TGCTGTCCAGGCAGGTAGATGAC	pMET219-XylR construction
pBR322_for_gib2	CTAGCAGCACGCCATAGTGACTG	pMET219-XylR construction

Figure S6: List of primers used to build the plasmids used in this study

Growth media

The culture were grown in M9 medium containing 33.7 mM Na₂HPO₄, 22 mM KH₂PO₄, 8.55 mM NaCl, 9.35 mM NH₄Cl, 1 mM MgSO₄, 0.3 mM CaCl₂, 13.4 mM EDTA, 3;1 mM FeCl₃-6H₂O, 0.62 mM ZnCl₂, 76 µM CuCl₂-2H₂O, 42 µM CoCl₂-2H₂O, 162 µM H₃BO₃, 8.1 µM MnCl₂-4H₂O, 1 µg.L⁻¹ thiamine and carbon sources at 90 mEqC (100% glucose, 100% xylose, 80% glucose + 20% xylose, 60% glucose + 40% xylose, 40% glucose + 60% xylose, or 20% glucose + 80% xylose). For the culture of strains containing plasmids, M9 medium was supplemented with 20 mg.L⁻¹ chloramphenicol.

Microplate cultivation

Cells were pre-grown overnight in 5 mL tubes with M9 glucose medium at 37 °C with agitation at 150 rpm. Overnight cultures were centrifuged at 5 000 rpm for 3 min and pellets were suspended in M9 glucose medium or M9 xylose medium at an initial OD_{600nm} of 0.05. The 96-well microplates were inoculated in triplicate with 200 µL of the cell suspensions or 200 µL of medium as control. Microplates with lids were read at OD_{600nm} at 17-minute intervals for 24 h with a Versamax microplate reader (Molecular devices, USA). Maximum growth rates $\mu_{\max}$ were determined during the exponential growth phase, and estimated using the coefficient of the linear regression of the curve representing ln(OD) versus time.

Regrowth lag times were considered as the difference in time needed to reach the same OD point in the exponential phase between growth with a time lag and growth that started immediately at the maximum growth rate, as defined by Enjalbert (Enjalbert *et al.*, 2015). Diauxic lag was considered as the time lost in comparison to hypothetical growth in which the first exponential phase is immediately followed by a second exponential phase. Formulas and demonstrations of how the lag was calculated are provided in supplementary data file S7.

Statistical validation was based on paired two-tailed Student's t tests, assuming equal or unequal variances according to the F test. Means were calculated from three biological replicates with three technical replicates each (n=9).

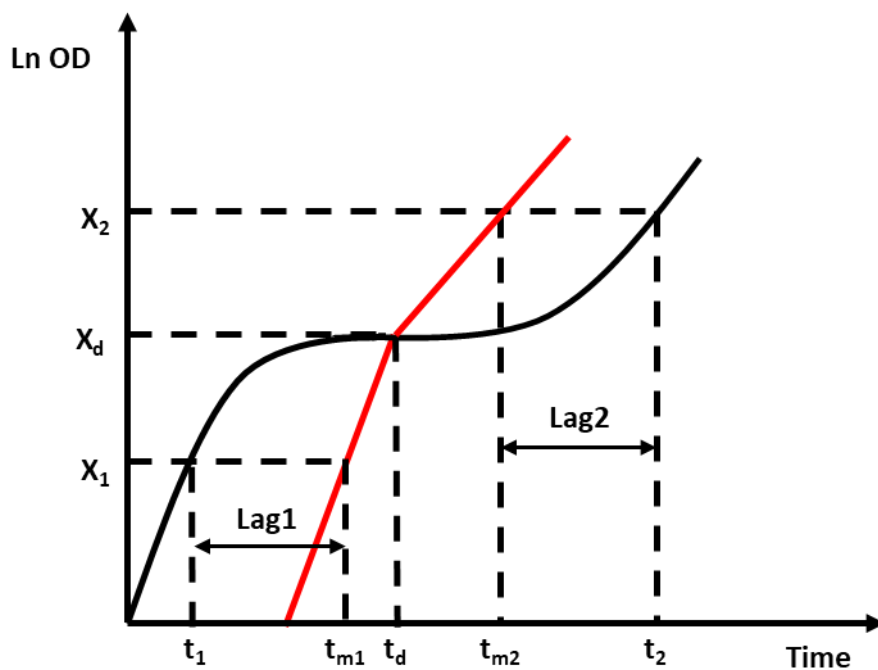


Figure S7: Demonstration of the diauxic lag formula

Here we consider the diauxic lag as the time lost in comparison to hypothetical growth where a first exponential phase is immediately followed by a second exponential phase (in red in the figure).

X_d : a biomass point reached during the diauxic lag.

μ_{max1} : the maximal growth rate on the first substrate.

μ_{max2} : the maximal growth rate on the second substrate.

X_1 : biomass reached during the exponential growth on the first substrate at time t_1 or at time tm_1 if the growth rate was theoretically maximal until td

X_2 : biomass reached during the exponential growth on the second substrate at time t_2 or at time tm_2 if the growth rate was theoretically maximal from td .

The diauxic lag is thus equal to lag1 ($tm_1 - t_1$) for the glucose growth phase plus lag2 ($t_2 - tm_2$) for the xylose growth phase. Based on the formula proposed by Enjalbert (Enjalbert et al., 2015), we demonstrated the diauxic lag calculation as followed:

- Demonstration for lag1

$$\mu_{max1} = \ln(X_d/X_1)/(td - tm_1), \text{ thus } (td - tm_1) = \ln(X_d/X_1)/\mu_{max1}$$

Given that $\text{Lag1} = (tm1 - t1) = td - (td - tm1) - t1$

Thus $(tm1 - t1) = td - (\ln(Xd/X1)/\mu_{max1}) - t1$

- Demonstration for lag2

$\mu_{max2} = \ln(X2/Xd)/(tm2 - td)$, thus $(tm2 - td) = \ln(X2/Xd)/\mu_{max2}$

Given that $\text{lag2} = (t2 - tm2) = (t2 - td) - (tm2 - td)$

Thus $(t2 - tm2) = (t2 - td) - (\ln(X2/Xd)/\mu_{max2})$

Since diauxic lag = lag1 + lag2

Diauxic lag = $td - (\ln(Xd/X1)/\mu_{max1}) - t1 + (t2 - td) - (\ln(X2/Xd)/\mu_{max2})$

So that Diauxic lag = $t2 - t1 - (\ln(Xd/X1)/\mu_{max1}) - (\ln(X2/Xd)/\mu_{max2})$

Batch cultivation

Bioreactor batch cultures were performed in a Sartorius Biostat B plus bioreactor in 1 L of M9 medium with 90 mEqC of glucose and/or xylose mixes. The temperature was set at 37 °C and pH at 7. Non-limiting aeration conditions were obtained with an air flow set at 0.35 L.min⁻¹ and adaptation of stirring to maintain pO₂ > 20%. Growth was assessed by OD_{600nm} measurement at 30-min intervals with a LibraS4 spectrophotometer (Biochrom, UK). Glucose, xylose and acetate concentrations were quantified by HPLC (Agilent, USA) with an HPX 87H Ion exclusion column (Biorad) as described in (Morin *et al.*, 2017).

Flow cytometry

Cells were sampled at 30-min intervals and immediately frozen at -20 °C, tests showed that freezing has no influence on the level of fluorescence. After thawing and diluting in filtered 0.9% NaCl to obtain a cell concentration of 1x10⁶ cells.mL⁻¹, fluorescence intensities were measured with a Masquant VYB cytometer (Miltenyi Biotec, Germany) equipped with a 561 nm yellow laser for excitation of mRFP1 and a 405nm purple laser for mTagBFP. Data acquisition was set at 40 000 events per sample. FlowJo X software was used for cytometry data analysis. A gate was created in the dot plot of the forward scatter channel (FSC-H) versus the side scatter channel (SSC-H) to distinguish bacteria from technical noise. A second gate was created in the dot plot of the SSC-H versus the

SSC-A to select single cells. The mTagBFP fluorescence emissions were analyzed with the V1-H channel (BP450/25nm) and the mRFP1 emissions with the Y2-H channel (BP615/10 nm). The OD_{600nm} subpopulation was extrapolated from the total OD_{600nm} population and the percentage of each subpopulation obtained in flow cytometry, based on blue and red fluorescence.

Mathematical modeling

A detailed description of the model, including its rationale, parameters, and code, is provided in supplementary data file S8.

Memory effect experiment

Cells were pre-grown overnight in tubes containing 5 mL of M9 glucose medium or M9 xylose medium at 37 °C with agitation (150 rpm). Cells were centrifuged at 5 000 rpm for 3 min and pellets were suspended and inoculated at 0.1 OD_{600nm} in 100 mL Erlenmeyer flasks containing 20 mL of M9 glucose medium or M9 xylose medium. Once in the exponential phase (around 1 OD_{600nm}), the cultures were diluted (theoretical OD_{600nm} 0.1; 0.01; 0.001; 0.0001 and 0.00001) and inoculated in duplicate in 96-well microplates in 200 µL of M9 glucose, M9 xylose or M9 12.5% glucose and 87.5% xylose. Wells containing 200 µL of medium with no cells were used as control. Microplates with a lid were read at OD_{600nm} at 17-minute intervals for 24 h with a Versamax microplate reader (Molecular devices, USA). Two biological replicates each with two technical replicates were performed.

Chapitre 3 :

Optimisation de la construction
fluorescente

La construction pSB1C3 P_{ihfB} BFP- P_{xylA} RFP utilisée précédemment induit un phénotype de reprise de croissance exacerbé en comparaison avec la souche sauvage. Ce phénotype nous a permis d'identifier le rôle clé de XylR sur la durée de la transition. Il permet aussi d'induire un temps de transition permettant l'étude approfondie de l'état physiologique des cellules lors de cette étape (voir chapitre 4). Enfin, il nous a permis de discriminer facilement et efficacement les sous-populations présentes. Dans un premier temps, cette construction s'est donc avérée extrêmement utile pour explorer le rôle de XylR. Cependant, la construction entraîne une modification du comportement de la souche sauvage avec une phase de transition exacerbée.

L'étude plus approfondie de cette souche nous a montrée qu'elle possède des biais que nous avons jugés acceptables, mais améliorables. Par exemple, nous avons observé une tendance à former des corps d'inclusion en microscopie, notamment lors d'essais de culture en fermenteurs à très haute densité cellulaire (travaux de Julie Taba dans le cadre du projet 3BCAR HME). Un comptage en microscopie nous a permis de quantifier que 12% des cellules présentent des corps d'inclusion lors de la croissance sur glucose et 6.6% sur xylose. Nous avons considéré que ces corps d'inclusions constituaient un biais améliorable. De plus, une légère fuite de fluorescence dans le rouge lors de croissances sur glucose est observable, cette fuite a été jugée acceptable car le niveau de fluorescence est 20 fois supérieur sur xylose (chapitre 1, II). À cause de ces observations, nous avons cherché à optimiser la construction. Pour cela, nous avons inversé le sens de transcription des protéines fluorescentes, afin de s'affranchir de la fuite de fluorescence rouge sur glucose. Afin d'impacter le moins possible la transition et éviter la formation de corps d'inclusion, nous avons cherché à intégrer la construction dans le génome de la souche BW25113 ainsi que dans un vecteur plasmidique à nombre de copies plus faible.

I. Le plasmide pSB1C3 avec inversion du sens de transcription

I.1. Design et Construction du plasmide

Chez la souche portant le plasmide pSB1C3 P_{ihfB} -mTagBFP P_{xyIA} -mRFP1il est possible d'observer une « fuite » de transcription faible, mais qui permet l'expression d'un niveau moyen de fluorescence rouge sur glucose qui n'est pas présent chez la souche contenant le plasmide pSB1C3 P_{xyIA} mRFP1. Dans la construction des nouveaux plasmides, le sens de transcription des gènes codants pour les protéines fluorescentes a été inversé, permettant ainsi de limiter la transcription de la protéine mRFP1 liée à l'activation du promoteur *ihfB* (Figure 32). Une fois inversées, les zones régulant l'expression des protéines se retrouvent accolées. Afin d'éviter toutes interférences entre elles, un terminateur de transcription a été inséré entre les deux (séquence provenant de la brique BBa_0015 de l'iGEM et correspondant au terminateur de transcription du gène *rrnB*) (Figure 32). Cette nouvelle construction a été insérée dans le vecteur plasmidique pSB1C3. Cette construction a été réalisée en utilisant le principe du clonage Gibson (Gibson *et al.*, 2009) (cf Matériel et Méthodes I.2). Le plasmide construit a été vérifié par séquençage et utilisé pour transformer la souche *E. coli* BW25113.

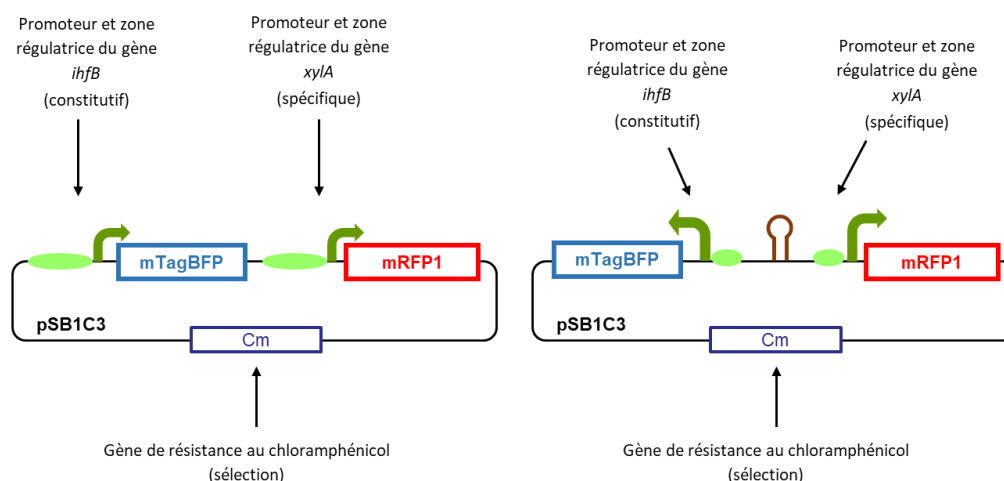


Figure 32 : Schéma du plasmide pSB1C3- P_{ihfB} BFP- P_{xyIA} RFP utilisé précédemment (à gauche) et du nouveau plasmide avec inversion du sens de transcription des protéines fluorescentes (à droite).

I. 2. Test de fluorescence

Cette souche a été cultivée sur M9 glucose et sur M9 xylose, avec pour contrôles les souches contenant les plasmides pSB1C3 P_{xyIA} mRFP1 ou le plasmide pSB1C3 P_{ihfB} -mTagBFP P_{xyIA} -mRFP1. Des échantillons ont été prélevés en phase exponentielle de croissance à $DO_{600nm} = 1$ et analysés en cytométrie en flux. La souche présentant la double fluorescence sans inversion du sens de transcription des deux protéines fluorescentes (pSB1C3 P_{ihfB} -mTagBFP P_{xyIA} -mRFP1) présente bien, sur glucose, un niveau de fluorescence rouge supérieur à la souche contenant le plasmide pSB1C3- P_{xyIA} mRFP1 (avec la fluorescence rouge uniquement) (Figure 33). Les tests de fluorescence sur la souche contenant le plasmide avec inversion de marqueur montrent une réduction du signal fluorescent rouge sur glucose, équivalente à celle observée avec le seul marqueur rouge (Figure 33).

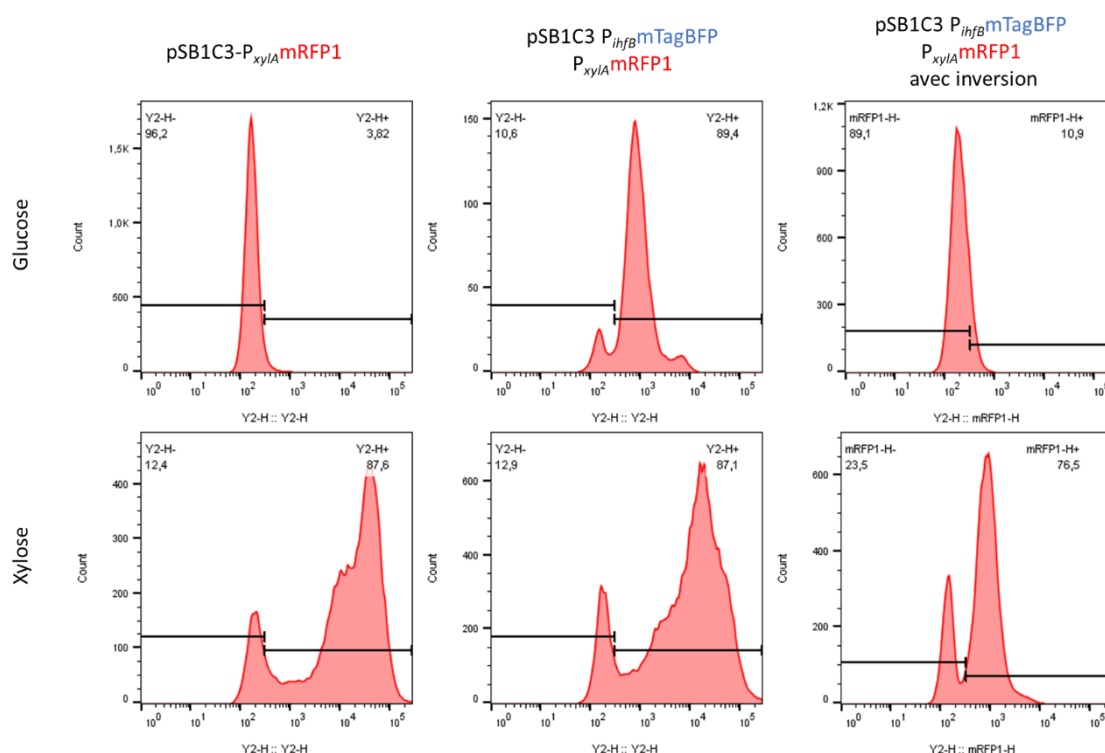


Figure 33 : Histogrammes des niveaux de fluorescence rouge en cytométrie en flux sur le canal Y2-H. L'acquisition des données a été faite avec le MACSQuant VYB. Les prélèvements ont été effectués à partir de cultures des souches contenant les plasmides pSB1C3 P_{xyIA} -mRFP1, pSB1C3 P_{ihfB} -mTagBFP P_{xyIA} -mRFP1, et la nouvelle construction pSB1C3 P_{ihfB} -mTagBFP P_{xyIA} -mRFP1avec inversion lors de cultures sur glucose ou sur xylose. Les barres noires présentent le niveau négatif de fluorescence (Y2-H -) contre le niveau positif de fluorescence (Y2-H +) fixé à l'aide de cellules d'*Escherichia coli* non fluorescentes.

Pour la souche pSB1C3 P_{ihfB} -mTagBFP P_{xylA} -mRFP1 le niveau d'expression de la fluorescence rouge sur xylose est 15 fois supérieur à celui sur glucose. Pour la souche avec inversion du sens de transcription, cette spécificité d'expression sur xylose est toujours présente avec une expression 10 fois supérieure sur xylose que sur glucose.

La construction avec inversion possède l'avantage complémentaire de maintenir un niveau de fluorescence bleue similaire sur glucose et sur xylose, contrairement à la construction sans inversion du sens de transcription où le niveau de fluorescence bleue diminue légèrement sur xylose (Figure 34).

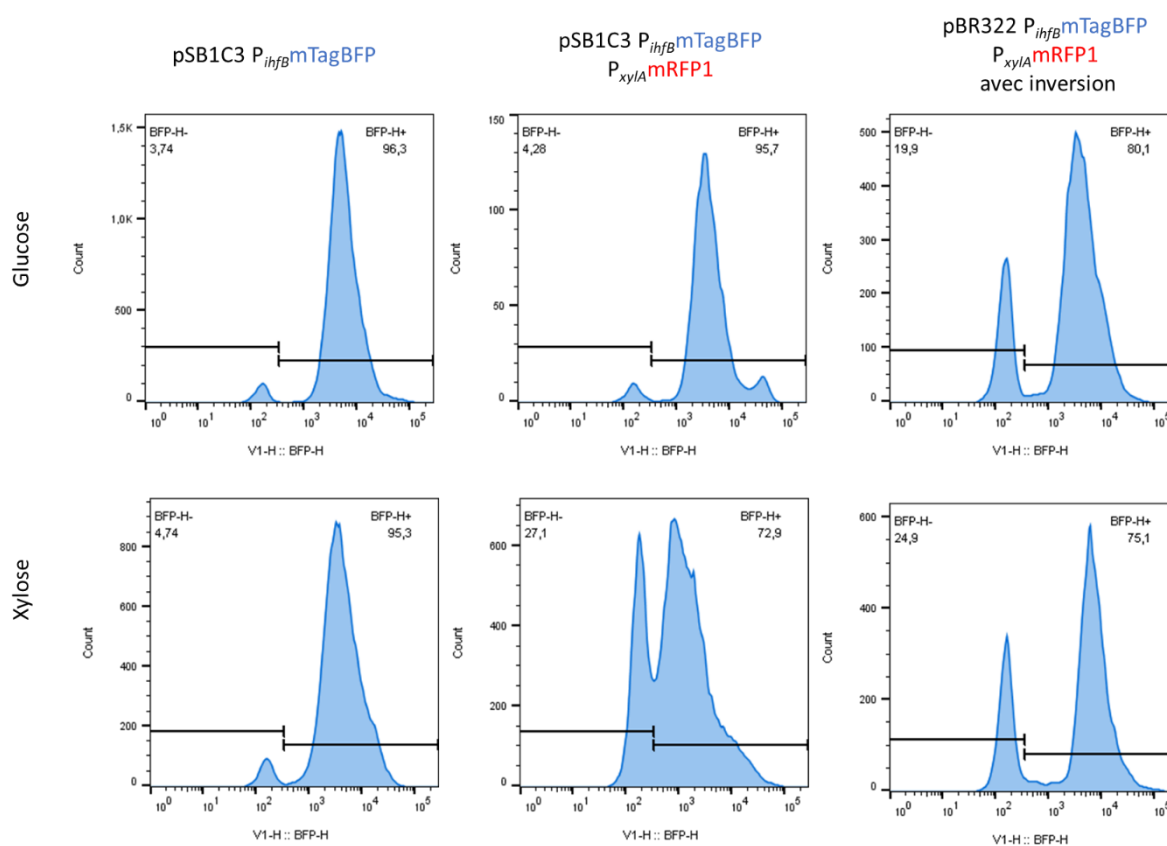


Figure 34 : Histogrammes des niveaux de fluorescence bleue en cytométrie en flux sur le canal V1-H. L'acquisition des données a été faite avec le MACSQuant VYB. Les prélèvements ont été effectués à partir de cultures des souches contenant les plasmides pSB1C3 P_{xylA} -mRFP1, pSB1C3 P_{ihfB} -mTagBFP P_{xylA} -mRFP1, et la nouvelle construction pSB1C3 P_{ihfB} -mTagBFP P_{xylA} -mRFP1 avec inversion lors de culture sur glucose ou sur xylose. Les barres noires présentent le niveau négatif de fluorescence (V1-H -) contre le niveau positif de fluorescence (V1-H +) fixé à l'aide de cellules d'*Escherichia coli* non fluorescentes.

II. Intégration de la construction dans le génome de la souche BW25113

Nous avons montré que l'inversion du sens de transcription des protéines fluorescentes augmente la qualité du signal détecté en cytométrie. Cependant, la construction sur plasmide à haut nombre de copies présente un impact dramatique sur la durée de la transition du glucose au xylose. Sur la base des résultats du Chapitre 2, nous avons cherché à impacter le moins possible cette transition et avons pour cela opté pour une réduction du nombre de sites de fixation du facteur de transcription XylR, et donc une réduction du nombre de copies du promoteur *PxylA*. Pour cela, nous avons cherché à intégrer la construction dans le génome de la souche BW25113.

La construction permettant l'expression de la double fluorescence avec inversion du sens de transcription a été intégrée dans le génome de la souche BW25113 grâce à la méthode *lambda red* (Datsenko and Wanner, 2000). La souche a été transformée avec un plasmide contenant la construction flanquée des zones d'homologies avec la région *aslA* et *aslB* permettant l'intégration spécifique dans cette région du génome (cf Matériel et Méthode I.3).

La bonne intégration de la construction sur le locus génomique choisi a été validée par PCR. Cependant, les tests préliminaires n'ont pas permis de détecter les signaux de fluorescence en cytométrie en flux. Le niveau d'expression des protéines fluorescentes n'est probablement pas suffisant si la construction est présente à une seule copie après intégration génomique.

III. Utilisation d'un plasmide à moyen nombre de copies

III.1. Design et Construction du plasmide

Un compromis entre le nombre de copies trop élevé du pSB1C3 et la trop faible expression de la construction intégrée a été recherché avec un plasmide à moyen nombre de copies.

Nous avons changé le châssis de cette nouvelle construction en l'intégrant dans un vecteur pBR322 modifié. Ce plasmide pBR322 contient une origine de réplication pMB1 et le gène *rop* permettant de maintenir un nombre moyen de copies du plasmide dans la cellule à environ une vingtaine de copies (Lee *et al.*, 2006), bien en deçà du pSB1C3 et ses environ 200 copies dues à son origine de réplication dérégulée. Le vecteur pBR322 originel contient un gène de résistance à l'ampicilline et un gène de résistance à la tétracycline (NEB, ref N3033S). Par souci de cohérence, nous l'avons modifié pour remplacer le gène de résistance à l'ampicilline par le gène de résistance au chloramphénicol et le gène de résistance à la tétracycline a été supprimé pour éviter une charge métabolique inutile.

Ce plasmide a été construit par assemblage Gibson en amplifiant par PCR le vecteur pBR322 modifié et la construction fluorescente à partir du plasmide pSB1C3 *P_{ihfB}-mTagBFP P_{xylA}-mRFP1* portant la construction avec inversion du sens de transcription des protéines fluorescentes. Une fois le plasmide construit, il a été vérifié par séquençage et introduit dans la souche *E. coli* BW25113.

III.2. Test de fluorescence

Pour vérifier la fonctionnalité de la construction sur le plasmide pBR322, cette souche a été cultivée sur M9 glucose ou M9 xylose et un échantillon a été prélevé en phase exponentielle de croissance à $DO_{600nm} = 1$, avec comme contrôles les souches portant les plasmides pSB1C3 *P_{xylA}-mRFP1* et pSB1C3 *P_{ihfB}-mTagBFP P_{xylA}-mRFP1*. Le niveau d'autofluorescence des cellules sauvages a permis de fixer le seuil de différenciation entre les cellules fluorescentes Y2-H⁺ et les cellules non fluorescentes Y2-H⁻. Les résultats pour la fluorescence rouge de la souche contenant le plasmide à moyen nombre de copies avec

la construction, présentés en figure 35, montrent que le signal fluorescent rouge est spécifique au xylose. Aucun signal fluorescent rouge supérieur au niveau d'auto fluorescence des cellules n'est détecté lors de la croissance sur glucose. Lors de croissances sur xylose le signal fluorescent rouge est 100 fois supérieur (figure 35). Cette construction présente une amélioration de la qualité du signal fluorescent par rapport à la souche portant le plasmide pSB1C3 P_{ihfB} -mTagBFP P_{xylA} -mRFP1, le CV (coefficient of variation) est plus réduit, c'est-à-dire qu'il y a moins de disparité intercellulaire sur le niveau de fluorescence, le pic est plus recentré sur la médiane de fluorescence.

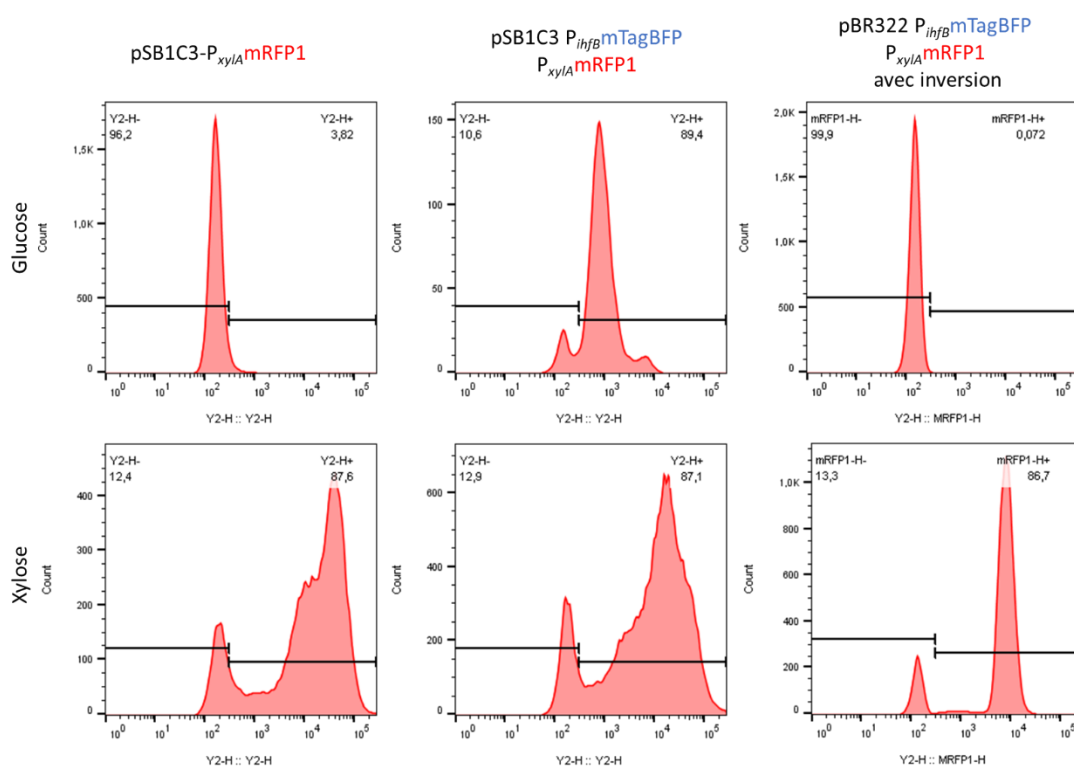


Figure 35 : Histogrammes des niveaux de fluorescence rouge en cytométrie en flux sur le canal Y2-H. L'acquisition des données a été faite avec le MACSQuant VYB. Les prélèvements ont été effectués à partir de cultures des souches contenant les plasmides pSB1C3 P_{xylA} -mRFP1, pSB1C3 P_{ihfB} -mTagBFP P_{xylA} -mRFP1, et la nouvelle construction pBR322 P_{ihfB} -mTagBFP P_{xylA} -mRFP1 avec inversion lors de culture sur glucose ou sur xylose. Les barres noires présentent le niveau négatif de fluorescence (Y2-H -) contre le niveau positif de fluorescence (Y2-H +) fixé à l'aide de cellules d'*Escherichia coli* non fluorescentes.

De plus cette souche présente une forte stabilité du niveau de la fluorescence bleue. En effet, le niveau de fluorescence bleue est similaire sur glucose (1×10^4 u.a.f) et sur xylose (2×10^4 u.a.f) (Figure 36). Cette souche semble donc intéressante pour la suite de l'étude puisqu'elle présente une amélioration de la qualité globale du signal fluorescent.

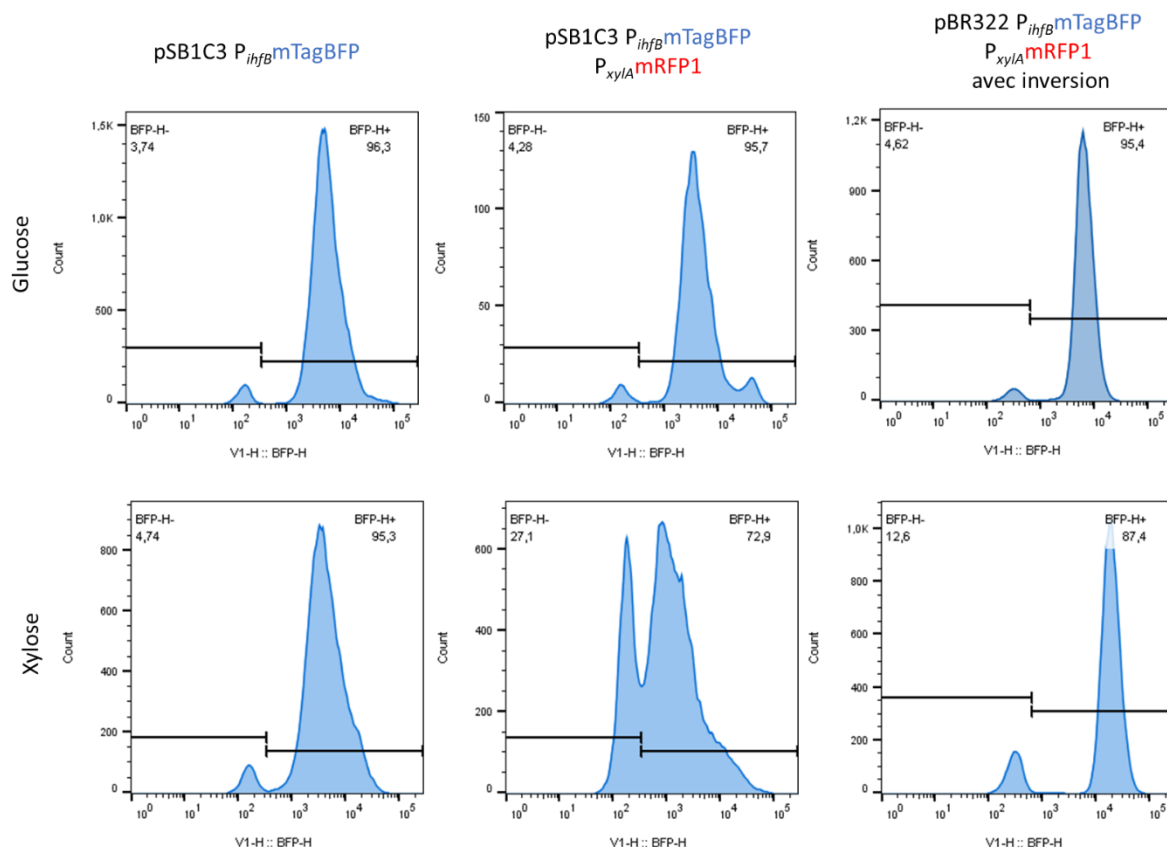


Figure 36 : Histogrammes des niveaux de fluorescence bleue en cytométrie en flux sur le canal V1-H. L'acquisition des données a été faite avec le MACSQuant VYB. Les prélèvements ont été effectués à partir de cultures des souches contenant les plasmides pSB1C3 P_{ihfB}-mTagBFP, pSB1C3 P_{ihfB}-mTagBFP P_{xyIA}-mRFP1, et la nouvelle construction pBR322 P_{ihfB}-mTagBFP P_{xyIA}-mRFP1 avec inversion lors de culture sur glucose ou sur xylose. Les barres noires présentent le niveau négatif de fluorescence (V1-H -) contre le niveau positif de fluorescence (V1-H +) fixé à l'aide de cellules d'*Escherichia coli* non fluorescentes.

III. 3. Test du nombre de copies

L'analyse en cytométrie de la nouvelle souche nous a montré une amélioration globale du signal fluorescent. Afin de comprendre l'origine de cette amélioration, nous avons cherché à quantifier le nombre de copies des différents plasmides dans la cellule. En observant les profils cytométriques, nous soupçonnons que le nombre de copies du pSB1C3 pourrait être variable. Pour cela, nous avons quantifié par qPCR le ratio gène *mRFP1* (présent sur le plasmide) sur gène *xyIA* (présent sur le chromosome). Ce test a été effectué pour la souche contenant le plasmide pSB1C3 contrôle sans promoteur, c'est-à-dire le plasmide pSB1C3 à haut nombre de copies contenant le gène codant pour la mRFP1 mais sans promoteur permettant l'expression de ce gène, la souche contenant l'intégration génomique de la construction, la souche avec la construction double fluorescente pSB1C3 (telle qu'utilisée dans le chapitre 2) ainsi que la souche contenant le plasmide pBR322 avec la double fluorescence en sens inverse de transcription. Les résultats sont présentés dans le tableau suivant (Tableau 3).

	pSB1C3 contrôle sans promoteur	Intégrée	pSB1C3 <i>P_{ihfB}BFP-P_{xyIA}RFP</i>	pBR322 <i>P_{ihfB}BFP-P_{xyIA}RFP</i>
moyenne sur glucose	105.9 +/- 11.7	1 +/- 0.3	166.8 +/- 29.9	15.6 +/- 4.3
moyenne sur xylose	101.2 +/- 15.2	0.9 +/- 0.3	18.9 +/- 1.3	17 +/- 5.5

Tableau 3 : Tableau du nombre de copies de la construction fluorescente par cellule. Ce nombre a été calculé à partir des ratios des résultats qPCR (Ct) : gène *mRFP1* (plasmidique) / gène *xyIA* (génomique) (n=4). Le tableau représente les ratios des Ct obtenus pour le gène *mRFP1* sur les Ct obtenus pour le gène *xyIA*.

Les résultats nous montrent que le plasmide pSB1C3 sans construction est bien présent à haut nombre de copies dans la cellule, une centaine lors des croissances sur glucose comme sur xylose. Ce nombre est en dessous des 200 copies par cellule attendues. Concernant la souche avec l'intégration, il y a bien une seule copie de la construction dans le génome. Cependant, lorsque la construction *P_{ihfB}BFP-P_{xyIA}RFP* est insérée dans le plasmide, on observe une chute drastique du nombre de copies du plasmide sur xylose, mais pas sur glucose. Cette réduction est possiblement liée au coût énergétique de production des deux fluorescences sur xylose. Cette chute du nombre de copies sur xylose

pourrait expliquer la réduction du niveau de fluorescence bleue observée lors de croissances sur milieu contenant uniquement du xylose comme source de carbone. Cependant, il est possible de normaliser par le signal fluorescent bleu afin de corriger cette différence de nombre de copies de plasmide. La souche avec le plasmide pBR322 présente une stabilité du nombre de copies de plasmide sur glucose et sur xylose. En effet, le plasmide pBR322 semble présent avec environ 16 copies par cellule en moyenne sur les deux substrats, ce qui est cohérent avec son origine de répllication décrite pour maintenir le plasmide à 20 copies par cellule environ (Lee *et al.*, 2006). Cette stabilité explique l'amélioration du signal fluorescent, ainsi que le fait que le signal fluorescent bleu soit plus constant sur glucose et sur xylose.

III. 4. Test de l'impact sur la croissance

La diminution du taux de croissance ainsi que du rendement en biomasse associé au substrat de culture sont des marqueurs d'une surcharge métabolique (Heins *et al.*, 2019). Afin de quantifier la surcharge métabolique possiblement due à la présence des plasmides portant les constructions fluorescentes, nous avons calculé ces paramètres pour les souches portant les plasmides pSB1C3 P_{ihfB} -mTagBFP P_{xylA} -mRFP1 et pBR322 P_{ihfB} -mTagBFP P_{xylA} -mRFP1, ainsi que pour la souche sauvage.

	WT	pSB1C3 P_{ihfB} - mTagBFP P_{xylA} - mRFP1	pBR322 P_{ihfB} - mTagBFP P_{xylA} - mRFP1
$Y_{X/S}$ g Biomasse. g^{-1} glucose	0.68	0.58	0.55
$Y_{X/S}$ g Biomasse. g^{-1} xylose	0.58	0.52	0.59
μ_{max} glucose (h^{-1})	0.79	0.63	0.69 +/- 0.01
μ_{max} xylose (h^{-1})	0.45	0.55	0.46 +/- 0.03

Tableau 4 : Tableau des rendements biomasse et taux de croissance pour chaque souche BW25113 avec ou sans plasmide lors de croissances sur glucose ou sur xylose en bioréacteur de 1L contrôlé (n=1 pour BW25113 WT et pSB1C3 ; n=5 pour pBR322)

La construction pBR322 ne semble pas impacter le taux de croissance lors de croissances en xylose, μ_{max} xylose pour la souche sauvage est de $0.45 h^{-1}$ et de $0.46 h^{-1}$ pour la souche portant le plasmide pBR322 avec la construction. De plus le rendement

biomasse est le même pour ces deux souches (respectivement 0.58 g Biomasse. g⁻¹ xylose et 0.59 g Biomasse. g⁻¹ xylose) (Tableau 4). Étonnament la présence du plasmide semble avoir un impact plus important lors de la croissance sur glucose, cependant l'impact sur le taux de croissance est moins important pour la souche portant le plasmide pBR322 avec la construction que pour la souche portant le plasmide pSB1C3 avec la construction (respectivement 0.63 h⁻¹ et de 0.69 h⁻¹).

Ainsi, le plasmide pBR322 semble résoudre la plupart des biais observés avec le plasmide pSB1C3. De plus, nous n'avons pas observé de corps d'inclusion avec la construction sur le chassis pBR322, même à haute densité optique. Il reste à tester le phénotype de la souche contenant le plasmide pBR322 lors de croissances en mélange glucose/xylose afin de regarder si cette construction nous permet d'observer des sous-populations tout en réduisant l'impact sur la physiologie des cellules.

IV. Caractérisation du comportement de la souche pBR322 avec inversion du sens de transcription lors de croissances en mélange glucose xylose

Dans cette partie nous allons caractériser la souche portant le plasmide pBR322 *P_{ihfB}-mTagBFP P_{xyIA}-mRFP1* lors de croissance en mélange et nous allons regarder les profils cytométriques comme effectué dans le chapitre 2 pour la souche portant le plasmide pSB1C3 *P_{ihfB}-mTagBFP P_{xyIA}-mRFP1*.

IV. 1. Caractéristiques de croissance de la souche avec le plasmide pBR322

Nous avons cultivé la souche contenant le plasmide pBR322 avec la construction fluorescente en mélange glucose/xylose afin d'observer l'impact de cette construction sur la physiologie de la souche. Pour cela, la souche BW25113 contenant le plasmide pBR322 avec la construction fluorescente a été cultivée en fermenteurs de 1L contrôlés avec régulation du pH, de l'oxygénation et de la température. Toutes les demi-heures un échantillon a été prélevé afin d'analyser la DO_{600nm}, la consommation des sucres et

l'expression de fluorescence. Les profils de croissance en OD_{600nm} sont présentés en Figure 37.

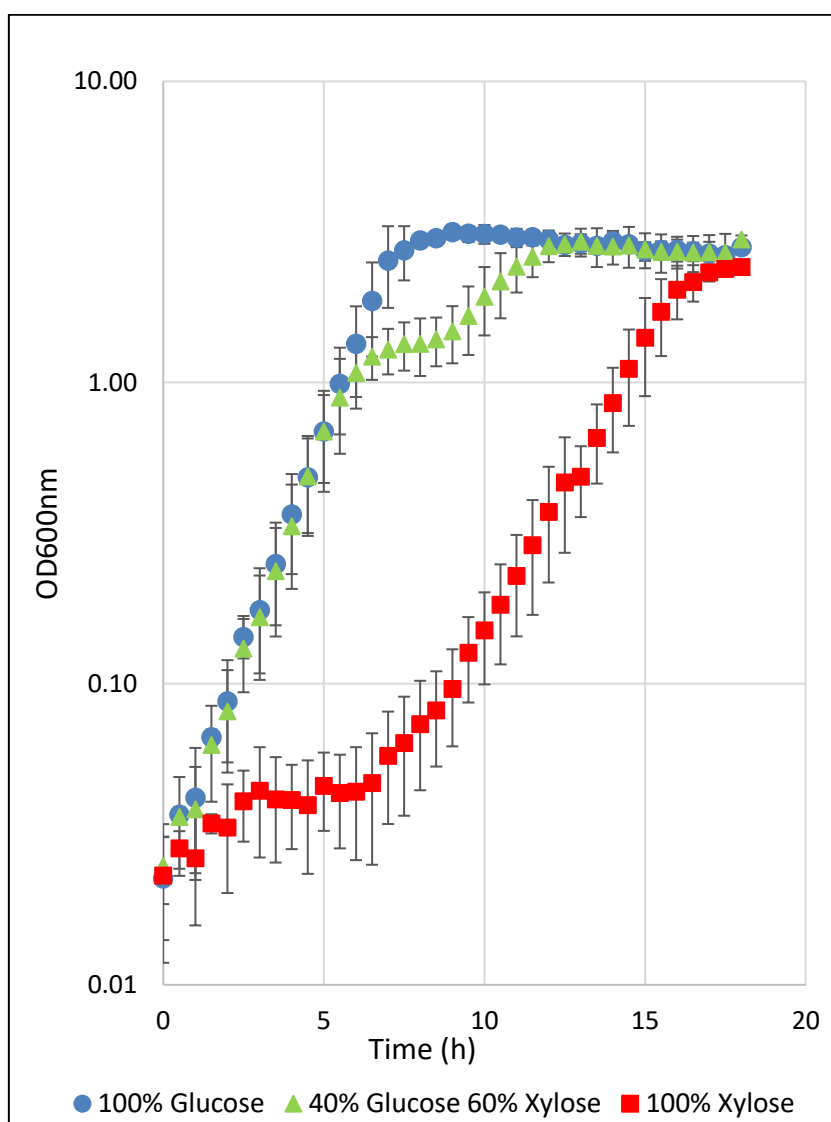


Figure 37 : Profil de croissance en fermenteurs contrôlés de 1L pour la souche BW25113 contenant le plasmide pBR322 sur M9 avec glucose comme unique source de carbone (rond bleu), sur M9 avec xylose comme unique source de carbone (carré rouge) ou sur mélange avec 40% Glucose 60% Xylose (à 90mMEqC en final) (triangle vert), n=4 pour chaque condition.

La souche sauvage BW25113 a été cultivée dans les mêmes conditions, ainsi que la souche BW25113 contenant le plasmide pSB1C3 sans promoteur, soit un plasmide contenant le gène codant pour la mRFP1 mais qui n'est pas exprimé car il n'y a pas de promoteur sur le plasmide. A partir des données de croissances, les temps de latence diauxiques ont été calculés lors de la croissance en mélange 40% glucose 60% xylose ainsi

que le temps de latence avant reprise de croissance sur 100% xylose à partir de cellules en pré-culture sur glucose qui n'ont donc jamais été exposées au xylose.

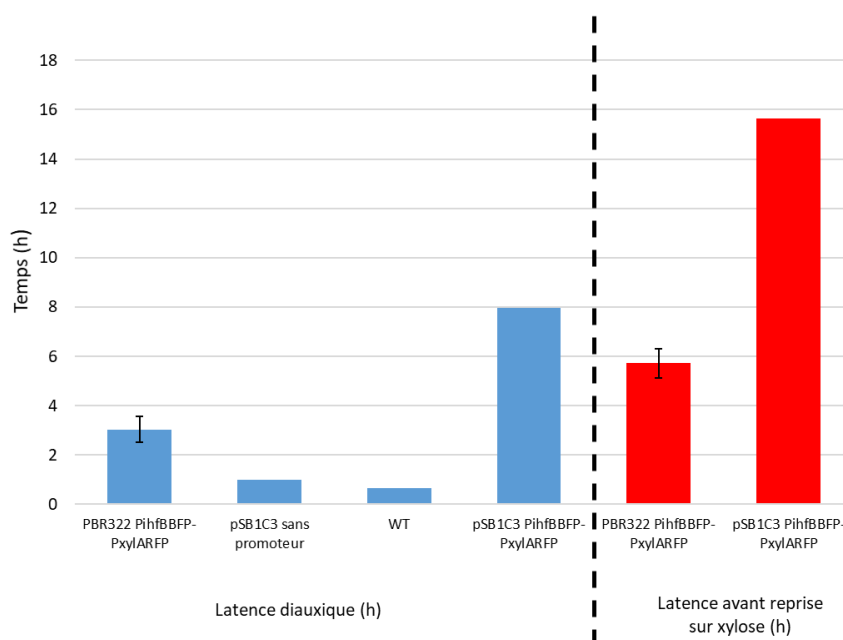


Figure 38 : Temps de latence diauxique pour les différentes souches (en bleu) et temps de latence avant la reprise de croissance sur 100% xylose à partir d'une pré-culture effectuée sur glucose (n=4 pour pBR322 et n=1 pour pSB1C3 avec la construction, pSB1C3 sans la construction et WT).

Ainsi, les cellules contenant le plasmide à moyen nombre de copies (pBR322) ont un temps de diauxie de 3.03 ± 0.53 h alors que la souche contenant le plasmide à haut nombre de copies avait un temps de diauxie de 7.9 h (Figure 38). Les temps de latence de la souche sauvage et de la souche contenant le plasmide vide sont respectivement de 0.65 h et de 1.01 h. Ainsi, comme attendu, l'extension de la latence diauxique par la titration de XylR est moindre dans la souche à moyen nombre de copies en comparaison avec la souche à haut nombre de copies. Ce temps de latence avec le pBR322 reste néanmoins trois fois plus long que celui des souches sans titration. De manière cohérente, le temps de reprise de croissance sur xylose est réduit à 5.72 ± 0.59 h pour la souche avec le pBR322 contre 15.66 h pour la souche avec le pSB1C3. Le temps de latence avant la reprise de croissance sur xylose ou le temps de transition du glucose au xylose pour la souche contenant le plasmide pBR322 avec la construction sont plus longs que chez la souche sauvage mais moins longs que pour la souche contenant le plasmide pSB1C3 avec

la construction. Ce temps intermédiaire est dû à la vingtaine de sites cibles de fixation de XylR présents chez le plasmide à moyen nombre de copies contrairement aux 200 chez le plasmide à haut nombre de copies.

IV. 2. Hétérogénéité de fluorescence de la souche avec le plasmide pBR322

Auparavant, nous avons montré qu'en utilisant une souche avec un nombre de copies élevé du promoteur P_{xylA} , une sous-population émergente est capable de reprendre la croissance sur xylose alors que la sous-population principale semble rester stable et inactive. Le phénotype des sous-populations lorsque l'effet de titration est réduit a donc été exploré de la même manière avec la construction sur le plasmide pBR322.

Nous avons prélevé des échantillons pour la cytométrie lors des cultures présentées précédemment pour la souche BW25113 contenant le plasmide pBR322 $P_{xylAMRFP1-P_{ihfBm}TagBFP$ (Figure 37). Les profils de fluorescence sont présentés avec le niveau de fluorescence inductible rouge, en abscisse, en fonction du niveau de fluorescence constitutive bleue, en ordonné, par cellule. Ce profil de fluorescence est présenté sur la figure 39 pour le 100% glucose et le 100% xylose pendant la phase de croissance exponentielle (premier et dernier dotplot respectivement), et pour le mélange 40% glucose et 60% xylose pendant la cinétique.

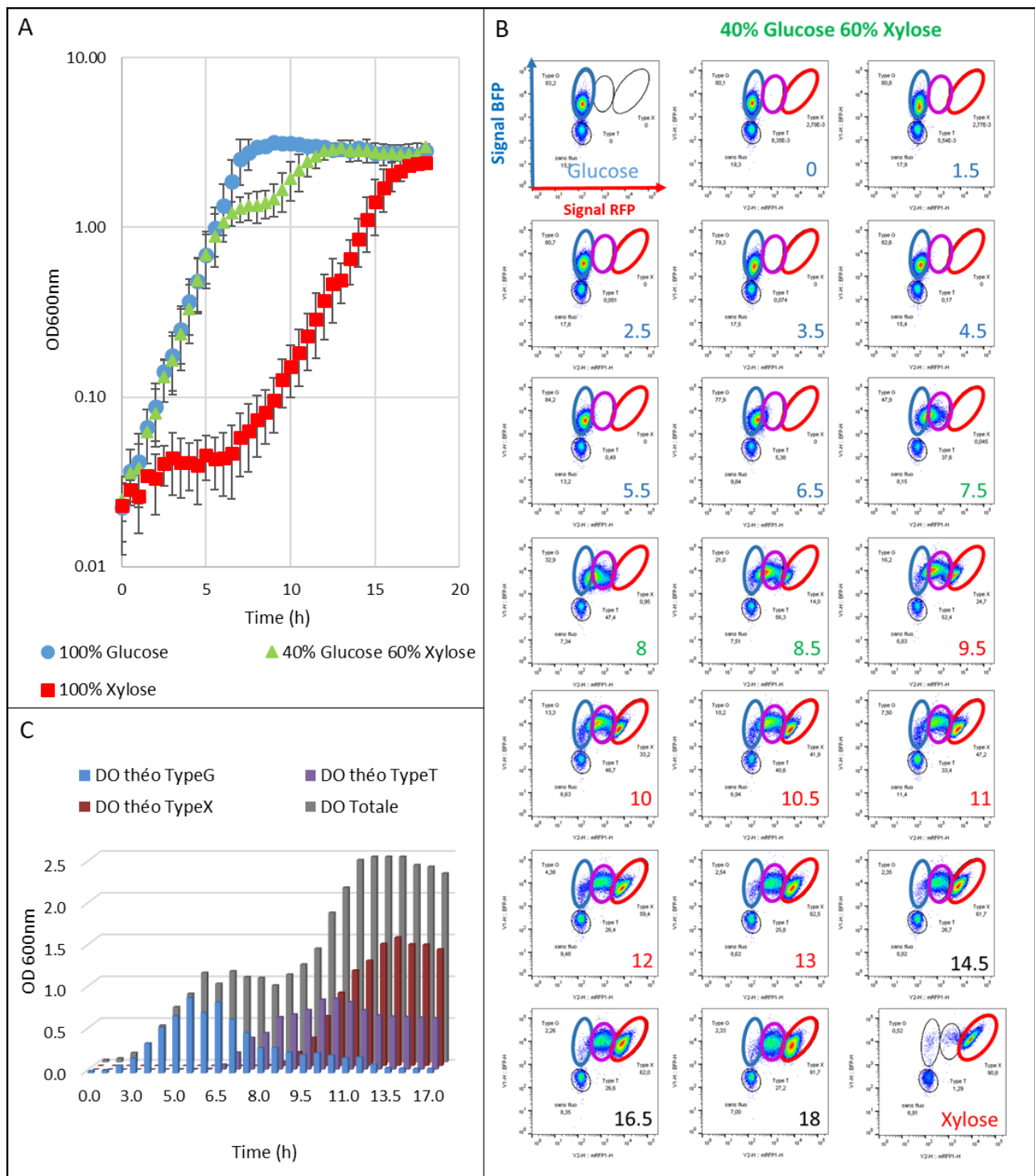


Figure 39 : Légende page suivante

Figure 39: Présence d'hétérogénéité d'expression du marqueur fluorescent rouge. Des fermentations en bioréacteurs de 1L ont été réalisées avec la souche possédant le plasmide pBR322 P_{thfB} -mTagBFP P_{xyIA} -mRFP1, en milieu minimum contenant du glucose, du xylose ou en mélange à 90mMEqC. **A)** Suivi de la production de biomasse en DO600nm en M9 avec glucose comme unique source de carbone (rond bleu), avec xylose comme unique source de carbone (carré rouge), ou sur mélange 40% Glucose 60% Xylose (à 90mMEqC en final) (triangle vert), n=4 pour chaque condition. **B)** Profils cytométriques obtenus lors de la cinétique de croissance avec 40% glucose 60%. L'axe des y affiche les niveaux de fluorescence bleue en unité arbitraire de fluorescence (u.a.f) sur le canal V1-H, et l'axe des x les niveaux de fluorescence rouge (u.a.f) sur le canal Y2-H. Le premier profil cytométrique est celui obtenu lors de croissance exponentielle sur 100% glucose et le dernier sur 100% xylose. Les heures de cultures sont indiquées sur les profils de fluorescence en bleu lors de la consommation de glucose, en vert lors de la consommation d'acétate, et en rouge lors de la consommation de xylose. Le cercle bleu représente la population comme vu avec le contrôle de glucose à 100%. Le cercle rouge représente la population comme vue avec le contrôle 100% xylose et le cercle violet représente la population intermédiaire en niveau de fluorescence rouge observée lors de croissance en mélange. Le cercle noir représente les cellules ne présentant pas de fluorescence. **C)** La biomasse théorique de chaque sous-population a été extrapolée à partir des pourcentages obtenus en cytométrie en flux lors de la croissance sur le mélange 40% glucose et 60% xylose (cellules de type glucose en bleu, cellules de type xylose en rouge, cellules avec un profil intermédiaire en violet et population entière en gris).

Durant la croissance il est possible de distinguer 3 types de profils fluorescents qui ont été appelés type G, type X et type T. Le type G est le profil observable lors de croissances en 100% glucose. Il s'agit d'une population exprimant uniquement le marqueur fluorescent bleu. Le niveau de fluorescence est homogène et présenté dans le cercle bleu (Figure 39). Le type X est le profil de fluorescence observable lors de croissances sur 100% xylose. Il s'agit d'une population exprimant à la fois la fluorescence bleue et la fluorescence rouge, avec un fort niveau de fluorescence rouge. Le niveau de fluorescence est homogène et présenté dans le cercle rouge (Figure 39). Le profil de fluorescence « type G » est observé lors de la première croissance exponentielle sur glucose lors de cultures en mélange 40% glucose et 60% xylose jusqu'à épuisement du glucose à 7 h. Après épuisement du glucose, le motif de fluorescence change pour la majeure partie de la population, la population transite vers un niveau de fluorescence rouge plus élevé, nommé « type T » (dans le cercle violet), mais qui reste inférieur à celui du type X. Simultanément, la sous-population de type X semble apparaître progressivement. Il est difficile de déterminer si la population de type xylose émerge *de novo* ou est due à des cellules du type T augmentant leur expression jusqu'à devenir des

types X. Dans la figure 39, nous avons présenté la biomasse théorique de chaque sous-population extrapolée à partir des pourcentages résultant des données de cytométrie en flux. La population type G disparaît progressivement à la disparition du glucose. La population type X augmente en parallèle de la consommation du xylose, et la population type T augmente directement à la disparition du glucose, avant que le glucose ne soit reconsommé.

V. Conclusion

En conclusion, dans ce chapitre, nous avons pu améliorer la qualité du signal fluorescent grâce à la construction du plasmide pBR322 P_{ihfB} BFP- P_{xylA} RFP avec inversion du sens de transcription des gènes codants pour les protéines fluorescentes. Il n'y a plus de « fuites » de fluorescence rouge sur glucose et le niveau de fluorescence bleue reste constant sur glucose et xylose. Le passage de la construction sur un plasmide à moyen nombre de copies nous a permis de réduire l'impact de la titration de XylR, et ainsi de réduire le temps de transition du glucose au xylose par rapport à la souche portant le plasmide pSB1C3. Cependant ce temps de latence est toujours exacerbé par rapport à la souche BW25113 sauvage. Nous avons aussi pu observer des sous-populations ayant des niveaux de fluorescence différents, cependant avec le plasmide pBR322 la discrimination des sous-populations est moins nette qu'avec le plasmide pSB1C3.

Chapitre 4:

Etude du phénotype des sous-populations observées lors de la transition glucose/xylose

Il a été montré que la souche avec un nombre de copies élevé de la région promotrice P_{xylA} (pSB1C3) présentait un retard diauxique prolongé lié au temps nécessaire à l'émergence d'une sous-population, tandis que la sous-population majeure semblait rester stable. Dans cette condition, la nouvelle sous-population émergente provient d'un petit nombre de cellules à croissance exponentielle et non de l'adaptation de l'ensemble de la population. L'émergence de cette sous-population est influencée par la disponibilité du facteur de transcription XylR. En utilisant un modèle mathématique stochastique, basé sur les données de la souche contenant le plasmide à haut nombre de copies et des données issues de la croissance de la souche de type sauvage, nous avons également établi théoriquement que la quantité de cellules capables de changer de substrat est de 8 760 cellules sur 10 000 par heure dans la souche sauvage.

Il a aussi été possible d'observer deux sous-populations présentant des niveaux de fluorescence rouge différents (population « type T » et population « type X ») chez la souche contenant le plasmide à moyen nombre de copies (pBR322), lors de croissance en mélange glucose/xylose.

Nous avons posé l'hypothèse que les sous-populations observées en fluorescence lors de croissances en mélange glucose/xylose ne présentent pas entre-elles les mêmes caractéristiques phénotypiques. Dans ce chapitre, afin de tester cette théorie, nous avons étudié les comportements des sous-populations à travers leur capacité à redémarrer leur croissance sur glucose ou sur xylose et à travers l'étude de l'expression des gènes durant la transition.

I. Reprise de croissance après tri cellulaire

Les sous-populations définies par leurs niveaux de fluorescence présentent-elles d'autres différences phénotypiques ? Nous avons dans un premier temps observé leur capacité à redémarrer leurs croissances sur différents substrats. Notre hypothèse était que les cellules présentant un fort niveau de fluorescence rouge sont en croissance sur xylose et de ce fait seraient déjà adaptées à ce substrat. Elles devraient donc être capables de redémarrer leur croissance plus rapidement sur ce substrat que les cellules présentant une intensité de fluorescence rouge plus faible. Au contraire les cellules présentant une intensité de fluorescence rouge plus faible qui étaient précédemment en croissance sur glucose pourraient redémarrer plus rapidement leur croissance sur glucose que les cellules s'étant déjà engagées dans l'utilisation d'un autre substrat.

Les souches BW25113 contenant soit le plasmide à haut nombre de copies pSB1C3 P_{ihfBM} TagBFP P_{xylAM} RFP1 soit le plasmide à moyen nombre de copies pBR322 P_{ihfBM} TagBFP P_{xylAM} RFP1 ont été cultivées dans un milieu M9 contenant 40% de glucose et 60% xylose comme sources de carbone. Ces cultures ont été échantillonnées pendant la transition et directement analysées par FACS. Les mêmes profils de fluorescence que ceux décrits précédemment lors de cultures en bioréacteurs ont été obtenus (Figure 39). Lors de la transition du glucose au xylose avec la souche contenant le plasmide pBR322, les cellules présentent uniquement de la fluorescence bleue lors de la phase de croissance sur glucose. Puis à la transition nous avons observé une sous-population à haut niveau de fluorescence rouge appelée Population « type X », qui présente le même niveau de fluorescence rouge que les cellules en croissance sur xylose comme seule source de carbone. Nous avons également observé une population avec un niveau intermédiaire de fluorescence rouge qui apparaît lors de la transition (appelée Population « type T » ; voir chapitre 3; Figure 40). Pour la souche contenant le plasmide pSB1C3, nous pouvons observer deux sous-populations (décrites dans le chapitre 2 ; Figure 40). La population 1 présentait le même motif de fluorescence que le contrôle 100% glucose et la population 2 présentait un motif similaire à celui obtenu sur le contrôle 100% xylose. Les profils de fluorescence pour la souche pBR322 présents au cours de la cinétique sont décrits dans le chapitre 3 et ceux pour la souche pSB1C3 dans le chapitre 2, les profils ont été rajoutés en figure 40 pour rappel.

Ainsi, durant la transition chaque souche présente deux sous-populations présentant des niveaux de fluorescences différents (Figure 40). Notre hypothèse de départ était que la population de « type T » et la population 1 sont identiques et que la population de « type X » et la population 2 sont identiques. Ces différentes populations ont été triées par FACS et ont été transférées en microplaque (volume final de 200 μ L par puits), contenant du milieu M9 avec soit du glucose soit du xylose comme unique source de carbone. Les puits de la microplaque ont été inoculés avec une seule cellule. La même expérience a été réalisée pour les deux souches. Les plaques inoculées ont été incubées pendant au moins 78 h à 37°C sous agitation et la DO_{600nm} a été mesurée toutes les 17 min. Pour chaque souche, l'expérience a été réalisée deux fois.

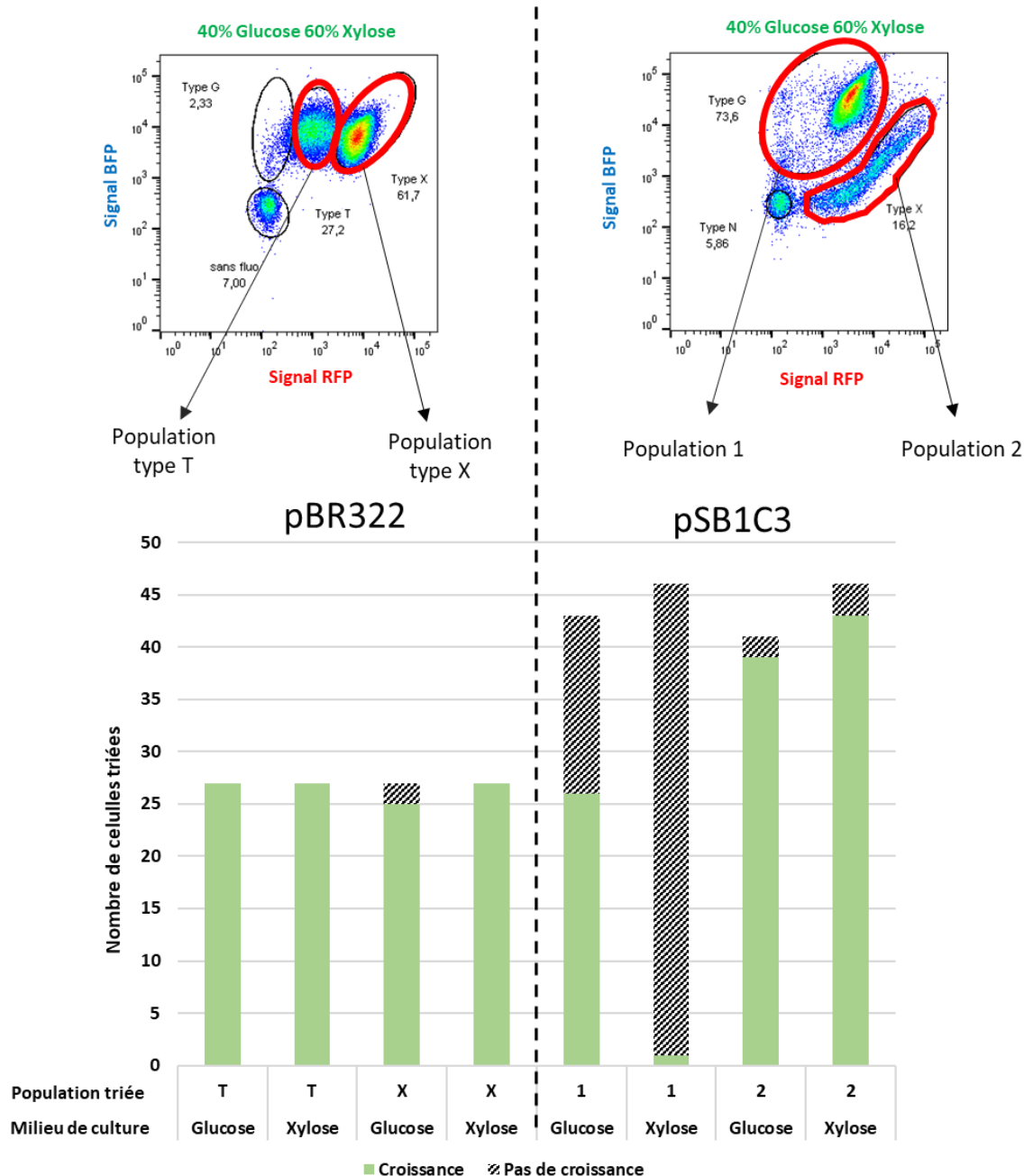


Figure 40 : A) Profils cytométriques pour les souches portant le plasmide pBR322 $P_{ihfBm}TagBFP$ $P_{xylAm}RFP1$ ou pSB1C3 $P_{ihfBm}TagBFP$ $P_{xylAm}RFP1$ durant la transition du glucose au xylose. Les sous-populations T et X sont représentées par des cercles rouges pour la souche BW25113 portant le plasmide pBR322, de la même manière les populations 1 et 2 sont représentées par des cercles rouges pour la souche portant le plasmide pSB1C3. Les sous-populations ont été triées pendant la transition et inoculées dans 200 μ L de M9 glucose ou de M9 xylose. B) Histogramme de la capacité de reprise de croissance après le tri des cellules, au moins 27 cellules ont été triées par conditions. Les puits où une croissance est détectable après 78 h sont représentés en vert et les puits sans croissance détectée sont représentés en noir hachuré. L'expérience a été réalisée en duplicat pour chaque souche.

Nous pouvons observer que, pour la souche contenant le plasmide à moyen nombre de copies, les deux populations (T et X) sont capables de redémarrer leur croissance à la fois sur glucose et sur xylose (Figure 40).

Nous avons calculé les temps de latence avant la reprise de croissance pour chaque puits où une croissance a été détectée. Le calcul du temps de latence a été effectué comme expliqué dans (Enjalbert *et al.*, 2015). La croissance d'une seule cellule n'est pas détectable en spectrophotométrie, de ce fait le calcul du temps de latence se fait à partir d'une extrapolation de la croissance observée afin de déterminer le moment où la cellule a commencé à croître. Du fait de la longue période entre inoculation d'une cellule unique et observation de la croissance, le calcul s'avère peu précis et des temps de latence négatifs peuvent apparaître, signifiant néanmoins une reprise de croissance très rapide. Pour la construction pBR322, les deux populations ont immédiatement redémarré leur croissance sur xylose avec un temps de latence médian de -1,3 h +/- 2,8 h pour la population T et -2,9 h +/- 2,4 h pour la population X (Figure 41). Les populations T et X semblent avoir besoin de plus de temps pour redémarrer la croissance sur glucose, avec des temps de latence médians de respectivement 4,9 h +/- 2,15 h et 5,4 h +/- 2,3 h (Figure 41). Les deux sous-populations présentent donc les mêmes caractéristiques de reprise de croissance.

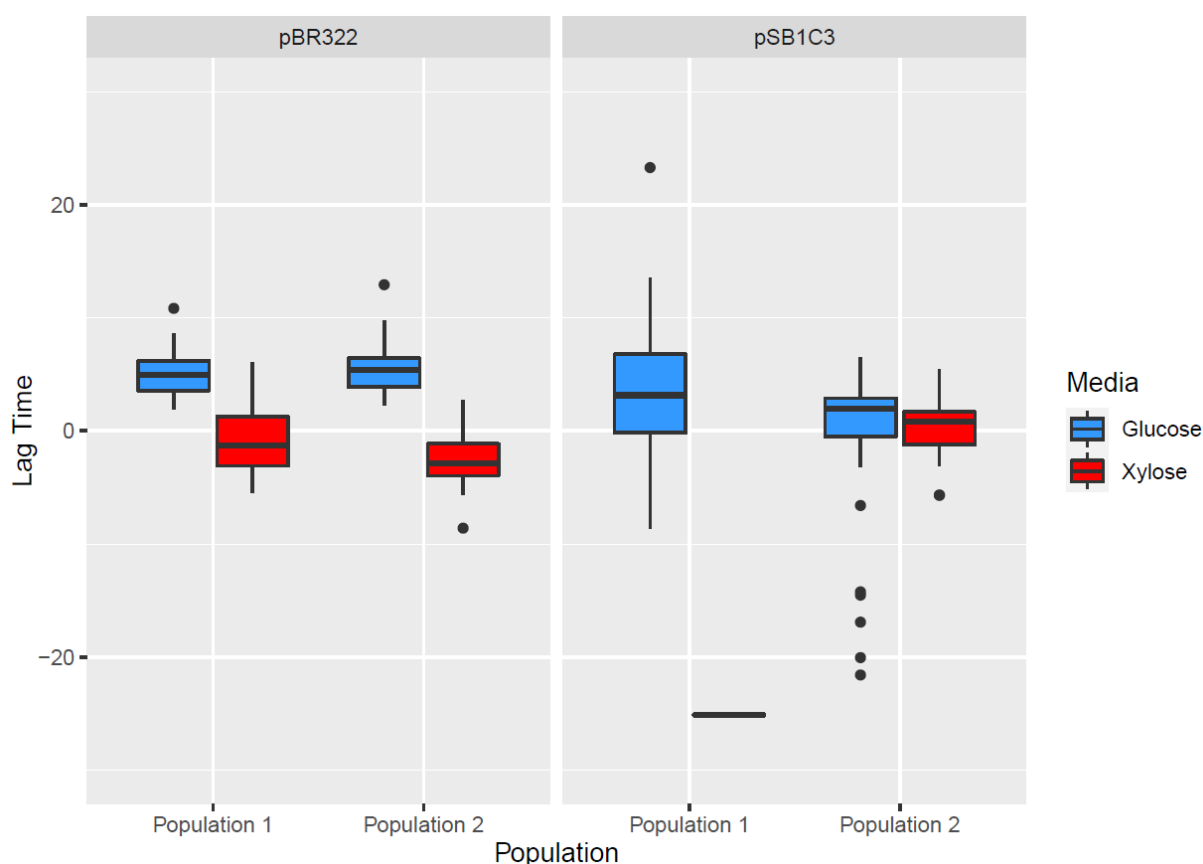


Figure 41 : Boîtes des temps de latence après le tri cellulaire dans différents milieux pour la souche BW25113 portant le plasmide pBR322 ou le plasmide pSB1C3. Pour chaque souche, au moins 27 cellules de chaque sous-population ont été triées pendant la transition et inoculées en microplaque 96 puits contenant 200 μ L du M9 glucose ou de M9 xylose. Le temps de latence avant la reprise de croissance a été extrapolé comme expliqué dans le Matériel et Méthode. Les temps de latence avant la reprise de croissance sur le glucose sont représentés en bleu et les temps de latence avant la reprise de croissance sur le xylose sont représentés en rouge.

A l'opposé, pour la souche contenant le plasmide à haut nombre de copie (pSB1C3), les deux populations ne présentent pas les mêmes capacités de reprise de croissance. La population 1 ne semble pas être en mesure de relancer sa croissance sur milieu contenant uniquement du xylose comme source de carbone (une seule cellule a redémarré sur 46 inoculées) (Figure 40). Cette population semble également avoir des difficultés à relancer sa croissance sur glucose (seules 26 cellules ont redémarré la croissance sur 43). Pour cette souche, la population 2 ne présente pas les mêmes difficultés : 39 cellules sur 41 sont capables de repartir en croissance sur glucose et 43 sur 46 sur xylose (Figure 40). En analysant les temps des latences (Figure 41), nous pouvons remarquer que la population 2 (type xylose) a en moyenne rapidement redémarré sa

croissance sur xylose (0,85 h +/- 7,75 h de temps de latence) comme sur glucose (1,86 h +/- 13,19 h) (Figure 41). Cependant, les temps de latence sont très variables entre les réplicats.

Les expériences de reprise de croissance après tri cellulaire ont montré que pour la souche portant le plasmide pSB1C3 P_{ihfB} TagBFP P_{xylA} mRFP1 à haut nombre de copies, les sous-populations présentes lors de la transition du glucose au xylose n'ont pas les mêmes capacités de reprise de croissance. Ainsi, la population 2, présentant le même signal fluorescent que les cellules en croissance sur xylose, sont capables de redémarrer leur croissance rapidement sur les deux substrats (glucose ou xylose), au contraire la population 1 qui semble présenter des difficultés à redémarrer sa croissance sur glucose et ne semble pas capable de croître sur xylose. Au contraire, pour la souche contenant le plasmide à moyen nombre de copies avec la construction fluorescente les deux sous-populations présentent exactement les mêmes caractéristiques de reprise de croissance. Les différences entre ces souches sont donc probablement dues à la différence du nombre de copies du plasmide, et donc, à un niveau différent de titration de XylR pour chaque souche. Le fort effet de titrage réduit la capacité des cellules à s'adapter à un autre substrat. De plus, durant une croissance en mélange glucose/xylose le temps de diauxie est de 3 h pour la souche contenant le plasmide pSB1C3 et de 8 h pour la souche contenant le plasmide pBR322. Le fait de posséder un nombre supérieur de cible de fixation plasmidique de XylR retarde la fixation de XylR sur le site cible génomique nécessaire à l'induction de l'expression des gènes de consommation du xylose. La présence du plasmide à haut nombre de copies portant les deux marqueurs fluorescents pourrait aussi augmenter la charge métabolique, contribuant à vider prématurément les ressources internes (notamment les réserves énergétiques) des cellules n'ayant pas encore transité sur le second substrat. Une autre possibilité est l'induction de la réponse stringente du fait de la charge métabolique accrue menant à un arrêt de la transcription ou de la traduction et ainsi induire des difficultés à redémarrer la croissance sur glucose. Cette particularité semble abolie dans la souche à moyen nombre de copies de plasmide.

II. Analyse transcriptomique lors de culture en mélange glucose/xylose

Au cours de l'étude de la souche portant le plasmide pBR322, nous avons montré que les deux sous-populations avec différents niveaux de fluorescence rouge présentaient les mêmes caractéristiques de croissance après le tri cellulaire (cellules de types T et de types X). Nous nous sommes demandés si les sous-populations sont vraiment distinctes avec des caractéristiques phénotypiques différentes ou bien si elles ne diffèrent que par l'expression de la fluorescence.

Pour répondre à cette question, nous avons dans un premier temps prévu d'effectuer une analyse transcriptionnelle des sous-populations. Pour cela, il est nécessaire de trier les cellules des différentes sous-populations avant l'extraction de l'ARN. La première limitation était le nombre de cellules nécessaires. L'analyse RNAseq nécessite au moins 200 ng d'ARN total. Un travail antérieur dans le cadre du projet 3BCAR CYTOM2 au sein du laboratoire a montré que 10^7 cellules d'*E. coli* sont nécessaires pour obtenir au moins 300 ng d'ARN total en utilisant le kit RNEASY Micro Kit (Qiagen), 10^8 étant une quantité souhaitable. Le FACS peut analyser entre 20 000 et 30 000 événements par seconde. Avec cette vitesse théorique, il est possible de trier 10^7 cellules en 8,3 min et 10^8 cellules en 1 h 20 min. Plus la population sera en faible quantité plus le temps de tri sera long, dans le cas de sous-populations représentant jusqu'à 25% de la population totale, cela pourrait prendre 25 min pour 10^7 cellules mais 5 h 20 min pour 10^8 cellules. Un temps de tri aussi long est hasardeux pour une analyse transcriptomique du fait du temps excessif passé dans le FACS à 4°C qui peut conduire à la dégradation de l'ARN dans les cellules. Ainsi, nous avons opté pour l'accumulation de tris de moins d'une heure en cherchant à minima 10^7 cellules.

Pendant la fermentation, des échantillons ont été prélevés et congelés dans l'azote liquide afin d'immédiatement stopper le métabolisme des cellules. Nous avons trié les cellules à un temps de cinétique où chaque sous-population représente environ 50% de la population totale. Le trieur a été maintenu 4°C et les cellules ont été diluées dans du PBS frais à 4°C. Nous avons effectué 6 tris différents sans dépasser 1 h de temps par tri, pour un total de 7.28×10^7 cellules pour la population de type T et de 9.28×10^7 cellules pour la population de type X. Les cellules ont été centrifugées à 16 000 g pendant 15 min

immédiatement après le tri et le culot théorique a été congelé (théorique car une si faible quantité de cellules n'est pas visible dans un tube Falcon de 50 ml, ce qui rend l'ensemble du processus hasardeux). Les cellules ont été décongelées dans la glace et les six lots par tri ont été centrifugés ensemble. L'extraction d'ARN a été réalisée avec le kit RNEasy Micro (Qiagen), puis nous avons quantifié les rendements par Nanodrop et Qubit et vérifié la qualité grâce au Bioanalyzer (Agilent). Pour la population de type T, aucun ARN n'a été détecté au Qubit et pour la population X, nous avons obtenu 18,1 ng / μL d'ARN mais cet ARN est apparu comme dégradé après analyse au Bioanalyzer (voir profil Figure 42). Nous avons essayé d'optimiser notre protocole, cependant des quantités et qualités d'ARN acceptables pour effectuer l'analyse en RNAseq n'ont pas pu être obtenus, ce qui nous a forcé à renoncer à une telle approche.

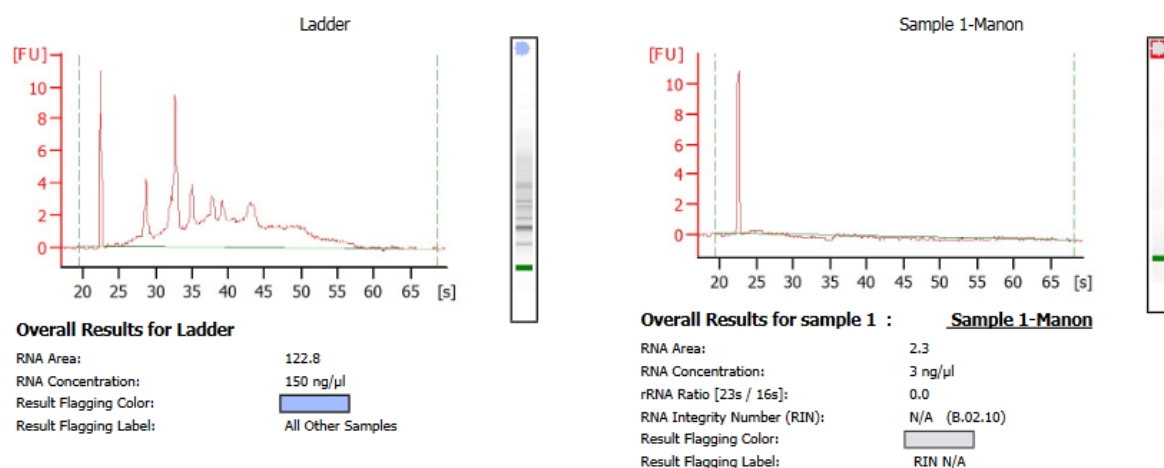


Figure 42 : Profil en électrophorèse capillaire des ARN extraits de la population X, le premier profil est un contrôle et présente le profil attendu si les ARN extraits ne sont pas dégradés, le second profil est celui obtenu avec un échantillon des ARN extraits de la population X, il n'y a qu'un seul pic les ARN sont dégradés.

L'option d'effectuer une analyse transcriptionnelle des sous-populations n'étant pas possible en raison de problèmes techniques, nous avons choisi d'effectuer une analyse de la population globale pendant la transition du glucose au xylose. Premièrement, il n'y a pas de données transcriptomiques pour cette transition dans la littérature. Deuxièmement, l'objectif était d'analyser les données RNAseq à travers un point de vue de sous-populations en utilisant les données de cytométrie pour essayer de déconvoluer les profils des sous-populations.

Une autre option aurait été d'essayer d'effectuer une analyse RNAseq sur cellules unique (single cell RNAseq) afin de s'affranchir du temps nécessaire pour trier 10^7 cellules

(Tang *et al.*, 2010; Imdahl *et al.*, 2020). Cependant la méthodologie n'est pas développée dans notre laboratoire et le coût aurait été important car de nombreux réplicats sont nécessaires pour que l'analyse soit robuste.

II. 1. Choix des échantillons :

Nous avons suivi la culture sur mélange glucose/xylose. Pendant, la fermentation, plusieurs échantillons ont été prélevés : un échantillon pendant la phase de croissance exponentielle sur glucose (M1), un échantillon pendant la phase de latence entre les deux substrats (M2), un autre au début de la reprise de croissance sur xylose (M3), et un dernier lors de la phase de croissance sur xylose (M4). Des croissances uniquement sur glucose ou xylose ont été réalisées pour disposer de contrôles en phase exponentielle sur glucose (G1) ou xylose (X1), ou en phase stationnaire après la croissance sur glucose (G2) (Figure 43). Trois répétitions indépendantes ont été effectuées.

Les profils de fluorescences pour la souche pBR322 présente au cours de la cinétique sont les mêmes que décrits dans le chapitre 3.

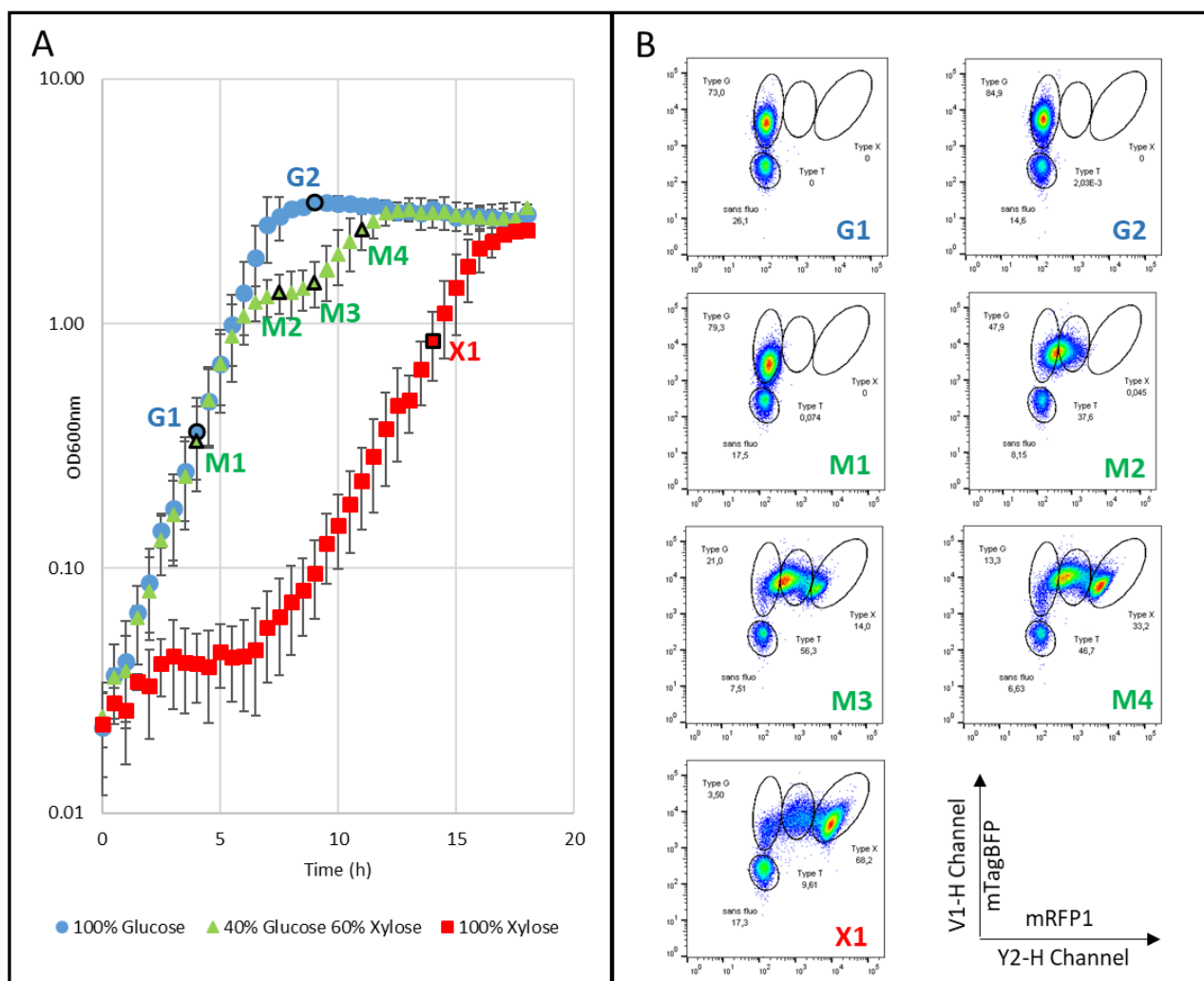


Figure 43 : Échantillonnage de l'analyse RNAseq. A) Courbes de croissance de la souche BW25113 contenant le plasmide pBR322 P_{ihfB} mTagBFP P_{xyA} mRFP1. Les cellules ont été pré-cultivées dans un milieu M9 glucose en Erlenmeyer, et ont été transférées dans un bioréacteur contrôlé de 1L contenant 100% de glucose (ronds bleus), 100% de xylose (carrés rouges) ou un mélange de 40% de glucose à 60% de xylose (triangles verts) à 90 mMeqC ($n = 4$ pour les cultures ; $n = 3$ pour les prélèvements ARN). Les échantillons utilisés lors de l'analyse transcriptomique sont entourés par des cercles noirs, les extractions ont été réalisées en triplicats. B) Profils cytométriques des échantillons obtenus à partir d'un des réplicats biologiques, les cercles noirs représentent les cellules sans fluorescence, de type G, de type T et de type X comme expliqués dans le chapitre 3.

Les échantillons des cellules ont été traités pour extraire l'ARN grâce à un traitement trizol/phénol. Puis des ERCC ont été rajoutés comme contrôle pour la normalisation, avant ribodéplétion afin d'éliminer les ARN ribosomiques. L'analyse RNAseq a été effectuée en IonTorrent (Ion S5, Thermofischer) à la plateforme GETbiopuce. Les données ont été normalisées grâce au package R suivant la méthodologie TMM_exact développée dans l'équipe.

II.2. Etude de l'expression génique lors de la transition glucose/xylose à l'échelle de la population

II.2.A Description générale des résultats d'expression

Dans un premier temps, chaque condition est comparée à la condition témoin G1 (phase exponentielle de croissance sur glucose). Pour chaque condition, les gènes avec une expression différentielle ont été identifiés avec une $p\text{-valeur} < 10^{-5}$ et un seuil de variation de niveau d'expression de 4 pour les gènes régulés positivement et 0,25 pour les gènes régulés négativement.

La condition G2 vs G1 est la comparaison entre la phase de croissance exponentielle et la phase stationnaire. Dans cette comparaison, le niveau d'expression d'un nombre important de gène varie (Figure 44a).

X1 vs G1 est la comparaison entre la phase exponentielle de croissance sur glucose et la phase exponentielle de croissance sur xylose. Le nombre de gène dont l'expression varie est faible (Figure 44b). Seul 42 gènes sont différentiellement exprimés (Tableau 5), ce qui est assez cohérent car il s'agit de la comparaison de deux phases exponentielles de croissance où les taux de croissance sont élevés (0.69 h^{-1} sur glucose et 0.47 h^{-1} sur xylose) et que la voie métabolique du xylose ne diffère que très peu par rapport à celle du glucose.

Nous n'avons observé aucun gène différentiellement exprimé en comparant M1 à G1 (Figure 44c), ce qui signifie que pendant la phase exponentielle sur glucose, les cellules ne sont pas impactées par la présence de xylose dans le milieu. Cette observation est cohérente avec le fait que la répression catabolique est présente dans les deux conditions, et qu'aucune sous-population n'a été observée pendant la phase de croissance sur glucose. Ainsi, la présence de xylose dans le milieu n'induit pas de réponse transcriptionnelle lors de la croissance sur le glucose. Les données M1 et G1 sont très corrélées (Figure 44c).

A l'inverse un grand nombre de gènes sont différentiellement exprimés en M2 comparé à G1 (Figure 44d), puis ce nombre diminue en M3 comparé à G1 et en M4 comparé à G1 (Figure 44e, 44f).

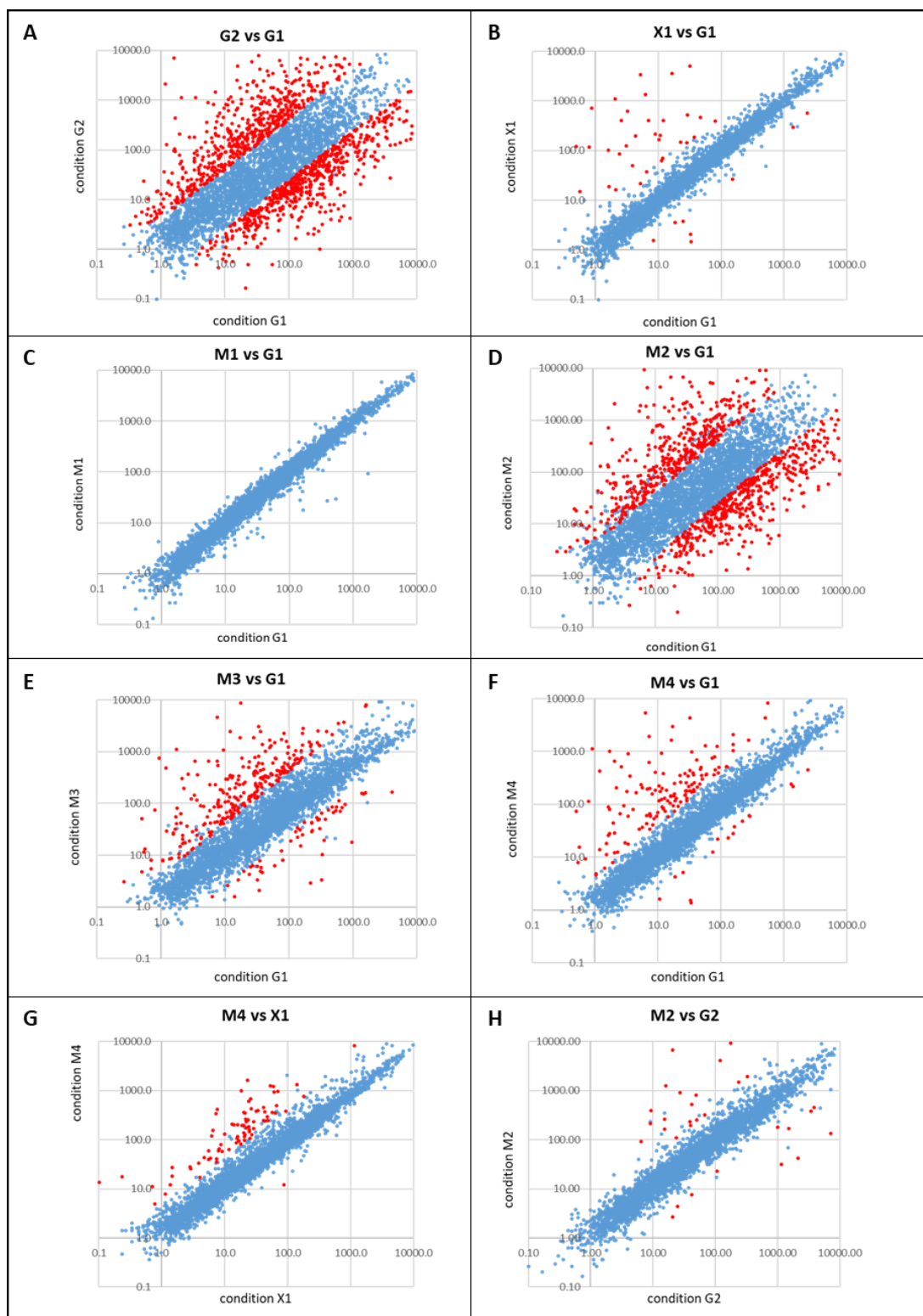


Figure 44 : Légende en page suivante

Figure 44 : Comparaison des niveaux d'expression des gènes d'*Echerichia coli* obtenus en RNA seq pour les différentes conditions. Pour chaque condition n=3. En rouge sont représentés les gènes différentiellement exprimés avec une p- value $<10^{-5}$ et un seuil à 4 pour les gènes régulés positivement et 0,25 pour les gènes régulés négativement. A) Comparaisons des conditions G2 (phase stationnaire après disparition du glucose) et G1(en phase exponentielle de croissance sur glucose). B) Comparaisons des conditions X1 (phase exponentielle de croissance sur xylose) et G1. C) Comparaisons des conditions M1 (croissance exponentielle sur glucose avant la transition) et G1. D) Comparaisons des conditions M2 (début de la transition après disparition du glucose) et G1. E) Comparaisons des conditions M3 (fin de la transition) et G1. F) Comparaisons des conditions M4 (croissance sur xylose après transition) et G1 G) Comparaisons des conditions M4 et X1 H) Comparaisons des conditions M2 et G2

Nous avons comparé M4 et X1 qui sont deux phases de croissance sur xylose, avec X1 en phase exponentielle de croissance lorsque le xylose est la seule source de carbone et M4 en croissance sur xylose 2 h après la fin de la transition du glucose au xylose. Les deux conditions sont similaires mais 80 gènes sont différentiellement exprimés (Figure 44g).

Nous nous sommes aussi intéressés à la différence potentielle entre un arrêt de croissance lié à une carence en carbone (G2) et un arrêt de croissance lié à un changement de source de carbone utilisée. L'analyse de l'expression des gènes entre les conditions M2 vs G2 (respectivement phase stationnaire après épuisement du glucose lorsque le glucose était la seule source de carbone, et début de la transition après épuisement du glucose lors de la croissance sur mélange glucose/xylose), a montré que ces conditions sont très semblables (Figure 44h).

II. 2. B Analyse fonctionnelle des gènes différentiellement exprimés

Nous avons estimé le nombre de gènes différentiellement exprimés dans chaque comparaison avec les critères statistiques définis plus haut (une p-valeur $< 10^{-5}$ et un seuil de variation de niveau d'expression de 4 pour les gènes régulés positivement et 0,25 pour les gènes régulés négativement) et répertoriés les résultats en Tableau 5.

	G2/G1	X1/G1	M1/G1	M2/G1	M3/G1	M4/G1	M2/G2	M4/X1
Nombre de gènes surexprimés	566	34	0	448	260	137	20	79
Nombre de gènes réprimés	595	8	0	503	54	16	11	1

Tableau 5 : Nombre de gènes différentiellement exprimés pour les comparaisons G2 vs G1, X1 vs G1, M1 vs G1, M2 vs G1, M3 vs G1, M4 vs G1, M2 vs G2 et M4 vs X1.

Ainsi les conditions M1 et G1 sont identiques, ainsi des cellules en phase exponentielle de croissance sur glucose ne sont pas impactées par la présence de xylose. De plus les conditions d'arrêt de croissances M2 et G2 semblent similaires (seulement 31 gènes différentiellement exprimés) avec une réponse transcriptomique très différente de celle en G1. De plus les phases exponentielles de croissances sur glucose et sur xylose comme seule source de carbone sont similaires et présentent uniquement 42 gènes différentiellement exprimés. La réponse transcriptomique des cellules en croissance exponentielle sur xylose (X1) et de cellules en croissance sur xylose après transition (M4) et similaire mais pas identique nous retrouvons 80 gènes différentiellement exprimés.

Le nombre de gènes différentiellement exprimés par rapport à G1 et très important au début de la transition (M2) puis semble diminuer au cours de la cinétique (M3 puis M4). Nous avons comparé les gènes surexprimés et réprimés par rapport à G1 pour les conditions X1, M1, M2 M3 et M4 afin de regarder si la diminution du nombre de gènes différentiellement exprimés au cours de la cinétique est liée à une atténuation de la réponse observée en M2 (Figure 45).

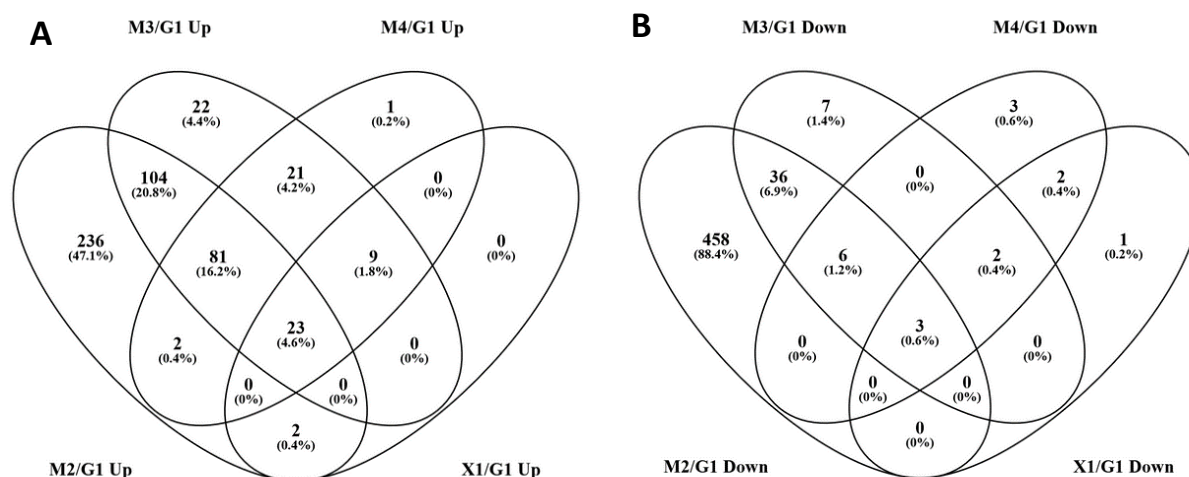


Figure 45 : Comparaison des conditions X1, M1, M2, M3 et M4. A) Diagramme de Venn des gènes surexprimés en conditions X1, M1, M2 M3 et M4 par rapport à G1. B) Diagramme de Venn des gènes réprimés dans les conditions X1, M1, M2 M3 et M4 par rapport à G1. Les données ont été obtenues par analyse RNA seq, les gènes sont différenciellement exprimés avec une p-value associé $<1,10^{-5}$ et un seuil de coupure à 4 pour les gènes positivement régulés et 0,25 pour les gènes négativement régulés.

Il y a 448 gènes surexprimés en M2 par rapport à G1, 208 de ces gènes sont encore surexprimés en M3 par rapport à G1 et plus que 106 en M4 par rapport à G1 parmi ces 106 gènes 23 sont aussi surexprimés en X1 par rapport à G1 (Figure 45). Nous observons donc une diminution de la réponse mise en place en M2 au cours de la cinétique.

Nous avons étudié les gènes qui étaient surexprimés par rapport à G1 dans les conditions de croissance en mélange glucose/xylose (M2, M3 et M4) mais pas lors de croissance sur xylose seul (X1). Ces gènes ont été regardés par rapport à G1 pour faciliter la comparaison au cours de la transition. Ces gènes représentent 16% des gènes (81 gènes au total) et sont surexprimés uniquement en condition de croissance en mélange (M2, M3 ou M4) (Figure 45). Nous avons réalisé un enrichissement fonctionnel (en Gene ontology) sur ces 81 gènes. Ces gènes sont impliqués les fonctions de « carboxylic catabolic process » et plus spécifiquement dans le catabolisme de l'arginine en glutamate ou en succinate. En ce qui concerne les gènes réprimés, seul M2 présente un grand nombre de gènes réprimés, en raison de l'arrêt de croissance. Les gènes codant pour un transporteur de glycine ABC (*proV*, *proW* et *proX*) sont réprimés dans les conditions M2, M3, M4 mais pas dans X1. Il serait intéressant de doser les acides aminés dans le surnageant.

A partir des gènes différentiellement exprimés comptabilisés (Tableau 5) nous avons réalisé des enrichissements fonctionnels en Gene Ontology via Ecocyc afin d'étudier les variations d'expression des fonctions biologiques entre les différentes conditions. Nous avons regardé quels étaient les processus biologiques sur-représentés avec une p-valeur $< 10^{-3}$ (Figure 46).

GO term for biological function	G2/G1	X1/G1	M1/G1	M2/G1	M3/G1	M4/G1	M2/G2	M4/X1	Some representative genes
GO:0006412 translation	2.36E-26			4.73E-35					rplK, rplW, rpmB, pheS, infA, rpmA, ...
GO:0008652 cellular amino acid biosynthetic process	1.31E-08			1.81E-08					gltB, ilvE, cysI, cysM, thrA, serA, ...
GO:0046394 carboxylic acid biosynthetic process	7.33E-10			1.52E-09					folE, cysM, gltB, thrA, serA, ilvE, ...
GO:0006754 ATP biosynthetic process	1.07E-07			1.13E-06					atpB, atpG, atpA, atpD, atpF, atpE, ...
GO:0009165 nucleotide biosynthetic process	7.42E-07			5.39E-05					upp, aceF, pyrF, purK, guaB, atpA, ...
GO:0043414 macromolecule methylation	1.33E-04			3.03E-04					trmD, mnmG, dcm, prmB, rsmC, hsdM, ...
GO:0019290 siderophore biosynthetic process	7.90E-05			1.13E-06	7.20E-09		1.33E-06		entA, entC, entD, entE, ybdZ, entB
GO:0006879 cellular iron ion homeostasis	1.38E-03			3.36E-04					exbB, fhuB, fepG, fecA, efeB, fecR, ...
GO:0009110 vitamin biosynthetic process	2.89E-04								bioB, serC, bioC, thiL, folB, ribE, ...
GO:0042732 D-xylose metabolic process	2.35E-04	1.39E-11		1.40E-06	9.46E-08	4.46E-09			xylB, xylA, xylF, xylH, xylG
GO:0046395 carboxylic acid catabolic process	2.62E-08			3.06E-11	6.83E-08	4.36E-07	1.31E-03	3.15E-06	astA, prpB, gadA, rhmD, acs, astB, ...
GO:0008643 carbohydrate transport	1.01E-04	6.85E-10		3.86E-12	1.64E-05	1.15E-07	4.76E-14	6.91E-03	ytfQ, araG, xylF, ytfT, sgcB, nanT, ...
GO:0045852 pH elevation	5.27E-03			2.80E-03	5.87E-04	9.82E-05	1.60E-07	1.99E-05	gadC, gadA, gadB
GO:0009063 cellular amino acid catabolic process	8.52E-06			5.67E-08	1.45E-07	4.35E-07		2.97E-07	gadA, astC, dadA, lhgD, gadB, astA, ...
GO:0015794 glycerol-3-phosphate transmembrane transport	4.20E-06			1.40E-06					glpT, ugpC, ugpA, ugpE, ugpB
GO:0009062 fatty acid catabolic process	1.69E-04			1.33E-03	4.32E-04		2.23E-03	4.27E-03	prpB, fadA, prpD, paaJ, prpE, fadD, ...
GO:0009310 amine catabolic process	3.35E-03			3.95E-03				5.96E-04	patA, puuE, puuD, purK, puuA, eutD, ...
GO:0042594 response to starvation	5.16E-05			1.87E-03	7.84E-04				yeaG, gabP, cstA, ddpX, dps, mgtS, ...
GO:0009450 gamma-aminobutyric acid catabolic process	2.35E-04			9.89E-05					gabT, gabD, gabP, puuE, ...
GO:0015768 maltose transport						1.02E-06	5.22E-09		malK, malF, lamB, malG

Figure 46 : Processus biologiques sur-représentés dans les listes de gènes différentiellement surexprimés pour chaque condition comparée à la condition G1, pour M2 comparée à G2 ou pour M4 comparée à X1. L'enrichissement en GO a été réalisé sur les gènes différentiellement exprimés entre les conditions. Les données ont été obtenues par analyse RNAseq avec une p-value $< 10^{-5}$ et un seuil à 4 pour les gènes régulés positivement et 0,25 pour le gène régulé négativement. Les « GO term », la p-value associé et une liste de quelques gènes illustratifs sont représentés. Les p-value des enrichissements des gènes positivement régulés sont représentés avec un gradient vert foncé à vert clair et celles des gènes négativement avec un gradient violet à rose.

Nous avons réalisé des enrichissements fonctionnels des processus biologiques sur les 1161 gènes différentiellement exprimés entre G2 et G1. Comme attendu, les gènes impliqués dans la croissance sont réprimés comme le montre l'enrichissement significatif des groupes fonctionnels : « translation », « cellular amino acid biosynthetic process », « ATP biosynthetic process », « nucleotide biosynthetic process » et « vitamin biosynthetic process » (Figure 46). De la même manière les gènes impliqués dans les fonctions « siderophore biosynthetic process », « cellular iron ion homeostasis » et « macromolecule methylation » sont réprimés. Ainsi dans la condition G2, les cellules sont en phase stationnaire impliquant la diminution de la synthèse de protéine, la diminution en besoin énergétique, ainsi que la diminution globale de l'anabolisme (Figure 46). À l'inverse, beaucoup de fonctions cataboliques sont induites par rapport à la phase exponentielle (Figure 46) comme le montrent les enrichissements des groupes fonctionnels « carboxylic acid catabolic process », « cellular amino acid catabolic process » ou « fatty acid catabolic process » cette réponse pourrait être induite pour palier à la carence en carbone et ainsi utiliser des réserves internes.

La principale différence entre ces conditions X1 et G1 réside dans l'expression de la voie métabolique d'utilisation du D-xylose qui est induite pendant la croissance sur xylose (Figure 46). Lors de croissance sur xylose, le groupe fonctionnel « carbohydrate transport » est enrichi, certainement en raison de la levée de la répression catabolique, menant à l'expression des transporteurs de substrats alternatifs au glucose (par exemples, *mglC*, *mglB*, *araG*, *araH*, *xylG*, *xylF* et *xylH* impliqués dans les transports du galactose, de l'arabinose et du xylose).

Dans la comparaison, M2 vs G2, seuls 31 gènes sont différentiellement exprimés. L'analyse des enrichissements fonctionnels des gènes différentiellement exprimés entre les conditions G2 et M2 illustrent une surexpression des gènes impliqués dans le transport des carbohydrates dans la condition M2, ainsi qu'une moindre expression des fonctions « pH elevation » et « siderophore biosynthesis process » dans la condition M2 (Figure 46). G2 et M2 sont deux conditions dans lesquelles la croissance est arrêtée, cependant, dans la condition G2, l'arrêt de la croissance est dû à une carence en carbone, tandis que dans la condition M2, l'arrêt de la croissance n'est que transitoire. La réponse entre les deux

conditions est similaire cependant la présence de xylose semble entraîner une surexpression des transporteurs de carbohydrates (Figure 46).

Nous avons étudié les gènes qui étaient surexprimés dans les conditions de croissance en mélange glucose/xylose (M2, M3 et M4). Afin de faciliter l'analyse, toutes les conditions ont été comparées à la phase exponentielle de croissance sur glucose G1. Pendant la transition du glucose au xylose, le métabolisme du xylose et la réponse à la carence en substrat sont d'abord induits avant la reprise de croissance sur xylose (condition M2 1 à 2h après l'épuisement du glucose). Ainsi la réponse est similaire à celle observée lors de la phase stationnaire (G2). Les fonctions impliquées dans la biosynthèse d'ATP, des nucléotides, d'acides carboxyliques, de sidérophores et de vitamines sont arrêtées et les fonctions de catabolisme des acides aminés, des acides carboxyliques et des acides gras sont surexprimées. Nous pouvons observer une augmentation de l'expression des transporteurs de substrats alternatifs au glucose. À la fin de la transition, en comparaison avec la phase exponentielle de croissance (condition M3/G1), les fonctions de biosynthèses ne sont plus réprimées comme en M2 sauf pour la biosynthèse des sidérophores. De plus, les fonctions cataboliques de catabolisme des acides aminés, des acides carboxyliques et des acides gras sont toujours induites par rapport à G1 (sauf le catabolisme des amines et les acides gamma-aminobutyriques). Par la suite, pendant la phase de croissance sur xylose comparée à la croissance exponentielle sur glucose (M4/G1), le métabolisme du D-xylose et le transport des glucides sont toujours induits alors que la réponse à la carence et le catabolisme des acides gras ne sont plus induits.

Cependant en M4, le profil d'expression des gènes n'est pas identique à celui observé pour des cellules en croissance sur xylose comme unique source de carbone. De manière intéressante, les fonctions « pH elevation », « cellular amino acid catabolic process », « carboxylic acid catabolic process », « carbohydrate transport », « amine catabolic process » et « fatty acid catabolic process » sont surexprimées en M4 par rapport à X1. Ces fonctions sont aussi surexprimées quand on compare M2 à G1, ainsi cette surexpression présente en M4 par rapport à X1 pourrait provenir du passage par la phase M2.

II. 2. C Analyse de gènes ciblés

À partir de ce que nous avons observé à travers les analyses de Gene Ontology et de diagramme de Venn, nous avons étudié le comportement de certains gènes sélectionnés pour présenter des profils d'expression qui ont suscité notre intérêt. Les gènes des opérons *xyl* sont surexprimés directement après l'épuisement du glucose (Figure 47) pendant la croissance sur un mélange glucose/xylose, et une expression plus élevée est également détectable en phase stationnaire (G2), sauf pour le gène *xylR*. Les gènes de l'opéron arabinose possèdent le même profil d'expression que les gènes des opérons xylose. Il est à noter que la souche BW25113 chez qui cette étude a été réalisée ne possède pas les gènes *araBAD*. Ces gènes sont plus fortement exprimés sur M2 que sur G2. Sans présence de xylose, l'expression de ces gènes augmente sans doute à cause de la levée de la répression catabolique, la présence de xylose dans le milieu permettant une expression bien plus forte (Figure 47).

La famille des gènes *his* codant pour la voie de biosynthèse de l'histidine est réprimée pendant la phase stationnaire et la transition (G2 et M2). La répression diminue lentement lorsque l'on observe une croissance sur xylose (M3, M4), montrant une réponse lente à l'échelle de la population. Ces réponses d'expression sont exprimées de manière logique en réponse à l'utilisation progressive du xylose. Cette famille de gènes est un exemple de nombreux gènes impliqués dans l'anabolisme présentant ce type de profils d'expression. Cependant tous les gènes ne présentent pas ce profil de réponse, certains gènes sont exprimés différenciellement entre M4 et X1. Ainsi, l'opéron *gad* est spécifiquement surexprimé pendant la transition du glucose au xylose mais pas pendant la croissance sur xylose comme seule source de carbone (Figure 47). Une réponse similaire est observée pour les gènes de la famille *prp* impliqués dans le catabolisme du propionate, ainsi que pour les gènes *hde*, impliqués dans la réponse au stress acide. De manière remarquable, ces familles de gènes (*gad*, *prp* et *hde*) semblent réprimées sur M2 par rapport à G2. D'autres familles de gènes semblent spécifiquement surexprimées pendant la transition mais pas pendant la croissance sur xylose comme seule source de carbone, il s'agit de la famille *ast* (catabolisme de l'arginine), la famille *lsr* et la famille de gènes *puu* (voie d'utilisation de la putrescine) (Figure 47). Ces gènes pourraient être différemment exprimés entre M4 et X1 à cause de la croissance précédente sur glucose, soit parce qu'ils font partie de la signature des cellules non adaptées au xylose, soit que

les cellules en reprise de croissance sur xylose après transition ont gardé des traces de leur histoire métabolique.

Gene code	Gene name	G2/G1	Pvalue	X1/G1	Pvalue	M1/G1	Pvalue	M2/G1	Pvalue	M3/G1	Pvalue	M4/G1	Pvalue	M2/G2	Pvalue	M4/X1	Pvalue
b3565	xyIA	12.44	3.8E-75	437.70	2.6E-100	2.15	3.9E-01	53.68	8.0E-14	270.61	1.7E-42	416.72	6.7E-159	4.47	2.1E-02	0.97	9.8E-01
b3564	xyIB	8.04	2.6E-36	159.35	7.8E-81	1.13	9.8E-01	20.37	3.3E-18	90.84	3.9E-27	135.38	4.3E-102	2.63	2.5E-02	0.86	8.2E-01
b4031	xyIE	3.18	4.3E-05	538.73	7.6E-34	1.58	9.8E-01	43.31	2.8E-21	143.97	7.3E-66	320.71	3.3E-44	14.17	3.4E-11	0.59	5.8E-01
b3566	xyIF	16.70	1.7E-27	816.12	6.4E-13	2.00	9.8E-01	383.46	7.1E-08	813.36	7.0E-74	1259.78	3.8E-71	23.07	4.0E-03	1.62	7.7E-01
b3567	xyIG	2.82	4.9E-04	190.32	1.1E-22	1.05	1.0E+00	70.73	2.7E-07	194.01	2.2E-42	267.09	2.7E-64	25.84	2.2E-04	1.45	7.0E-01
b3568	xyIH	6.05	1.4E-14	157.87	7.3E-79	2.14	9.8E-01	62.23	1.5E-11	139.86	1.4E-49	197.28	7.2E-75	10.51	5.6E-05	1.26	6.6E-01
b3569	xyIR	0.58	6.5E-03	16.15	2.2E-38	0.81	9.8E-01	3.70	1.1E-03	14.17	2.8E-23	20.97	4.9E-65	6.62	1.1E-04	1.32	5.2E-01
b0064	araC	7.75	3.7E-27	6.62	4.9E-13	1.02	1.0E+00	14.90	8.7E-12	5.64	2.1E-20	4.83	4.9E-08	2.00	3.5E-01	0.74	7.3E-01
b2841	araE	1.24	2.1E-01	217.14	3.2E-51	1.39	9.8E-01	386.17	2.6E-10	500.95	3.4E-34	176.53	6.7E-12	321.31	1.8E-08	0.82	9.2E-01
b1901	araF	68.50	2.5E-98	635.63	2.2E-31	1.10	9.9E-01	4733.95	5.0E-17	4831.63	1.9E-58	2119.20	3.5E-53	71.83	1.9E-06	3.41	7.5E-02
b1900	araG	28.45	2.3E-45	214.21	1.8E-41	1.48	9.8E-01	1443.29	5.8E-16	1686.80	4.6E-73	846.28	4.5E-48	52.43	4.2E-06	4.00	8.9E-03
b4460	araH	17.15	1.1E-38	57.01	1.4E-38	1.00	1.0E+00	564.97	4.1E-17	623.19	1.9E-58	270.05	2.2E-40	34.32	9.5E-07	4.78	2.8E-04
b0396	araJ	1.31	1.1E-01	17.73	9.1E-46	1.31	9.8E-01	5.08	2.3E-04	16.89	1.4E-33	11.56	1.7E-13	4.04	2.2E-02	0.66	5.6E-01
b1747	astA	82.96	5.7E-120	0.86	9.6E-01	0.97	1.0E+00	73.46	5.8E-42	26.53	5.4E-12	9.12	8.8E-13	0.92	9.6E-01	10.66	4.5E-09
b1745	astB	91.36	1.6E-158	0.95	1.0E+00	1.19	9.8E-01	65.42	2.4E-45	26.38	9.8E-17	11.46	2.6E-18	0.75	6.7E-01	12.12	6.0E-12
b1748	astC	120.03	6.3E-103	0.89	9.9E-01	1.14	9.8E-01	117.83	8.3E-40	36.74	1.8E-11	10.84	2.5E-12	1.03	1.0E+00	12.19	3.4E-09
b1746	astD	82.56	1.4E-111	0.89	9.9E-01	0.98	1.0E+00	69.05	8.3E-35	24.24	2.5E-15	9.73	3.5E-12	0.88	9.2E-01	11.02	1.0E-08
b1744	astE	94.37	2.5E-188	0.80	9.4E-01	1.18	9.8E-01	57.86	9.0E-48	25.88	3.6E-18	10.27	5.1E-17	0.64	3.5E-01	12.88	3.7E-10
b3517	gadA	60.81	1.3E-35	0.95	1.0E+00	0.75	9.8E-01	6.45	8.3E-08	25.41	1.4E-17	22.27	4.6E-13	0.11	2.7E-06	23.55	1.2E-09
b1493	gadB	44.16	2.2E-43	0.70	7.8E-01	0.73	9.8E-01	5.03	5.5E-05	17.14	8.6E-18	14.26	2.6E-11	0.12	1.3E-05	20.61	7.3E-12
b1492	gadC	67.25	1.3E-76	0.95	1.0E+00	0.71	9.8E-01	7.40	4.0E-07	21.37	1.9E-34	15.68	4.1E-16	0.11	1.1E-05	16.57	2.3E-11
b3512	gadE	63.24	6.8E-39	0.92	1.0E+00	0.85	9.9E-01	10.29	1.6E-03	23.25	1.9E-22	16.51	4.0E-14	0.16	6.3E-02	18.73	4.5E-09
b3510	hdeA	45.96	3.4E-55	0.65	9.4E-01	0.77	9.9E-01	6.62	5.9E-02	17.12	5.4E-22	13.34	8.7E-12	0.15	2.7E-01	21.10	1.9E-04
b3509	hdeB	60.61	1.3E-60	0.67	9.4E-01	0.76	9.9E-01	9.73	6.5E-03	23.73	1.3E-26	16.83	2.2E-14	0.17	1.7E-01	26.04	5.9E-06
b3511	hdeD	19.97	2.3E-50	0.97	1.0E+00	0.71	9.8E-01	3.47	3.8E-07	7.28	2.6E-15	5.99	1.6E-07	0.18	4.8E-08	6.24	5.5E-05
b1513	lsrA	534.00	1.1E-125	7.63	9.0E-05	1.00	1.0E+00	925.75	5.8E-52	86.00	7.3E-29	92.91	4.0E-40	1.80	3.5E-01	12.51	6.8E-09
b1516	lsrB	162.34	9.2E-159	2.46	5.0E-01	0.50	9.8E-01	336.15	5.6E-11	40.08	4.1E-20	37.29	1.1E-28	2.13	6.5E-01	15.56	3.7E-06
b1514	lsrC	258.48	5.9E-137	5.24	1.6E-03	1.54	9.8E-01	407.92	1.1E-50	46.00	1.6E-13	47.61	2.5E-31	1.63	3.7E-01	9.05	6.9E-08
b1515	lsrD	1.89	4.6E-06	1.49	7.8E-01	0.78	9.8E-01	3.02	4.8E-05	1.15	6.8E-01	1.46	1.1E-01	1.66	3.2E-01	1.00	1.0E+00
b1517	lsrF	125.09	5.1E-163	2.88	1.3E-01	0.75	9.8E-01	268.37	1.2E-08	33.18	1.9E-17	29.26	1.4E-37	2.22	6.7E-01	10.37	6.9E-08
b1518	lsrG	105.97	3.3E-96	2.52	4.2E-01	0.97	1.0E+00	207.89	2.2E-08	31.07	2.0E-14	22.39	4.5E-24	2.00	7.0E-01	9.06	1.9E-07
b1511	lsrK	68.65	7.7E-114	1.82	5.4E-01	0.89	9.9E-01	124.40	3.6E-15	14.11	6.0E-13	13.64	4.8E-24	1.87	5.7E-01	7.66	9.4E-08
b1512	lsrR	47.06	2.0E-92	2.72	1.9E-02	0.89	9.9E-01	98.48	8.3E-20	12.49	8.3E-14	13.60	2.2E-28	2.17	3.3E-01	5.12	9.3E-06
b0331	prpB	4433.17	0.0E+00	11.37	5.5E-09	0.88	9.9E-01	78.37	9.9E-11	642.13	1.7E-37	603.22	3.4E-61	0.02	4.6E-09	54.22	3.9E-23
b0333	prpC	1806.54	6.0E-231	6.31	4.5E-04	0.81	9.9E-01	35.83	9.2E-06	404.03	1.4E-36	358.25	6.0E-46	0.02	1.6E-06	55.92	9.7E-22
b0334	prpD	325.21	3.3E-127	2.09	1.9E-01	0.57	9.8E-01	8.78	1.3E-03	95.81	3.9E-31	95.32	1.1E-38	0.03	2.1E-06	45.87	3.3E-28
b0335	prpE	208.20	9.5E-84	3.43	5.1E-04	1.22	9.8E-01	9.43	6.3E-04	87.14	5.3E-27	85.59	4.2E-41	0.05	1.7E-04	25.20	1.3E-19
b0330	prpR	9.37	6.7E-24	1.06	1.0E+00	1.37	9.8E-01	71.37	3.0E-15	5.00	5.3E-06	2.93	8.7E-08	7.94	3.4E-04	2.76	6.4E-03
b1297	puuA	4.64	2.9E-21	0.68	3.4E-01	0.66	9.8E-01	4.16	1.2E-04	3.24	5.8E-02	1.50	1.3E-01	0.93	9.8E-01	2.25	2.2E-03
b1301	puuB	7.90	1.6E-21	0.27	2.0E-07	0.57	9.8E-01	5.69	3.5E-04	6.04	2.7E-02	1.74	5.7E-02	0.75	8.7E-01	6.56	3.5E-09
b1300	puuC	8.60	7.5E-21	0.29	1.7E-05	0.60	9.8E-01	6.45	1.2E-04	6.30	2.6E-02	1.68	1.0E-01	0.78	9.0E-01	5.91	3.0E-08
b1298	puuD	9.86	1.3E-23	0.46	6.8E-02	0.71	9.8E-01	6.44	7.3E-05	4.86	2.0E-02	1.87	5.1E-02	0.68	7.2E-01	4.09	4.7E-08
b1302	puuE	9.00	9.6E-30	0.33	2.0E-04	0.58	9.8E-01	7.22	2.3E-05	6.75	1.4E-02	2.19	4.9E-04	0.83	9.3E-01	6.78	1.6E-09
b1296	puuP	7.69	1.8E-36	0.70	5.6E-01	0.78	9.8E-01	4.07	2.2E-05	3.82	1.7E-02	1.81	1.6E-02	0.55	3.5E-01	2.65	9.8E-04
b1299	puuR	8.13	1.2E-31	0.46	5.7E-03	0.65	9.8E-01	3.92	6.7E-03	3.47	5.7E-02	1.50	1.8E-01	0.50	4.7E-01	3.29	3.9E-06
b2024	hisA	0.14	2.6E-51	0.79	7.6E-01	0.96	1.0E+00	0.17	1.3E-07	0.38	2.6E-02	0.61	3.7E-02	1.29	8.3E-01	0.78	6.8E-01
b2022	hisB	0.12	2.3E-49	0.72	4.6E-01	0.97	1.0E+00	0.16	7.5E-10	0.36	5.5E-03	0.62	1.3E-02	1.35	7.3E-01	0.88	8.4E-01
b2021	hisC	0.11	3.1E-54	0.82	8.7E-01	1.02	1.0E+00	0.18	5.2E-06	0.33	7.0E-04	0.63	1.8E-02	1.79	4.5E-01	0.78	6.5E-01
b2020	hisD	0.13	3.4E-43	1.00	1.0E+00	0.95	9.9E-01	0.42	1.3E-01	0.30	2.0E-04	0.64	1.8E-02	3.30	2.1E-01	0.65	4.8E-01
b2025	hisF	0.12	2.0E-42	0.82	8.7E-01	0.99	1.0E+00	0.17	5.4E-08	0.39	1.0E-02	0.67	6.9E-02	1.42	7.1E-01	0.83	7.8E-01
b2019	hisG	0.06	5.9E-45	1.02	1.0E+00	0.67	9.8E-01	0.26	3.1E-02	0.19	8.7E-08	0.46	4.2E-05	4.76	1.3E-01	0.45	1.7E-01
b2023	hisH	0.16	6.7E-35	0.78	6.7E-01	1.09	9.9E-01	0.18	1.1E-07	0.39	2.1E-02	0.66	3.3E-02	1.18	9.2E-01	0.85	7.9E-01
b2026	hisI	0.16	2.1E-43	0.76	6.2E-01	0.96	9.9E-01	0.26	1.1E-03	0.38	2.5E-02	0.60	8.1E-03	1.66	6.1E-01	0.79	6.7E-01

Figure 47 : Niveau d'expression en RNAseq et p-valeur pour une sélection de gènes d'intérêt. Les données sont présentées après normalisation par le R package TMM_exact où chaque condition est comparée à la condition G1.

II. 3. Déconvolution des données : Etude du profil d'expression génique au niveau des sous-populations.

Après avoir examiné les données d'expression des gènes à l'échelle de la population entière, nous avons entrepris de déconvoluer les résultats au niveau des sous-population. En effet, trois profils types de fluorescence différents ont été précédemment décrits: le « type glucose » exprimant uniquement la fluorescence bleue, « le type xylose » avec un haut niveau de fluorescence rouge et « le type transition » avec un niveau de fluorescence modéré.

Les analyses cytométriques nous fournissent le pourcentage de chaque population pendant la transition. Une hypothèse est que les différentes sous-populations n'ont pas les mêmes niveaux d'expression génique. Lors de la croissance sur xylose ou glucose, les profils de fluorescence sont homogènes, il est donc raisonnable de penser que l'expression génique est homogène, puisque sans sous-population. Nous avons précédemment émis l'hypothèse que la transition était due à l'émergence de certaines cellules capables de transiter et non à une adaptation homogène de l'ensemble de la population.

Dans notre cas, différents schémas sont envisageables. Le premier est que les cellules de type « G » ont la même expression que les cellules présentes lors de la croissance exponentielle sur le glucose (G1), que les cellules de type « X » sont des cellules qui poussent sur xylose (donc présentant le même profil que X1) et que les cellules de type « T » sont des cellules qui ne poussent pas et présentent un profil d'expression des gènes similaires aux cellules en phase stationnaire (comme pour G2).

Pour tester ce scénario, nous avons calculé le niveau théorique M2, M3 et M4 d'expression génique sur la base des données obtenues en G1, G2 et X1, et le pourcentage de chaque sous-population. Pour cela nous avons utilisé la formule suivante :

$$\% \text{cellules type G} * \text{expression G1} + \% \text{cellules type T} * \text{expression G2} + \% \text{cellules type X} * \text{expression X1}$$

Nous avons ensuite effectué des corrélations de Pearson entre les profils d'expression théoriques et les valeurs expérimentales obtenues pour M2, M3 et M4 (Figure 48). La première hypothèse ou la population de type G est resté avec un profil G1, la population

de type T avec un profil G2 et la population de type X avec un profil de type X1 s'est avérée peu concluante, notamment vis-à-vis de la condition M2.

	"popG" / "popT" / "popX"	M2	M3	M4
Hypothesis 1	G1 G2 X1	0.19	0.90	0.85
Hypothesis 2	G2 G2 X1	0.93	0.88	0.83
Hypothesis 3	G2 X1 X1	0.94	0.91	0.89
Hypothesis 4	X1 X1 X1	0.37	0.50	0.66
Hypothesis 5	G2 G2 G2	0.93	0.86	0.69

Figure 48 : Valeur de corrélation de Pearson pour l'hypothèse de comportement des différentes sous-populations. Le niveau d'expression théorique a été calculé avec une extrapolation du niveau d'expression multipliée par le pourcentage des sous-populations. La similitude entre les profils théoriques et mesurés a été comparée par corrélation de Pearson (un score de 0 indique une absence totale de similarité, un score de 1 une totale identité)

La population de type G étant majoritaire en M2, il semble très improbable que ces cellules soient restées avec un profil transcriptomique G1. Nous avons donc testé d'autres scénarios. La deuxième hypothèse choisie est que les cellules des populations G et T ont un profil d'expression G2 (Figure 48). Ce scénario présente une bien meilleure corrélation que le scénario précédant pour M2. Cependant, cette corrélation est plus basse pour M3 et M4 où les cellules présentant un niveau de fluorescence rouge intermédiaire (« type T ») sont nombreuses.

La troisième hypothèse est que les cellules de type T ont un profil d'expression X1 tout comme les cellules de type X, les cellules de type G étant de profil G2. Cette hypothèse 3 obtient les meilleurs scores statistiques tout au long de la transition, bien au dessus des autres scénarios testés (toutes les cellules avec un profil X1 ou G2) (Figure 49). Il semblerait donc que le profil d'expression du type T soit plus proche de X1 que de G2. Ce résultat corrobore ceux des tests de reprise de croissance après tri cellulaire. Lors de la transition, deux sous-populations avec des niveaux de fluorescence rouge différents apparaissent mais toutes deux semblent présenter des propriétés phénotypiques similaires en termes de capacité de reprise de croissance sur glucose ou xylose, et

possiblement aussi en termes de profil transcriptionnel. Nous en concluons que la souche portant le plasmide pBR322 et nos marqueurs ne présente pas ou peu de phénomène de bistabilité, la grande majorité des cellules réorientant leur métabolisme vers la consommation du xylose. Néanmoins, nous observons que les profils cytométriques et, par extrapolation, les vitesses d'adaptations entre cellules, sont très hétérogènes, certaines cellules acquérant rapidement le profil « xylose », d'autres nécessitant plusieurs heures.

III. Conclusion

Le modèle développé avait estimé que la souche pSB1C3 à haut nombre de copies présentait un taux d'apparition de consommateurs de xylose à partir de la population d'origine après épuisement du glucose de 2,02 cellules sur 10 000 par heure, alors que cette valeur était de 8 760 cellules sur 10 000 par heure pour la souche sauvage. Ici, nous avons utilisé ce modèle pour prédire la fréquence de transition cellulaire du glucose au xylose pour la souche portant le plasmide pBR322. Ainsi, nous avons prédit un taux d'apparition de 102 cellules sur 10 000 par heure. Ce résultat semble en accord avec les observations cytométriques et l'évolution du profil transcriptionnel au cours de la diauxie.

La construction sur plasmide à moyen nombre de copies ne semble pas engendrer d'état bistable. Cependant, la comparaison des profils d'expression des cellules en condition M4 comparées aux cellules en condition X1 a montré des différences alors que ces deux conditions sont des croissances sur xylose. Il semblerait qu'une partie des gènes surexprimés lors de l'arrêt de croissance au moment de la transition (M2) le soit toujours après la reprise de croissance sur xylose (M4). De plus, lors de croissance sur xylose après la transition glucose/xylose les cellules présentent des niveaux de fluorescence rouge très étalés. Or, ces cellules semblent se comporter de manière similaire sur la reprise de croissance après tri. Ainsi, il est possible que les cellules s'adaptent progressivement (entraînant différents niveaux de fluorescence rouge) avec un temps d'adaptation variable d'une cellule à l'autre. Il est aussi possible que les profils d'expressions obtenus en M4 soient dus à la présence de certaines cellules restées bloquées en arrêt de croissance. Il reste quelques cellules de type G après la transition représentant environ 10

% des cellules en M4, ces cellules ne présentent pas de fluorescence rouge, il serait intéressant de tester leurs capacités de reprise de croissance sur les différents substrats, elles pourraient représenter des cellules qui n'ont pas encore transitées sur xylose. La caractérisation des phénotypes des cellules est complexe et mériterait d'être approfondie.

Discussion et Perspectives

I. Conclusions

La compréhension des mécanismes d'adaptation mis en place par les cellules bactériennes présente un intérêt majeur dans de nombreux domaines. Cet intérêt est aussi bien présent dans le domaine des biotechnologies avec la compréhension des facteurs pouvant impacter le bioprocédé (en terme de productivité ou de rendement) (Binder *et al.*, 2017), dans le domaine de la santé avec la compréhension des phénomènes de persistance (Brauner *et al.*, 2016), dans le domaine agroalimentaire pour la conservation des aliments, que pour la bioéconomie avec la valorisation de déchets carbonés complexes (Dien *et al.*, 2002; Cirino *et al.*, 2006). Cependant, l'adaptation bactérienne reste encore mal connue et de nombreux paramètres restent à élucider, notamment à l'échelle de la cellule unique. Nombre d'études récentes décrites dans l'introduction bibliographie montrent que certaines conditions de stress pourraient favoriser l'apparition de sous-populations présentant des comportements phénotypiques différents, à l'encontre du dogme précédent qui considérait une adaptation homogène de la population.

Dans ce travail, nous avons étudié une transition métabolique d'intérêt bioéconomique, la transition glucose/xylose chez la bactérie *E. coli*, en nous focalisant plus particulièrement sur le comportement individuel des cellules afin d'établir si cette adaptation s'accompagne ou non de phénomènes d'hétérogénéité. La première partie du travail a consisté à mettre au point une méthodologie nous permettant d'observer les phénotypes des cellules au sein de la population et de pouvoir observer ou non d'une variabilité phénotypique lors de croissances en mélange glucose/xylose. Cette méthodologie repose sur l'expression de protéines fluorescentes induite par la consommation d'un substrat ciblé. Pour cela nous avons construit un jeu de plasmide permettant l'expression de gènes codant pour des protéines fluorescentes de manière constitutive ou en réponse à l'utilisation du xylose. Le promoteur P_{xyIA} a été utilisé afin de marquer spécifiquement la consommation du xylose. L'utilisation du plasmide pSB1C3 à haut nombre de copies permettant l'expression de fluorescence rouge sous contrôle du promoteur P_{xyIA} et l'expression de fluorescence bleue sous contrôle de P_{ihfB} (constitutif) nous a permis de décortiquer les mécanismes de la transition du glucose au xylose. Le

temps de transition était regardé jusqu'à présent comme le temps nécessaire pour produire les enzymes impliquées dans le catabolisme du second substrat. Ici, nous avons pu mettre en évidence que le facteur clé influençant ce temps de transition du métabolisme du glucose vers celui du xylose est la disponibilité du facteur de transcription XylR. Ce rôle a été vérifié par la surexpression de *xylR* réduisant le temps de transition et la titration de XylR augmentant ce temps de transition. Ce résultat montre que le temps de transition n'est pas une simple contrainte technique pour permettre aux cellules de produire la machinerie nécessaire à la consommation du second substrat. De plus l'étude de la souche permettant la titration et présentant un temps de transition artificiellement rallongé nous a montré que certaines cellules s'adaptent rapidement à l'utilisation du xylose et sont à l'origine de l'émergence de la population en croissance sur xylose. Dans ce cas, une partie de la population ne transite pas sur le second substrat mais reste stable en nombre d'individus et semble arrêter sa croissance. La modélisation a permis d'appuyer sur le caractère stochastique de la capacité à s'adapter au second substrat. Le temps de transition pourrait donc être redéfini comme le temps nécessaire pour qu'une majorité de la population bascule sur le nouveau substrat. Dans cette conception, le paramètre central n'est plus le temps de production de la machinerie métabolique mais la probabilité de bascule vers le nouveau substrat.

Nous avons par la suite construit un nouveau plasmide avec un nombre de copies réduit à une vingtaine contre les deux cents copies théoriques du pSB1C3 afin de réduire l'impact sur la durée de la transition glucose/xylose en réduisant la titration de XylR induite par les sites de fixation de XylR présents dans le promoteur P_{xylA} . S'il est toujours possible d'observer des sous-populations basées sur les niveaux de fluorescence avec cette nouvelle construction, la différence entre les sous-populations est moins nette qu'avec le plasmide pSB1C3 portant la construction fluorescente. Les différentes sous-populations ont été triées pour les deux souches afin de les remettre en croissance. Pour la souche avec le plasmide à haut nombre de copies où la titration de XylR est très forte, les cellules présentant un signal fluorescent similaire aux cellules en croissance sur xylose ré-initient leurs croissances sur glucose ou sur xylose rapidement lors de la transition. Au contraire les cellules présentant un signal fluorescent similaire aux cellules en croissance sur glucose sont incapables de repartir sur xylose et seule la moitié est capable de repartir en croissance sur glucose. A l'inverse, les cellules issues des différentes sous-populations de la souche portant la construction sur le pBR322 portant la construction fluorescente

recommencent à croître rapidement sur xylose ou glucose quelle que soit la sous-population dont elles sont issues. Au niveau transcriptomique, nous avons pu observer qu'il n'y a pas de différence entre des cellules en phase exponentielle de croissance sur glucose en présence ou non de xylose. De plus, lors de croissance en mélange glucose/xylose, les cellules expriment, à la disparition du glucose, une réponse similaire mais pas identique que du xylose soit présent ou non. Lors de la reprise de croissance sur xylose, le profil d'expression des gènes tend petit à petit vers celui de cellules en croissance sur xylose comme unique source de carbone, mais des différences sont observables et potentiellement due à une hétérogénéité au sein de la population. Ces différences pourraient être dues à variabilité dans le temps d'adaptation des cellules, ou à la présence de cellules en arrêt de croissance.

II. Critique de la Méthodologie

II.1. Limites de notre méthodologie

La méthodologie choisie pour cette étude présente des limites. La présence sur le plasmide de sites de fixation d'un facteur de transcription clé pour la croissance sur xylose impacte la physiologie des cellules. Ainsi, le fait d'augmenter les cibles de fixation de XylR par l'insertion plasmidique de nombreuses copies du promoteur P_{xylA} induit une titration. Ce type de titration a été modélisé, notamment pour l'opéron lactose et le régulateur LacI, et semblerait suivre une corrélation dose/réponse influençant l'expression des gènes (Lee and Maheshri, 2012; Rydenfelt *et al.*, 2014; Brewster *et al.*, 2014). Ainsi la présence de nombreuses cibles non chromosomiques de XylR empêche l'activation de l'opéron. Cette titration est moins importante chez la souche portant le plasmide à moyen nombre de copies (pBR322 P_{ihfB} BFP P_{xylA} MRF1), car le nombre de cibles non chromosomiques est réduit. Chez cette souche, la transition du glucose au xylose présente une durée intermédiaire de 3 h, entre celle de la souche sauvage de 0.6 h et celle de la souche à haut nombre de copies plasmidiques (pSB1C3) de 8h.

Cependant, même réduite, la présence de cette titration impacte le métabolisme des cellules menant à des comportements lors de la transition non-présents ou difficilement détectables chez la souche sauvage. Ainsi la présence de la titration de XylR induit un

phénotype où l'acétate est consommé à la disparition du glucose avant la consommation du xylose. Normalement, chez les souches d'*Escherichia coli* sauvages, la hiérarchie d'utilisations des substrats dépend du taux de croissance capable d'être assuré par le substrat (Aidelberg *et al.*, 2014), ainsi chez la souche sauvage le xylose est consommé avant l'acétate.

L'autre point à prendre en compte est la surcharge métabolique possiblement induite par notre approche. Notre méthodologie induit deux sources de surcharge métabolique potentielles : la première induite par la présence de plasmide au sein de la cellule, et la seconde par la surproduction de deux protéines recombinantes (Bentley *et al.*, 1990; Silva *et al.*, 2012). Le cout énergétique dû à la maintenance d'un plasmide au sein de la cellule peut induire une diminution du taux de croissance ainsi que du rendement en biomasse associé au substrat de culture. Ces paramètres ont donc été observés (Ow *et al.*, 2006; Heins *et al.*, 2019). De plus, la diminution du taux de croissance peut entraîner une variation dans l'expression de certains gènes, notamment de gènes impliqués dans le transport de nutriment ou dans le métabolisme énergétique de la cellule (Ow *et al.*, 2006).

Dans nos conditions la présence du plasmide semble induire une surcharge métabolique plus forte chez la souche portant le plasmide à haut nombre de copies avec la construction que chez la souche portant le plasmide à moyen nombre de copies avec la construction. Cependant cette surcharge n'est pas visible sur xylose pour la souche portant le plasmide pBR322 avec construction. La surproduction de protéines recombinantes ainsi que la présence de plasmide peuvent entraîner une augmentation de l'expression de certaines réponses aux stress, notamment une augmentation de la transcription des gènes impliqués dans la réponse au choc thermique (Gill *et al.*, 2000; Ow *et al.*, 2006). Notre étude transcriptomique étant entièrement basée sur la souche disposant de la construction sur pBR322, elle ne permet pas de vérifier la réalité de ce biais mais le limite car il serait présent dans toutes les conditions.

Il est possible que l'expression de fluorescence soit en partie toxique pour les cellules. Cette toxicité est faible et n'a pas été étudiée pour les protéines fluorescentes utilisées dans cette étude, cependant nous avons écarté le marqueur sYFP2 car son impact sur le taux de croissance était trop élevé. Il a été rapporté que l'expression de GFP chez des

cellules eucaryotes entraînerait un comportement apoptotique chez 3% de la population (Liu *et al.*, 1999). Cependant cet effet a été peu observé chez les bactéries. Les protéines fluorescentes tetramériques pourraient entraîner une toxicité chez les bactéries, mais les protéines monomériques semblent non toxiques (Shaner *et al.*, 2005). Ainsi la toxicité des protéines fluorescentes ne semble pas être un problème de notre étude. Un autre facteur à prendre en compte est le temps de maturation des protéines fluorescentes, la maturation de la mRFP1 est de 1h maximum (Campbell *et al.*, 2002) alors que celle de la mTagBFP est de 12 min (Subach *et al.*, 2011).

Pour certaines études des marqueurs non génétiques ont été utilisés. Kotte *et al.*, (2014) avaient utilisé un marqueur membranaire dont l'intensité diminuait avec la division des cellules permettant d'observer des cellules se divisant de celle ne se divisant pas. Cependant ces méthodologies semblent invasives pour la cellule et les protocoles associés semblent induire des biais important (Heinemann *et al.*, 2020).

II. 2. Perspectives d'amélioration

Pour approfondir l'étude de l'hétérogénéité de phénotype lors de la transition glucose/xylose, il est nécessaire de limiter l'impact de la méthodologie d'observation. Pour cela nous devons nous affranchir au maximum du nombre de copies de P_{xyIA} alternatifs, afin de réduire le nombre de sites de fixation de XylR.

L'intégration génomique de la construction permettrait de n'introduire qu'une seule copie alternative du promoteur P_{xyIA} . Notre premier test d'intégration n'était pas concluant car le niveau d'expression des protéines fluorescentes était trop faible et non détectable en cytométrie. Les perspectives sont d'améliorer ce signal. Il existe différents leviers d'action pour cela : utiliser une protéine fluorescente avec une intensité de fluorescence plus forte, intégrer la construction dans un nouveau site génomique qui serait potentiellement plus actif, ou utiliser un promoteur permettant une expression plus forte. Cette dernière option semble la plus difficilement réalisable puisque différents promoteurs spécifiques de l'utilisation du xylose ont déjà été testé et P_{xyIA} permettait l'expression la plus forte. Effectuer un changement des protéines fluorescentes utilisées dans l'étude semble être le levier le plus prometteur. Il serait possible de choisir une protéine fluorescente avec une intensité de fluorescence plus forte. Il faudrait aussi

privilégier un temps de maturation plus court que celui de la mRFP1 qui a été utilisé dans ce travail (estimé à un maximum de 1h d'après Campbell *et al.*, 2002). Ce choix de protéine fluorescente est un compromis entre le temps de maturation, l'intensité de fluorescence et la stabilité au photoblanchiment. Ce dernier paramètre peut être important en fonction de la méthodologie d'analyse de la fluorescence utilisé. Ainsi, une utilisation en microscopie nécessitera une meilleure photostabilité de la protéine.

Une autre possibilité prometteuse serait de s'affranchir d'ajout de séquence P_{xylA} en intégrant le gène codant pour une protéine fluorescente directement au locus génomique de l'opéron xylose en aval ou en amont du gène *xylA*, afin de créer une protéine chimère XylA-Protéine fluorescente. Dans ce cas, il est important de vérifier la fonctionnalité de la protéine XylA, mais aussi l'expression des gènes de l'opéron.

III. Le rôle de XylR et des facteurs impactant la transition

III.1. Le rôle de XylR

Durant cette thèse nous avons pu montrer que la disponibilité du facteur de transcription XylR est un des facteurs principaux impactant le temps de transition du glucose au xylose. L'importance de la disponibilité en XylR pour la transition étant établie, nous avons regardé le profil d'expression du gène *xylR* pendant la transition glucose/xylose chez la souche pBR322. Ainsi, il est possible de remarquer que le gène *xylR* est légèrement surexprimé pendant la croissance exponentielle sur xylose, mais cependant beaucoup moins que les autres gènes constituant l'opéron xylose. De plus, *xylR* n'est pas surexprimé en phase stationnaire de croissance après la disparition du glucose contrairement aux autres gènes de l'opéron. Cette relative stabilité d'expression peut s'expliquer par la disposition des gènes dans l'opéron même si on ne peut pas exclure l'implication de phénomènes de régulation au niveau de la stabilité du transcrit *XylR*. Le gène *xylR* se situe en effet à la fin de l'opéron *xylFGH* et il est connu que les gènes situés en position terminale d'un opéron sont de manière générale moins exprimés que les gènes en position initiale (Lim *et al.*, 2011). *XylR* semble de plus posséder son propre promoteur

(Laikova *et al.*, 2001), identifié *in silico* mais non testé expérimentalement. Ainsi le niveau d'expression de *xylR* pourrait dépendre de ce promoteur. Nous avons pu remarquer que son niveau d'expression lors de croissance sur glucose est plus élevé que celui des gènes de l'opérons *xylFGH*. Ainsi la quantité de XylR dans la cellule semble être maintenue constante afin de promouvoir l'activation de l'opéron xylose à la disparition du glucose et en présence de xylose. Enfin, nous avons démontré que la surexpression de *xylR* n'a pas impacté le taux de croissance de la souche. Il est possible de se demander alors pourquoi le niveau d'expression de *xylR* n'est pas plus important. Potentiellement le promoteur *xylR* permettrait une expression basale plus forte et moins soumise aux variations environnementales. Chez des souches de *E. coli* KO11 et LY01, *xylR* semble plus fortement exprimé que *xylFG* lors de croissance sur glucose (Gonzalez *et al.*, 2002). Il est ainsi possible que l'expression de *xylR* sur glucose soit maintenue par la présence de son promoteur propre et que l'augmentation de l'expression provienne, au moins en partie du promoteur de l'opéron *xyFGH*.

La dynamique des systèmes facteurs de transcription/sites cibles est importante dans l'expression génique. Ceci est illustré notamment par le fait que le gène codant pour un facteur de transcription et ses sites de fixation sont généralement co-localisés sur le génome chez les procaryote (Kolesov *et al.*, 2007). L'activation des promoteurs grâce à des facteurs de transcription a longtemps été regardé comme un système ON/OFF : en ON s'il y a fixation du facteur de transcription sur sa cible, et OFF si non. Cependant, ces systèmes semblent plus complexes puisqu'il semble aussi y avoir une notion de temporalité. Ainsi le temps de contact entre le facteur de transcription et sa cible semble être de 5 min en moyenne (Hammar *et al.*, 2014). Il s'agit d'un facteur supplémentaire influençant la disponibilité en XylR au sein de la cellule. Nous avons décrit précédemment que la disponibilité en XylR peut jouer comme un seuil de déclenchement de l'utilisation ou non du xylose. Cela est supporté par l'expression bistable de l'opéron xylose observé par (Afroz *et al.*, 2014). Leur étude a aussi reporté une expression bistable de l'opéron nécessaire à l'utilisation de L-rhamnose, et de l'opéron arabinose lorsque le L-arabinose est présent en faible quantité. Il serait intéressant de regarder le rôle des régulateur AraC et RhamS lors de transition glucose/arabinose et glucose/rhamnose. Les opérons étant aussi décrits comme bistables il serait aussi intéressant de regarder si les cellules se comportent de manière homogène durant la transition. L'organisation génomique,

impliquant la présence de facteurs de transcription et de sites cibles permettrait de promouvoir l'hétérogénéité dans le temps d'adaptation des cellules pour repartir en croissance sur un autre substrat. Cette organisation étant générique pour de nombreux opérons permettant l'utilisation de sucres la variabilité du temps d'adaptation pourrait être commune à l'utilisation des différents substrats.

III. 2. Les autres facteurs impactants la transition

Notre étude allie deux types de reprises de croissance différents : des reprises de croissance lors de la diauxie (donc lors de croissance en mélange de substrat), ainsi que des reprises de croissance après changement artificiel du milieu de culture (soit un temps de latence après inoculation). Dans notre étude, ces deux temps apparaissent différents mais corrélés. Une explication potentielle pourrait provenir du fait que la croissance sur glucose permettrait d'avoir des cellules suffisamment chargées en métabolite et composées énergétiques pour faciliter la bascule et rendre la transition plus rapide. Une accélération de la transition a été observée en ajoutant 0.002% de glucose lors de croissance sur gluconate pour des cellules de *E. coli* (Jöers and Tenson, 2016). Cependant, nous avons aussi observé que la latence de reprise de croissance et le temps de diauxie sont différents selon le mode de culture utilisé, qu'il soit en bioréacteur, en microplaque ou en Erlenmeyer. Ainsi, il semblerait que d'autres facteurs influençant ne soient pas pris en compte. Chez la levure *Saccharomyces cerevisiae*, le temps de reprise de croissance après changement artificiel dépend du temps passé sur le substrat. Ainsi, plus longtemps les cellules seront en croissance sur maltose entre deux cultures sur glucose, plus le temps de latence pour redémarrer sur glucose sera long (Cerulus *et al.*, 2018). L'antécédant des cellules semble donc impacter le temps de transition.

D'autres facteurs décrits dans des études précédentes pourraient aussi avoir un impact sur l'adaptation lors de la transition glucose/xylose, tel que l'âge des cellules, leur taille, la disponibilité en ribosome ou encore le niveau de la répression catabolique. Le rôle du complexe CRP-AMPc dans cette transition reste encore à investiguer. Nous avons réalisé des tests préliminaires de mutant en provenance de la collection KEIO afin d'étudier leurs phénotypes lors de transitions artificielles du glucose au xylose. La souche Δcya est incapable de redémarrer sa croissance sur xylose lors de transfert artificiel de milieu (Figure 49). Le complexe CRP-AMPc pourrait être nécessaire pour permettre un

niveau basal d'expression de *xylR* afin de rendre la transition possible. Cependant la souche Δcrp semble capable de redémarrer sa croissance sur xylose, malgré un temps de latence moyen de plus de 10h. Ces tests sont à poursuivre avec des délétions propres des gènes, c'est-à-dire en supprimant la cassette kanamycine restante pouvant influencer l'expression des gènes afin d'assurer la fiabilité des résultats.

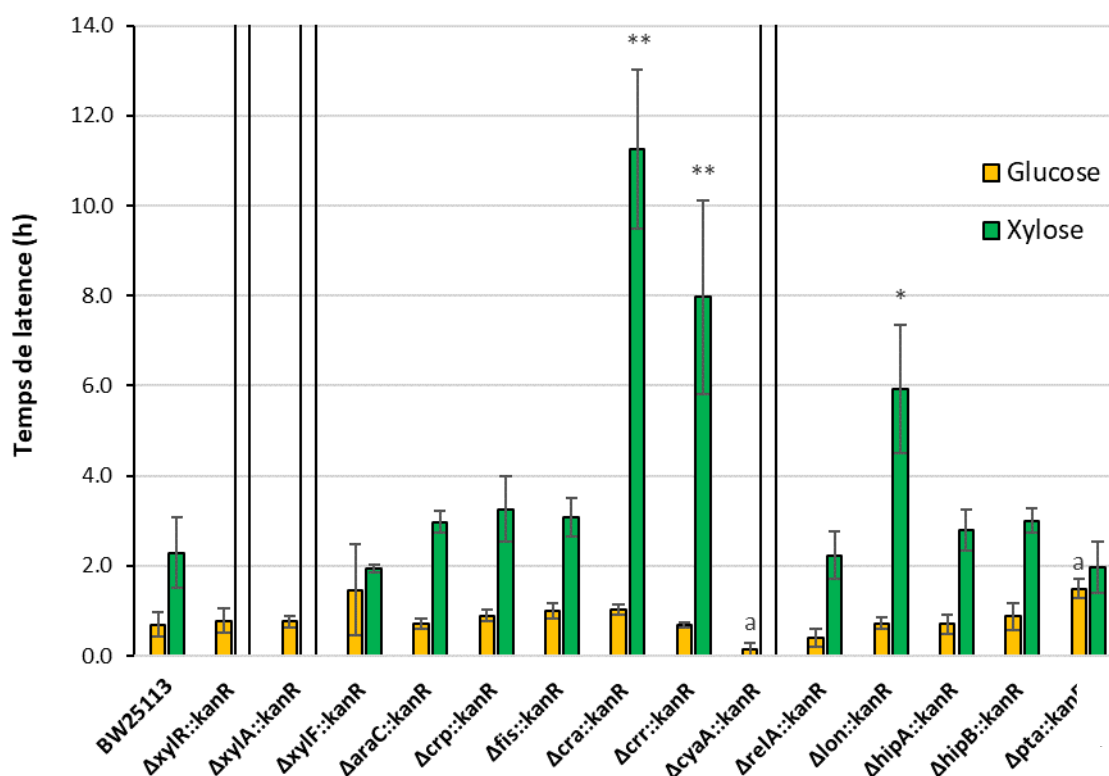


Figure 49 : Temps de latence de souche mutante d'*E. Coli* BW25113 avant reprise de la croissance sur xylose après une préculture effectuée sur glucose. Les souches cultivées sont la BW25113 et 14 souches mutantes de la collection Keio (Baba *et al.*, 2006) ($n = 9$). Les barres blanches représentent les souches qui n'ont pas repris leur croissance au bout de 24h. La significativité statistique des latences des souches par rapport à la souche de référence BW25113 est représentée par (a) pour la reprise de croissance sur glucose et (*) pour celle sur xylose en considérant $P < 0,05$ (a ou*) ou $P < 0,01$ (**) basé sur un test de student bilatéral.

Une autre piste d'étude serait de tester d'autres transitions et de chercher à démontrer si les facteurs de transcription sont aussi importants que lors de la transition glucose/xylose. Les transitions glucose/arabinose ou glucose/lactose sont des modèles particulièrement intéressants car la structuration de la régulation de l'arabinose est similaire à celle du xylose et que la régulation du lactose fonctionne différemment. De plus des études ont montré un comportement hétérogène lors de la transition glucose/lactose (Boulineau *et al.*, 2013; Solopova *et al.*, 2014). L'étude de la transition glucose/acétate

semble aussi intéressante car l'acétate est produit lors de croissance sur les autres substrats avant d'être reconsommé. C'est un modèle étudié par l'équipe et nous avons commencé la sélection de promoteurs d'intérêt.

IV. Comportement de la souche BW25113 avec ou sans construction

IV. 1. La souche portant le plasmide pSB1C3

Nous avons montré que les comportements des deux souches portant les plasmides à différents nombres de copies ne sont pas similaires. En ce qui concerne la souche à haut nombre de copies, les tests de reprise de croissance sur glucose et sur xylose confirment le fait que seules quelques cellules s'adaptent au xylose et sont à l'origine de l'émergence de la population en croissance sur xylose. Ainsi une partie importante n'est pas capable de s'adapter, et semble métaboliquement inactive. Ce fait est dû au manque de disponibilité en XylR. Ce plasmide entraîne certes un phénotype exacerbé sur la transition, mais cette transition rallongée nous a permis d'étudier l'importance du facteur de transcription XylR. Le système d'expression de l'opéron xylose repose sur l'activation d'une boucle de régulation positive. L'activation de cette boucle semble dépendre d'une concentration seuil en XylR disponible au sein de la cellule, induisant une bistabilité de l'expression de l'opéron déjà observé (Afroz *et al.*, 2014) entre les cellules ayant franchi cette concentration seuil et celle ne l'ayant pas franchi. Chez afroz cette bistabilité n'est observée qu'à très faible concentration en xylose, soit entre 10 et 100 μM , ainsi il semblerait que la limitation en complexe XylR-xylose induise cette bistabilité dans notre cas la limitation provient de xylR et n'est plus observée chez la souche où XylR n'est pas titré et dans le cas d'Afroz *et al.*, (2014) la limitation provient de la concentration en xylose, cette bistabilité n'est plus observée quand la concentration en xylose est plus importante.

La phase de transition prolongée observée dans la souche contenant le plasmide à haut nombre de copies est intéressante car, au vu des résultats de reprise de croissance, elle permettrait de générer des cellules non croissantes. Le devenir des cellules incapables de croître sur un nouveau substrat reste à déterminer. Afin de déterminer leurs

propriétés physiologiques, il faudrait approfondir les tests de leur activité métabolique. Dans le cas d'infection par des salmonelles, une sous-population est non répliquative mais métaboliquement active et capable de répondre à un signal (Helaine *et al.*, 2010). Les cellules non croissantes pourraient présenter des capacités similaires. Il serait aussi intéressant de tester leur résistance aux antibiotiques étant donné que le passage par une transition diauxique pourrait générer des cellules persistantes (Amato *et al.*, 2013).

Lors de la transition du glucose au xylose, l'étude transcriptomique montre une induction similaire à celle présente lors de carence en carbone. Une réponse similaire a été observée lors de croissance en mélange glucose/lactose associé à une activation de la réponse stringente et de la réponse au stress médié par RpoS dans les 10 min suivant disparition du glucose (Traxler *et al.*, 2006). Cette réponse pourrait induire un blocage de certaines cellules et empêcher la reprise de croissance sur un nouveau substrat (Solopova *et al.*, 2014; Radzikowski *et al.*, 2016). Il serait donc intéressant de tester les niveaux intracellulaires de ppGpp pour chaque sous-population.

IV. 2. La souche portant le plasmide pBR322

La souche d'*Escherichia coli* portant le plasmide pBR322 présente, lors de la transition, des sous-populations avec des niveaux différents de fluorescence rouge. Cependant l'étude de reprise de croissance après tri montre que toutes les cellules exprimant la fluorescence rouge ont le même comportement lors de reprise de croissance sur glucose ou xylose. Toutes les cellules semblent capables de transiter sur un nouveau substrat. Cependant la croissance en microplaque ne peut être regardée qu'à partir d'une certaine biomasse. Ainsi, les temps de latence sont extrapolés à partir du taux de croissance observé, entraînant une imprécision dans la mesure. Pour palier à ce problème, nous commençons à étudier les reprises de croissance en microgouttellettes en collaboration avec l'ESPCI de Paris afin de pouvoir observer la reprise de croissance dès le début. L'étude transcriptomique nous suggère que les deux sous-populations semblent avoir un profil d'expression des gènes similaires. Cependant, la comparaison des profils d'expression des gènes des cellules en croissances sur xylose pur et des cellules en croissance sur xylose après la transition diauxique montre des différences potentiellement dues à une hétérogénéité au sein de la population. Ces différences pourraient être dues à la variabilité dans le temps d'adaptation des cellules ou à la

présence de cellules en arrêt de croissance. Afin d'apporter des réponses à ces hypothèses il sera nécessaire d'étudier directement les différences entre les sous-populations. Maintenant que nous avons identifié environ 80 gènes présentant un comportement différent sur M4 et X1, il devient possible d'analyser leur niveau d'expression en RT-qPCR dans des cellules isolées. Ainsi nous pourrions déterminer si ces gènes sont différenciellement exprimés avec des niveaux variables dans les cellules, laissant penser à une adaptation hétérogène, ou si les différences sont dues à la présence de cellules non adaptées et donc à une bistabilité dans la population.

Les différents niveaux de fluorescence pourraient provenir de différents niveaux de transcrit mais pourraient aussi avoir d'autres origines. En effet, la méthodologie employée dans cette étude comporte des limites dues à l'utilisation de plasmides. Un premier point à prendre en compte est la division et le maintien du nombre de copies de plasmide. Le plasmide à haut nombre de copies ne possède pas de système de partition actif et sa répartition dans les cellules filles semble uniquement due à la diffusion aléatoire (Million-Weaver and Camps, 2014). Il pourrait y avoir une variabilité du nombre de copies de plasmide d'une cellule à l'autre. Chez la souche contenant le plasmide pBR322, le nombre de copies est moins important en raison de la présence du gène *rop* sur le plasmide. *Rop* permet de réduire le nombre de copies du plasmide en impactant indirectement la formation du complexe ADN-ARN nécessaire à l'initiation de la réplication du plasmide (Cesareni *et al.*, 1982; Sozhamannan and Stitt, 1997). Cependant ce plasmide ne possède pas de système de partition actif lui non plus. Ainsi, potentiellement, le nombre de copies de plasmide pourrait différer entre les cellules filles et pourrait entraîner une variation dans le niveau d'intensité observé. Cependant, pour cette souche seul le niveau de fluorescence rouge est variable et non le niveau de fluorescence bleue, de plus le nombre de copies du plasmide ne varie pas entre une croissance sur glucose et une croissance sur xylose, il est peut probable que les variations du niveau d'intensité de fluorescence rouge soient dues à une variation du nombre de copies de plasmide.

Il est aussi possible que des propriétés dues à la nature des protéines fluorescentes, notamment leur temps de maturation, aient un impact sur les profils fluorescents des cellules. En effet, la maturation de la mRFP1 est de 1h maximum (Campbell *et al.*, 2002)

alors que celle de la mTagBFP est de 12 min (Subach *et al.*, 2011). Afin de savoir si la nature des protéines fluorescentes est responsable des différents niveaux de fluorescence rouge, un autre test complémentaire nécessaire implique la construction des plasmides avec inversion des gènes codants pour les protéines fluorescentes afin de voir si les profils obtenus restent similaires. Le gène codant pour la mTagBFP sera donc placé sous contrôle du promoteur *xylA* et le gène codant pour la mRFP1 sera sous contrôle du promoteur *ihfB* (Figure 52). Ces constructions sont en cours de réalisation.

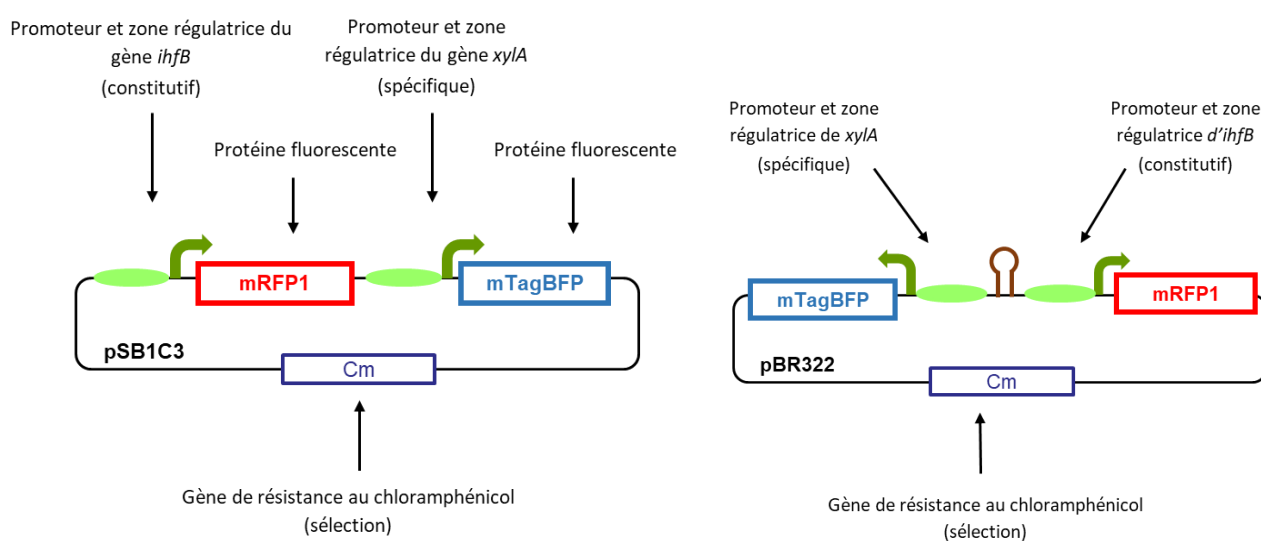


Figure 52 : Schéma des constructions plasmidiques avec le gène codant pour la mTagBFP sous contrôle du promoteur *xylA* et le gène codant pour la mRFP1 sous contrôle du promoteur *ihfB* dans le vecteur à haut nombre de copies pSB1C3 ou à moyen nombre de copies pBR322.

IV. 3. Chez la souche sauvage

Dans le cas de transition d'un substrat à un second substrat non préférentiel, les facteurs de transcription jouent un rôle important. La souche possédant le plasmide à haut nombre de copies, nous a permis d'étudier ce mécanisme. Nous supposons que le mécanisme impliqué dans la transition, est similaire chez la souche sauvage, bien que celle-ci est une durée de transition très courte (environ 0,6 h). La différence résiderait dans le nombre de cellules restant « inactives » par rapport au nombre de cellules

capables de reprendre la croissance sur le second substrat, paramètre appelé taux de transition ou « switch rate » dans le modèle développé. Ainsi chez la souche sauvage, contrairement à la souche avec XylR titré, le nombre de cellule capable d'effectuer la transition sur le second substrat est majoritaire. Nous l'avons estimé à 8 760 cellules sur 10 000 par heure contre 102 cellules sur 10000 par heure pour la souche contenant le plasmide pBR322 et 2 cellules sur 10000 par heure pour la souche contenant le plasmide pSB1C3. L'étude plus approfondie de ce taux de transition chez la souche sauvage nécessite l'utilisation d'une construction intégrée dans le chromosome afin de limiter le nombre de cible de fixation de XylR. Nous avons décrit précédemment que la disponibilité en XylR peut jouer comme un seuil de déclenchement de l'utilisation ou non du xylose. L'amélioration de la méthodologie permettrait de venir se rapprocher du phénotype de croissance de la souche sauvage.

V. Conclusion finale

La question restante est quelle forme prend l'hétérogénéité lors de transitions du glucose au xylose, est-ce qu'une réelle bistabilité est présente avec des cellules capables de repartir en croissance et d'autres incapables de transiter sur le second substrat ? Cette bistabilité a été observée par Kotte *et al.*, (2014) lors de croissance en mélange glucose/acétate et dans notre étude en utilisant le plasmide à haut nombre de copies pSB1C3 contenant la construction fluorescente lors de croissance en mélange glucose/xylose. Cependant dans ces deux cas la méthodologie d'étude a un impact, provenant dans le premier cas des importants lavages et dans le second cas de la construction plasmidique (Heinemann *et al.*, 2020 et chapitre 2).

Il est probable que chez la souche sauvage toutes les cellules soient capables de repartir en croissance mais que l'hétérogénéité provienne du temps de latence des cellules pour s'adapter au second substrat. Plusieurs études semblent arriver à ces conclusions. Boulineau *et al.*, (2013) ont observé chez des cellules d'*Escherichia coli* en croissance sur glucose/lactose que 15% des cellules ne subissent pas de temps de latence et que le reste de la population présente une réelle phase de latence, dont la durée est variable entre les cellules. Chez *Methylobacterium extroquens* sur mélange méthanol/succinate (McClelland *et al.*, 2020), certaines cellules présentent un temps de

latence de plusieurs heures tandis que d'autres ne recommencent pas à croître. De plus chez *E. coli* lors de transitions artificielles de glucose à acétate, la durée de latence est variable entre les cellules (Basan *et al.*, 2020). Il semblerait que nous observons un phénomène similaire avec la souche portant le plasmide pBR322 contenant la construction fluorescente.

Lors de transitions du glucose au xylose, il semblerait que toutes les cellules soient capables de repartir en croissance sur le second substrat mais que le temps de latence soit variable, l'exacerbation des phénotypes, induit par exemple par la présence du plasmide pSB1C3 empêche les cellules de repartir en croissance, mais ce phénomène semble artéfactuel. Malgré les limitations présentes dans notre étude, il s'agit d'un des premiers travaux concernant l'étude de l'hétérogénéité métabolique lors de croissance en mélange glucose/xylose en bioréacteur contrôlé. Ce travail a permis d'établir une base importante pour la suite. De nombreuses pistes sont encore à explorer en termes d'amélioration technique et d'intérêt scientifique. Ce travail apporte de la compréhension des mécanismes impliqués dans l'hétérogénéité métabolique, utile dans le cadre du développement de procédés microbiens fiables, économiques et respectueux de l'environnement s'inscrivant dans le secteur de la bioéconomie.

Matériels et méthodes

I. Construction des souches

I. 1. Clonage en ligation/digestion

Les constructions dans le vecteur pSB1C3 ont été réalisées par digestion ligation. 25 µg de vecteur pSB1C3 ont été digérés par les enzymes de restriction EcoRI et XbaI dans un volume de 120 µL. De manière similaire 1,8 µg de chaque insert ont été digérés par les enzymes de restriction EcoRI et SpeI. Le vecteur et les inserts digérés ont été purifiés grâce au kit gel extraction (Qiagen, Allemagne), puis leurs concentrations ont été quantifiées au Nanodrop ® (ND-1000- spectrophotometer). Ils ont été assemblés par ligation avec un rapport molaire de 1:3 (vecteur : insert), dans un volume final de 20 µL avec 10 µL de tampon quick Ligase Reaction (NEB, USA), 1 µL de Quick Ligase (NEB, USA), et un ajout d'eau miliQ pour atteindre les 20 µL. La réaction a été effectuée pendant 1h à température ambiante. Le produit de la ligation est ensuite transformé dans la souche d'*E. coli* NEB 5-alpha (NEB, USA) et les colonies sont sélectionnées sur LB agar contenant 40 µg/mL de chloramphénicol. L'ADN plasmidique de 5 clones a ensuite été extrait en utilisant le kit QIAprep Spin Miniprep Kit (Qiagen, chine) et vérifié par digestion enzymatique et par séquençage (Eurofins, France) afin de valider la construction.

I. 2. Clonage Gibson

Le plasmide pSB1C3PihfbBFPxylARFP en sens de transcription inverse et le plasmide pBR322 avec inversion du sens de transcription ont été réalisés par assemblage Gibson. Pour cela les fragments ont été amplifiés à partir de plasmide purifiés ou d'ADN génomique d'*Escherichia coli* BW25113 avec des amorces permettant une homologie de 20 pb entre chaque fragment. L'assemblage a été réalisé grâce au mix DNA HiFi assembly (New England Biolab), avec 100ng de vecteur et un rapport molaire de 1:2 (vecteur : insert) et incubé 1h à 50°C. Le produit de la réaction est ensuite transformé dans la souche d'*E. coli* NEB 5-alpha (NEB, USA) et les colonies sont sélectionnées sur LB agar contenant 40 µg/mL de chloramphénicol. L'ADN plasmidique de 5 clones a ensuite été extrait en utilisant le kit QIAprep Spin Miniprep Kit (Qiagen) et vérifié par digestion enzymatique et par séquençage (Eurofins, France) afin de valider la construction.

I. 3. Intégration génomique

L'intégration a été réalisée par lambda Red en suivant la méthode de Datsenko and Wanner (2000). Pour cela un plasmide a été construit par méthode gibbon (I.2) contenant la construction fluorescente entre les zones d'homologie *aslA* et *aslB* du génome d'*E. coli* MG1655. Un produit PCR a été réalisé à partir de ce plasmide et d'amorces (p666 et p669 donnés par Sébastien Nouaille) s'hybridant sur les zones d'homologies *aslA* et *aslB* permettant de générer un fragment de 4799 bp contenant la construction en sens inversé et une cassette chloramphénicol entouré de séquences FRT entre les deux zones d'homologies *aslA* et *aslB*. Ce fragment PCR a été inséré dans des cellules compétentes de BW25113 préalablement transformées avec le plasmide pKD46. Le plasmide pKD46 permet l'expression de Gam, Exo et Beta nécessaire à la recombinaison entre les fragments d'homologie sur le substrat PCR et le chromosome. La construction a été vérifiée par PCR et séquençage (Eurofins).

II. Milieux de cultures

Différents milieux de cultures ont été utilisés dans cette étude. Pour les clonages, le milieu LB a été utilisé (composition : 10g NaCl, 10g tryptone, 5g d'extrait de levure) avec 40 mg/L de chloramphénicol. Pour les cultures, le milieu minimum M9 a été utilisé (composition : Na₂HPO₄ 33,7 mM, KH₂PO₄ 22,0 mM, NaCl 8,55 mM, NH₄Cl 9,35 mM, MgSO₄ 1 mM, CaCl₂ 0,3 mM, EDTA 13,4 mM, FeCl₃-6H₂O 3,1 mM, ZnCl₂ 0,62 mM, CuCl₂-2H₂O 76 µM, CoCl₂-2H₂O 42 µM, H₃BO₃ 162 µM, MnCl₂-4H₂O 8,1 µM, Thiamine 1 µg/L) avec 20 mg/L chloramphénicol et une concentration en source de carbone de 90 mM d'équivalent carbone (soit 15 mM de glucose, 18 mM de xylose ou 45 mM d'acétate).

III. Culture en bioréacteurs

Les cultures en Batch ont été réalisées dans un bioréacteur Sartorius Biostat B plus dans 1 L de milieu M9 avec 90 mMeqC de mélanges de glucose et / ou de xylose. La température a été réglée à 37 ° C et le pH à 7. Des conditions d'aération non limitatives ont été obtenues avec un débit d'air réglé à 0,35 L.min⁻¹ et une adaptation de l'agitation pour maintenir pO₂ > 30%. La croissance a été évaluée par une mesure OD_{600nm} à des

intervalles de 30 minutes avec un spectrophotomètre LibraS4 (Biochrom, Royaume-Uni). Les concentrations de glucose, de xylose et d'acétate ont été quantifiées par HPLC (Agilent, USA) avec une colonne d'exclusion d'ions HPX 87H (Biorad) comme décrit dans (Morin et al., 2017).

IV. Calcul des $\mu_{\max}$ et des temps de latence

Les taux de croissance maximum $\mu_{\max}$ ont été déterminés pendant la phase de croissance exponentielle et estimés à l'aide du coefficient de la régression linéaire de la courbe représentant $\ln(\text{OD})$ en fonction du temps. Le décalage diauxique a été considéré comme le temps perdu par rapport à la croissance hypothétique dans laquelle la première phase exponentielle est immédiatement suivie d'une seconde phase exponentielle. Les formules et les démonstrations de la façon dont le décalage a été calculé sont fournies dans le fichier de données supplémentaires S7 du chapitre 2. Les temps de latence ont été considérés comme la différence de temps nécessaire pour atteindre le même point de DO dans la phase exponentielle entre la croissance avec un décalage de temps et la croissance qui a commencé immédiatement au taux de croissance maximal, selon la définition d'Enjalbert (Enjalbert et coll., 2015).

V. Analyse de la cytométrie de flux

Des échantillons de cellule ont été prélevés à des intervalles de 30 minutes et immédiatement congelées à -20°C , car les tests ont montré que la congélation n'a aucune influence sur le niveau de fluorescence. Après décongélation et dilution dans du NaCl filtré à 0,9 % pour obtenir une concentration cellulaire de 1×10^6 cellules.mL⁻¹, les intensités de fluorescence ont été mesurées avec un cytomètre Masquant VYB (Miltenyi Biotec, Allemagne) équipé d'un laser jaune 561 nm pour l'excitation de mRFP1 et d'un laser violet 405nm pour mTagBFP. L'acquisition des données a été fixée à 40 000 événements par échantillon. Le logiciel FlowJo X a été utilisé pour l'analyse des données de cytométrie. Une gate a été créée dans le tracé des points du canal de diffusion avant (FSC-H) par rapport au canal de diffusion latéral (SSC-H) pour distinguer les bactéries du bruit technique. Une deuxième gate a été créée dans le graphique en points du SSC-H par rapport

au SSC-A pour sélectionner des cellules uniques. Les émissions de fluorescence mTagBFP ont été analysées avec le canal V1-H (BP450/25nm) et les émissions mRFP1 avec le canal Y2-H (BP615/10 nm). L'OD_{600nm} de chaque sous-population a été extrapolé à partir de la population totale de OD_{600nm} et du pourcentage de chaque sous-population obtenu en cytométrie de flux, basé sur la fluorescence bleue et rouge.

VI. Tri cellulaire pour l'expérience de reprise de croissance

Les cellules précultivées sur M9 glucose ont été transférées sur M9 40 % de glucose 60 % de xylose. La première culture a été inoculée à 0,01 OD_{600nm} et plusieurs dilutions à 1/10 ont été effectuées afin d'avoir au moins une culture en phase de transition (passage du glucose au xylose) le lendemain. Le tri cellulaire a été effectué sur les niveaux de fluorescence, lorsque la culture a atteint entre 1,5 et 2 OD_{600nm}. Une cellule a été triée par puits dans une plaque de 96 puits, remplie de 200 µL de glucose M9 ou de xylose M9. Au moins 27 cellules de chaque sous-population ont été triées dans chaque milieu. Les plaques ont été incubées pendant 78 h à 37 °C avec agitation et filmées pour éviter la contamination et l'évaporation. L'OD_{600nm} a été mesuré toutes les 17 min.

VII. Analyse transcriptionnelle

VII. 1. Prélèvement et préparation des échantillons

Pour l'analyse transcriptomique, nous avons réalisé des prélèvements pendant les cultures en bioréacteurs contrôlés. Pour chaque échantillon nous avons prélevé un équivalent de 4 unités de DO. Les échantillons ont été centrifugés 1 min à vitesse maximale (13 500 tr/min). Les culots ont été immédiatement congelés dans de l'azote liquide et stockés à une température de -80 °C. Les cellules ont été décongelées dans de la glace et mises en suspension dans du tampon Tris-EDTA (TE), puis lysées mécaniquement au Fastprep (MP Biomedical) avec des billes de verre. L'extraction d'ARN a été effectuée par ajout de Trizol et de phénol. La contamination potentielle de l'ADN a été éliminée grâce au kit TURBO DNA-free™ (Ambion). Des ERCC (ERCC RNA Spikes-In Control Mix 1, AMBION) a été ajouté avant ribodepletion avec 1 µL d'ERCC dilué à 1/10 pour 5000 ng

d'ARN total. La ribodepletion a été réalisée avec le kit Pan-Karyote SiTools ribo POOL (siTOOLS Biotech).

VII. 2. RNAseq

La qualité et la quantité des ARN ont été testés par Nanodrop et BioAnalyzer (Agilent). L'analyse RNAseq réalisée en IonTorrent (Ion S5, Thermofischer) à la plateforme GET-biopuces (TBI, Toulouse). La normalisation des données a été effectuée par le package R TMM_exact. Les gènes avec une expression différentielle ont été identifiés avec une p-valeur $<10^{-5}$ et un seuil de variation de niveau d'expression de 4 pour les gènes régulés positivement et 0,25 pour les gènes régulés négativement. Les enrichissement fonctionnels ont été réalisés sur les processus biologiques avec une p-valeur $<10^{-3}$ grâce à la base de données Ecocyc (Karp *et al.*, 2018).

Références

- Ackermann, M. (2015) A functional perspective on phenotypic heterogeneity in microorganisms. *Nat Rev Microbiol* **13**: 497–508.
- Afroz, T., Biliouris, K., Kaznessis, Y., and Beisel, C.L. (2014) Bacterial sugar utilization gives rise to distinct single-cell behaviours. *Mol Microbiol* **93**: 1093–1103.
- Agirrezabala, X., Fernández, I.S., Kelley, A.C., Cartón, D.G., Ramakrishnan, V., and Valle, M. (2013) The ribosome triggers the stringent response by RelA via a highly distorted tRNA. *EMBO Rep* **14**: 811–816.
- Aidelberg, G., Towbin, B.D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014) Hierarchy of non-glucose sugars in Escherichia coli. *BMC Syst Biol* **8**: 133.
- Amato, S.M., and Brynildsen, M.P. (2015) Persister Heterogeneity Arising from a Single Metabolic Stress. *Curr Biol* **25**: 2090–2098.
- Amato, S.M., Orman, M.A., and Brynildsen, M.P. (2013) Metabolic Control of Persister Formation in Escherichia coli. *Mol Cell* **50**: 475–487.
- Ammar, E.M., Wang, X., and Rao, C.V. (2018) Regulation of metabolism in Escherichia coli during growth on mixtures of the non-glucose sugars: arabinose, lactose, and xylose. *Sci Rep* **8**: 1–11.
- Arnoldini, M., Vizcarra, I.A., Peña-Miller, R., Stocker, N., Diard, M., Vogel, V., *et al.* (2014) Bistable Expression of Virulence Genes in Salmonella Leads to the Formation of an Antibiotic-Tolerant Subpopulation. *PLOS Biol* **12**: e1001928.
- Atkinson, G.C., and Hauryliuk, V. (2012) Evolution and Function of the RelA/SpoT Homologue (RSH) Proteins. In *eLS*. American Cancer Society,.
- Atkinson, G.C., Tenson, T., and Hauryliuk, V. (2011) The RelA/SpoT Homolog (RSH) Superfamily: Distribution and Functional Evolution of ppGpp Synthetases and Hydrolases across the Tree of Life. *PLoS ONE* **6**
- Axelrod, K., Sanchez, A., and Gore, J. (2015) Phenotypic states become increasingly sensitive to perturbations near a bifurcation in a synthetic gene network. *eLife* **4**: e07935.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., *et al.* (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006.0008.
- Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004) Bacterial Persistence as a Phenotypic Switch. *Science* **305**: 1622–1625.
- Basan, M., Honda, T., Christodoulou, D., Hörl, M., Chang, Y.-F., Leoncini, E., *et al.* (2020) A universal trade-off between growth and lag in fluctuating environments. *Nature* **584**: 470–474.

- Battesti, A., and Bouveret, E. (2006) Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. *Mol Microbiol* **62**: 1048–1063.
- Bentley, W.E., Mirjalili, N., Andersen, D.C., Davis, R.H., and Kompala, D.S. (1990) Plasmid-encoded protein: The principal factor in the “metabolic burden” associated with recombinant bacteria. *Biotechnol Bioeng* **35**: 668–681.
- Bertrand, R.L. (2019) Lag Phase Is a Dynamic, Organized, Adaptive, and Evolvable Period That Prepares Bacteria for Cell Division. *J Bacteriol* **201**
- Bettenworth, V., Steinfeld, B., Duin, H., Petersen, K., Streit, W.R., Bischofs, I., and Becker, A. (2019) Phenotypic Heterogeneity in Bacterial Quorum Sensing Systems. *J Mol Biol* **431**: 4530–4546.
- Bhogale, P.M., Sorg, R.A., Veening, J.-W., and Berg, J. (2014) What makes the lac-pathway switch: identifying the fluctuations that trigger phenotype switching in gene regulatory systems. *Nucleic Acids Res* **42**: 11321–11328.
- Bigger, Joseph W. (1944) Treatment of staphylococcal infections with penicillin by intermittent sterilisation. *The Lancet* **244**: 497–500.
- Binder, D., Drepper, T., Jaeger, K.-E., Delvigne, F., Wiechert, W., Kohlheyer, D., and Grünberger, A. (2017) Homogenizing bacterial cell factories: Analysis and engineering of phenotypic heterogeneity. *Metab Eng* **42**: 145–156.
- Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., *et al.* (1997) The Complete Genome Sequence of Escherichia coli K-12. *Science* **277**: 1453–1462.
- Boulineau, S., Tostevin, F., Kiviet, D.J., Wolde, P.R. ten, Nghe, P., and Tans, S.J. (2013) Single-Cell Dynamics Reveals Sustained Growth during Diauxic Shifts. *PLOS ONE* **8**: e61686.
- Brauner, A., Fridman, O., Gefen, O., and Balaban, N.Q. (2016) Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat Rev Microbiol* **14**: 320–330.
- Brewster, R.C., Weinert, F.M., Garcia, H.G., Song, D., Rydenfelt, M., and Phillips, R. (2014) The Transcription Factor Titration Effect Dictates Level of Gene Expression. *Cell* **156**: 1312–1323.
- Briggs, K.A., Lancashire, W.E., and Hartley, B.S. (1984) Molecular cloning, DNA structure and expression of the Escherichia coli D-xylose isomerase. *EMBO J* **3**: 611–616.
- Brown, D.R. (2019) Nitrogen Starvation Induces Persister Cell Formation in Escherichia coli. *J Bacteriol* **201**.
- Brown, L., Gentry, D., Elliott, T., and Cashel, M. (2002) DksA affects ppGpp induction of RpoS at a translational level. *J Bacteriol* **184**: 4455–4465.

- Browning, D.F., and Busby, S.J.W. (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**: 57–65.
- Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K. (1969) Factor stimulating transcription by RNA polymerase. *Nature* **221**: 43–46.
- Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (2002) A monomeric red fluorescent protein. *Proc Natl Acad Sci* **99**: 7877–7882.
- Casadesús, J., and Low, D.A. (2013) Programmed Heterogeneity: Epigenetic Mechanisms in Bacteria. *J Biol Chem* **288**: 13929–13935.
- Cashel, M., and Gallant, J. (1969) Two Compounds implicated in the Function of the RC Gene of Escherichia coli. *Nature* **221**: 838–841.
- Castañó-Cerezo, S., Bernal, V., Blanco-Catalá, J., Iborra, J.L., and Cánovas, M. (2011) cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in Escherichia coli. *Mol Microbiol* **82**: 1110–1128.
- Castañó-Cerezo, S., Pastor, J.M., Renilla, S., Bernal, V., Iborra, J.L., and Cánovas, M. (2009) An insight into the role of phosphotransacetylase (pta) and the acetate/acetyl-CoA node in Escherichia coli. *Microb Cell Factories* **8**: 54.
- Cerulus, B., Jariani, A., Perez-Samper, G., Vermeersch, L., Pietsch, J.M., Crane, M.M., *et al.* (2018) Transition between fermentation and respiration determines history-dependent behavior in fluctuating carbon sources. *eLife* **7**: e39234.
- Cesareni, G., Muesing, M.A., and Polisky, B. (1982) Control of ColE1 DNA replication: the rop gene product negatively affects transcription from the replication primer promoter. *Proc Natl Acad Sci* **79**: 6313–6317.
- Chang, D.-E., Smalley, D.J., Tucker, D.L., Leatham, M.P., Norris, W.E., Stevenson, S.J., *et al.* (2004) Carbon nutrition of Escherichia coli in the mouse intestine. *Proc Natl Acad Sci* **101**: 7427–7432.
- Chaudhuri, R.R., and Henderson, I.R. (2012) The evolution of the Escherichia coli phylogeny. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* **12**: 214–226.
- Chavarría, M., and Lorenzo, V. de (2018) The imbroglio of the physiological Cra effector clarified at last. *Mol Microbiol* **109**: 273–277.
- Choi, P.J., Cai, L., Frieda, K., and Xie, X.S. (2008) A Stochastic Single-Molecule Event Triggers Phenotype Switching of a Bacterial Cell. *Science* **322**: 442–446.
- Choudhury, D., Gayen, K., and Saini, S. (2018) Dynamic control of arabinose and xylose utilization in E. coli. *Can J Chem Eng* **96**: 1881–1887.
- Chowdhury, N., Kwan, B.W., and Wood, T.K. (2016) Persistence Increases in the Absence of the Alarmone Guanosine Tetraphosphate by Reducing Cell Growth. *Sci Rep* **6**: 1–9.

- Chu, D., and Barnes, D.J. (2016) The lag-phase during diauxic growth is a trade-off between fast adaptation and high growth rate. *Sci Rep* **6**: 25191.
- Cirino, P.C., Chin, J.W., and Ingram, L.O. (2006) Engineering *Escherichia coli* for xylitol production from glucose-xylose mixtures. *Biotechnol Bioeng* **95**: 1167–1176.
- Cole, J.A., Kohler, L., Hedhli, J., and Luthey-Schulten, Z. (2015) Spatially-resolved metabolic cooperativity within dense bacterial colonies. *BMC Syst Biol* **9**: 15.
- Crown, S.B., Long, C.P., and Antoniewicz, M.R. (2015) Integrated ¹³C-metabolic flux analysis of 14 parallel labeling experiments in *Escherichia coli*. *Metab Eng* **28**: 151–158.
- Curtis, S.J., and Epstein, W. (1975) Phosphorylation of D-glucose in *Escherichia coli* mutants defective in glucosephosphotransferase, mannosephosphotransferase, and glucokinase. *J Bacteriol* **122**: 1189–1199.
- Dal Co, A., Brannon, C., and Ackermann, M. (2018) Division of labor in bacteria. *eLife* **7**: e38578.
- Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640–6645.
- Death, A., and Ferenci, T. (1993) The importance of the binding-protein-dependent Mgl system to the transport of glucose in *Escherichia coli* growing on low sugar concentrations. *Res Microbiol* **144**: 529–537.
- Death, A., Notley, L., and Ferenci, T. (1993) Derepression of LamB protein facilitates outer membrane permeation of carbohydrates into *Escherichia coli* under conditions of nutrient stress. *J Bacteriol* **175**: 1475–1483.
- Delvigne, F., and Goffin, P. (2014) Microbial heterogeneity affects bioprocess robustness: dynamic single-cell analysis contributes to understanding of microbial populations. *Biotechnol J* **9**: 61–72.
- Delvigne, F., Zune, Q., Lara, A.R., Al-Soud, W., and Sørensen, S.J. (2014) Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. *Trends Biotechnol* **32**: 608–616.
- Desai, T.A., and Rao, C.V. (2010) Regulation of Arabinose and Xylose Metabolism in *Escherichia coli*. *Appl Environ Microbiol* **76**: 1524–1532.
- Deutscher, J. (2008) The mechanisms of carbon catabolite repression in bacteria. *Curr Opin Microbiol* **11**: 87–93.
- Deutscher, J., Francke, C., and Postma, P.W. (2006) How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. *Microbiol Mol Biol Rev* **70**: 939–1031.
- Dien, B.S., Nichols, N.N., and Bothast, R.J. (2002) Fermentation of sugar mixtures using *Escherichia coli* catabolite repression mutants engineered for production of L-lactic acid. *J Ind Microbiol Biotechnol* **29**: 221–227.

- Dörr, T., Vulić, M., and Lewis, K. (2010) Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. *PLoS Biol* **8**: e1000317.
- Dubnau, D., and Losick, R. (2006) Bistability in bacteria. *Mol Microbiol* **61**: 564–572.
- Durfee, T., Hansen, A.-M., Zhi, H., Blattner, F.R., and Jin, D.J. (2008) Transcription Profiling of the Stringent Response in *Escherichia coli*. *J Bacteriol* **190**: 1084–1096.
- Eisenberg, R.C., and Dobrogosz, W.J. (1967) Gluconate Metabolism in *Escherichia coli*. *J Bacteriol* **93**: 941–949.
- Eiteman, M.A., Lee, S.A., and Altman, E. (2008) A co-fermentation strategy to consume sugar mixtures effectively. *J Biol Eng* **2**: 3.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002) Stochastic Gene Expression in a Single Cell. *Science* **297**: 1183–1186.
- Enjalbert, B., Coccagn-Bousquet, M., Portais, J.-C., and Letisse, F. (2015) Acetate exposure determines the diauxic behavior of *Escherichia coli* during the glucose-acetate transition. *J Bacteriol* **JB.00128-15**.
- Enjalbert, B., Letisse, F., and Portais, J.-C. (2013) Physiological and Molecular Timing of the Glucose to Acetate Transition in *Escherichia coli*. *Metabolites* **3**: 820–837.
- Enjalbert, B., Millard, P., Dinclaux, M., Portais, J.-C., and Létisse, F. (2017) Acetate fluxes in *Escherichia coli* are determined by the thermodynamic control of the Pta-AckA pathway. *Sci Rep* **7**: 42135.
- Esquerré, T., Bouvier, M., Turlan, C., Carpousis, A.J., Girbal, L., and Coccagn-Bousquet, M. (2016) The Csr system regulates genome-wide mRNA stability and transcription and thus gene expression in *Escherichia coli*. *Sci Rep* **6**: 25057.
- Evans, C.R., Kempes, C.P., Price-Whelan, A., and Dietrich, L.E.P. (2020) Metabolic Heterogeneity and Cross-Feeding in Bacterial Multicellular Systems. *Trends Microbiol*
- Evans, D.J., and Evans, D.G. (1983) Classification of pathogenic *Escherichia coli* according to serotype and the production of virulence factors, with special reference to colonization-factor antigens. *Rev Infect Dis* **5 Suppl 4**: S692-701.
- Ferenci, T. (1996) Adaptation to life at micromolar nutrient levels: the regulation of *Escherichia coli* glucose transport by endoinduction and cAMP. *FEMS Microbiol Rev* **18**: 301–317.
- Fernández-Coll, L., and Cashel, M. (2018) Contributions of SpoT Hydrolase, SpoT Synthetase, and RelA Synthetase to Carbon Source Diauxic Growth Transitions in *Escherichia coli*. *Front Microbiol* **9**
- Ferullo, D.J., and Lovett, S.T. (2008) The Stringent Response and Cell Cycle Arrest in *Escherichia coli*. *PLoS Genet* **4**

- Finkel, S.E. (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat Rev Microbiol* **4**: 113–120.
- Fischer, D., Teich, A., Neubauer, P., and Hengge-Aronis, R. (1998) The General Stress Sigma Factor σ^S of *Escherichia coli* Is Induced during Diauxic Shift from Glucose to Lactose. *J BACTERIOL* **4**.
- Ganini, D., Leinisch, F., Kumar, A., Jiang, J., Tokar, E.J., Malone, C.C., *et al.* (2017) Fluorescent proteins such as eGFP lead to catalytic oxidative stress in cells. *Redox Biol* **12**: 462–468.
- Gefen, O., and Balaban, N.Q. (2009) The importance of being persistent: heterogeneity of bacterial populations under antibiotic stress. *FEMS Microbiol Rev* **33**: 704–717.
- Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**: 343–345.
- Gill, R.T., Valdes, J.J., and Bentley, W.E. (2000) A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli*. *Metab Eng* **2**: 178–189.
- Gimenez, R., Nuñez, M.F., Badia, J., Aguilar, J., and Baldoma, L. (2003) The gene *yjcG*, cotranscribed with the gene *acs*, encodes an acetate permease in *Escherichia coli*. *J Bacteriol* **185**: 6448–6455.
- Gonzalez, R., Tao, H., Shanmugam, K.T., York, S.W., and Ingram, L.O. (2002) Global Gene Expression Differences Associated with Changes in Glycolytic Flux and Growth Rate in *Escherichia coli* during the Fermentation of Glucose and Xylose. *Biotechnol Prog* **18**: 6–20.
- Görke, B., and Stülke, J. (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol* **6**: 613–624.
- Gosset, G. (2005) Improvement of *Escherichia coli* production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system. *Microb Cell Factories* **4**: 14.
- Gottesman, S. (2004) The Small RNA Regulators of *Escherichia coli*: Roles and Mechanisms. *Annu Rev Microbiol* **58**: 303–328.
- Gourse, R.L., Ross, W., and Rutherford, S.T. (2006) General Pathway for Turning on Promoters Transcribed by RNA Polymerases Containing Alternative σ Factors. *J Bacteriol* **188**: 4589–4591.
- Grimbergen, A.J., Siebring, J., Solopova, A., and Kuipers, O.P. (2015) Microbial bet-hedging: the power of being different. *Curr Opin Microbiol* **25**: 67–72.
- Guantes, R., Benedetti, I., Silva-Rocha, R., and Lorenzo, V. de (2016) Transcription factor levels enable metabolic diversification of single cells of environmental bacteria. *ISME J* **10**: 1122–1133.

- Hammar, P., Walldén, M., Fange, D., Persson, F., Baltekin, Ö., Ullman, G., *et al.* (2014) Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation. *Nat Genet* **46**: 405–408.
- Harms, A., Maisonneuve, E., and Gerdes, K. (2016) Mechanisms of bacterial persistence during stress and antibiotic exposure. *Science* **354**.
- Haseltine, W.A., and Block, R. (1973) Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. *Proc Natl Acad Sci U S A* **70**: 1564–1568.
- Haseltine, W.A., Block, R., Gilbert, W., and Weber, K. (1972) MSI and MSII made on ribosome in idling step of protein synthesis. *Nature* **238**: 381–384.
- Hashimoto, M., Nozoe, T., Nakaoka, H., Okura, R., Akiyoshi, S., Kaneko, K., *et al.* (2016) Noise-driven growth rate gain in clonal cellular populations. *Proc Natl Acad Sci* **113**: 3251–3256.
- Hauryliuk, V., Atkinson, G.C., Murakami, K.S., Tenson, T., and Gerdes, K. (2015) Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat Rev Microbiol* **13**: 298–309.
- Heerden, J.H. van, Wortel, M.T., Bruggeman, F.J., Heijnen, J.J., Bollen, Y.J.M., Planqué, R., *et al.* (2014) Lost in Transition: Start-Up of Glycolysis Yields Subpopulations of Nongrowing Cells. *Science* **343**: 1245114.
- Heinemann, M., Basan, M., and Sauer, U. (2020) Implications of initial physiological conditions for bacterial adaptation to changing environments. *Mol Syst Biol* **16**: e9965.
- Heins, A.-L., Lundin, L., Nunes, I., Gernaey, K.V., Sørensen, S.J., and Lantz, A.E. (2019) The effect of acetate on population heterogeneity in different cellular characteristics of *Escherichia coli* in aerobic batch cultures. *Biotechnol Prog* **35**: e2796.
- Heins, A.-L., and Weuster-Botz, D. (2018) Population heterogeneity in microbial bioprocesses: origin, analysis, mechanisms, and future perspectives. *Bioprocess Biosyst Eng* **41**: 889–916.
- Helaine, S., Cheverton, A.M., Watson, K.G., Faure, L.M., Matthews, S.A., and Holden, D.W. (2014) Internalization of *Salmonella* by Macrophages Induces Formation of Nonreplicating Persisters. *Science* **343**: 204–208.
- Helaine, S., Thompson, J.A., Watson, K.G., Liu, M., Boyle, C., and Holden, D.W. (2010) Dynamics of intracellular bacterial replication at the single cell level. *Proc Natl Acad Sci* **107**: 3746–3751.
- Hermesen, R., Okano, H., You, C., Werner, N., and Hwa, T. (2015) A growth-rate composition formula for the growth of *E. coli* on co-utilized carbon substrates. *Mol Syst Biol* **11**: 801.
- Hernández-Montalvo, V., Martínez, A., Hernández-Chavez, G., Bolivar, F., Valle, F., and Gosset, G. (2003) Expression of galP and glk in a *Escherichia coli* PTS mutant restores

- glucose transport and increases glycolytic flux to fermentation products. *Biotechnol Bioeng* **83**: 687–694.
- Hirsch, M., and Elliott, T. (2002) Role of ppGpp in rpoS Stationary-Phase Regulation in *Escherichia coli*. *J Bacteriol* **184**: 5077–5087.
- Hollenstein, K., Dawson, R.J., and Locher, K.P. (2007) Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol* **17**: 412–418.
- Hollinshead, W.D., Rodriguez, S., Martin, H.G., Wang, G., Baidoo, E.E.K., Sale, K.L., *et al.* (2016) Examining *Escherichia coli* glycolytic pathways, catabolite repression, and metabolite channeling using Δ pfk mutants. *Biotechnol Biofuels* **9**: 212.
- Huang, C.-J., Lin, H., and Yang, X. (2012) Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol* **39**: 383–399.
- Huh, D., and Paulsson, J. (2011) Random partitioning of molecules at cell division. *Proc Natl Acad Sci* **108**: 15004–15009.
- Iguchi, A., Thomson, N.R., Ogura, Y., Saunders, D., Ooka, T., Henderson, I.R., *et al.* (2009) Complete Genome Sequence and Comparative Genome Analysis of Enteropathogenic *Escherichia coli* O127:H6 Strain E2348/69. *J Bacteriol* **191**: 347–354.
- Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E., and Vogel, J. (2020) Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat Microbiol* **5**: 1202–1206.
- Inada, T., Kimata, K., and Aiba, H. (1996) Mechanism responsible for glucose–lactose diauxie in *Escherichia coli*: challenge to the cAMP model. *Genes Cells* **1**: 293–301.
- Ishihama, A., Shimada, T., and Yamazaki, Y. (2016) Transcription profile of *Escherichia coli*: genomic SELEX search for regulatory targets of transcription factors. *Nucleic Acids Res* **44**: 2058–2074.
- Ito, Y., Toyota, H., Kaneko, K., and Yomo, T. (2009) How selection affects phenotypic fluctuation. *Mol Syst Biol* **5**: 264.
- Jacob, F., and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**: 318–356.
- Jahn, M., Vorpahl, C., Türkowsky, D., Lindmeyer, M., Bühler, B., Harms, H., and Müller, S. (2014) Accurate Determination of Plasmid Copy Number of Flow-Sorted Cells using Droplet Digital PCR. *Anal Chem* **86**: 5969–5976.
- Jöers, A., and Tenson, T. (2016) Growth resumption from stationary phase reveals memory in *Escherichia coli* cultures. *Sci Rep* **6**: 24055.
- Johnson, J.R., Johnston, B., Kuskowski, M.A., Nougayrede, J.-P., and Oswald, E. (2008) Molecular epidemiology and phylogenetic distribution of the *Escherichia coli* pks genomic island. *J Clin Microbiol* **46**: 3906–3911.

- Kaiser, M., Jug, F., Julou, T., Deshpande, S., Pfohl, T., Silander, O.K., *et al.* (2018) Monitoring single-cell gene regulation under dynamically controllable conditions with integrated microfluidics and software. *Nat Commun* **9**: 1–16.
- Karp, P.D., Ong, W.K., Paley, S., Billington, R., Caspi, R., Fulcher, C., *et al.* (2018) The EcoCyc Database. *EcoSal Plus* **8**
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6504970/>. Accessed January 4, 2021.
- Kim, Y., and Wood, T.K. (2010) Toxins Hha and CspD and small RNA regulator Hfq are involved in persister cell formation through MqsR in Escherichia coli. *Biochem Biophys Res Commun* **391**: 209–213.
- Kiviet, D.J., Nghe, P., Walker, N., Boulineau, S., Sunderlikova, V., and Tans, S.J. (2014) Stochasticity of metabolism and growth at the single-cell level. *Nature* **514**: 376–379.
- Klapper, I., Gilbert, P., Ayati, B.P., Dockery, J., and Stewart, P.S. (2007) Senescence can explain microbial persistence. *Microbiol Read Engl* **153**: 3623–3630.
- Kochanowski, K., Volkmer, B., Gerosa, L., Rijsewijk, B.R.H. van, Schmidt, A., and Heinemann, M. (2013) Functioning of a metabolic flux sensor in Escherichia coli. *Proc Natl Acad Sci* **110**: 1130–1135.
- Koirala, S., Wang, X., and Rao, C.V. (2016) Reciprocal Regulation of l-Arabinose and d-Xylose Metabolism in Escherichia coli. *J Bacteriol* **198**: 386–393.
- Kolesov, G., Wunderlich, Z., Laikova, O.N., Gelfand, M.S., and Mirny, L.A. (2007) How gene order is influenced by the biophysics of transcription regulation. *Proc Natl Acad Sci* **104**: 13948–13953.
- Kopf, S.H., McGlynn, S.E., Green-Saxena, A., Guan, Y., Newman, D.K., and Orphan, V.J. (2015) Heavy water and ¹⁵N labelling with NanoSIMS analysis reveals growth rate-dependent metabolic heterogeneity in chemostats. *Environ Microbiol* **17**: 2542–2556.
- Kotte, O., Volkmer, B., Radzikowski, J.L., and Heinemann, M. (2014) Phenotypic bistability in Escherichia coli's central carbon metabolism. *Mol Syst Biol* **10**: 736.
- Kraemer, J.A., Sanderlin, A.G., and Laub, M.T. (2019) The Stringent Response Inhibits DNA Replication Initiation in E. coli by Modulating Supercoiling of oriC. *mBio* **10**
- Kumari, S., Tishel, R., Eisenbach, M., and Wolfe, A.J. (1995) Cloning, characterization, and functional expression of acs, the gene which encodes acetyl coenzyme A synthetase in Escherichia coli. *J Bacteriol* **177**: 2878–2886.
- Kwan, B.W., Valenta, J.A., Benedik, M.J., and Wood, T.K. (2013) Arrested protein synthesis increases persister-like cell formation. *Antimicrob Agents Chemother* **57**: 1468–1473.
- Labhsetwar, P., Cole, J.A., Roberts, E., Price, N.D., and Luthey-Schulten, Z.A. (2013) Heterogeneity in protein expression induces metabolic variability in a modeled Escherichia coli population. *Proc Natl Acad Sci U S A* **110**: 14006–14011.

- Laikova, O.N., Mironov, A.A., and Gelfand, M.S. (2001) Computational analysis of the transcriptional regulation of pentose utilization systems in the gamma subdivision of Proteobacteria. *FEMS Microbiol Lett* **205**: 315–322.
- Lambert, G., and Kussell, E. (2014) Memory and Fitness Optimization of Bacteria under Fluctuating Environments. *PLOS Genet* **10**: e1004556.
- Lambert, T.J. (2019) FPbase: a community-editable fluorescent protein database. *Nat Methods* **16**: 277–278.
- Lara, A.R., Leal, L., Flores, N., Gosset, G., Bolívar, F., and Ramírez, O.T. (2006) Transcriptional and metabolic response of recombinant Escherichia coli to spatial dissolved oxygen tension gradients simulated in a scale-down system. *Biotechnol Bioeng* **93**: 372–385.
- Larsson, G., Törnkvist, M., Wernersson, E.S., Trägårdh, C., Noorman, H., and Enfors, S.-O. (1996) Substrate gradients in bioreactors: origin and consequences. *Bioprocess Eng* **14**: 281–289.
- Lawlis, V.B., Dennis, M.S., Chen, E.Y., Smith, D.H., and Henner, D.J. (1984) Cloning and sequencing of the xylose isomerase and xylulose kinase genes of Escherichia coli. *Appl Environ Microbiol* **47**: 15–21.
- Lazzarini, R.A., Cashel, M., and Gallant, J. (1971) On the regulation of guanosine tetraphosphate levels in stringent and relaxed strains of Escherichia coli. *J Biol Chem* **246**: 4381–4385.
- Lee, C., Kim, J., Shin, S.G., and Hwang, S. (2006) Absolute and relative QPCR quantification of plasmid copy number in Escherichia coli. *J Biotechnol* **123**: 273–280.
- Lee, T.-H., and Maheshri, N. (2012) A regulatory role for repeated decoy transcription factor binding sites in target gene expression. *Mol Syst Biol* **8**: 576.
- Leighty, R.W., and Antoniewicz, M.R. (2013) COMPLETE-MFA: Complementary parallel labeling experiments technique for metabolic flux analysis. *Metab Eng* **20**: 49–55.
- Lemke, J.J., Sanchez-Vazquez, P., Burgos, H.L., Hedberg, G., Ross, W., and Gourse, R.L. (2011) Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proc Natl Acad Sci* **108**: 5712–5717.
- Lewis, M. (2005) The lac repressor. *C R Biol* **328**: 521–548.
- Lim, H.N., Lee, Y., and Hussein, R. (2011) Fundamental relationship between operon organization and gene expression. *Proc Natl Acad Sci* **108**: 10626–10631.
- Liu, H.-S., Jan, M.-S., Chou, C.-K., Chen, P.-H., and Ke, N.-J. (1999) Is Green Fluorescent Protein Toxic to the Living Cells? *Biochem Biophys Res Commun* **260**: 712–717.
- Liu, M., Durfee, T., Cabrera, J.E., Zhao, K., Jin, D.J., and Blattner, F.R. (2005) Global Transcriptional Programs Reveal a Carbon Source Foraging Strategy by Escherichia coli. *J Biol Chem* **280**: 15921–15927.

- Lobell, R.B., and Schleif, R.F. (1990) DNA looping and unlooping by AraC protein. *Science* **250**: 528–532.
- Loomis, W.F., and Magasanik, B. (1967) Glucose-Lactose Diauxie in *Escherichia coli*. *J Bacteriol* **93**: 1397.
- Madar, D., Dekel, E., Bren, A., Zimmer, A., Porat, Z., and Alon, U. (2013) Promoter activity dynamics in the lag phase of *Escherichia coli*. *BMC Syst Biol* **7**: 136.
- Magnusson, L.U., Farewell, A., and Nyström, T. (2005) ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol* **13**: 236–242.
- Magnusson, L.U., Gummesson, B., Joksimović, P., Farewell, A., and Nyström, T. (2007) Identical, Independent, and Opposing Roles of ppGpp and DksA in *Escherichia coli*. *J Bacteriol* **189**: 5193–5202.
- Maisonneuve, E., and Gerdes, K. (2014) Molecular mechanisms underlying bacterial persisters. *Cell* **157**: 539–548.
- Marchès, O., Nougayrède, J.-P., Boullier, S., Mainil, J., Charlier, G., Raymond, I., *et al.* (2000) Role of Tir and Intimin in the Virulence of Rabbit Enteropathogenic *Escherichia coli* Serotype O103:H2. *Infect Immun* **68**: 2171–2182.
- Martin, P., Marcq, I., Magistro, G., Penary, M., Garcie, C., Payros, D., *et al.* (2013) Interplay between Siderophores and Colibactin Genotoxin Biosynthetic Pathways in *Escherichia coli*. *PLOS Pathog* **9**: e1003437.
- Martins, B.M., and Locke, J.C. (2015) Microbial individuality: how single-cell heterogeneity enables population level strategies. *Curr Opin Microbiol* **24**: 104–112.
- Mayer, C., and Boos, W. (2005) Hexose/Pentose and Hexitol/Pentitol Metabolism. *EcoSal Plus* **1**.
- McClelland, H.L.O., Jones, C., Chubiz, L.M., Fike, D.A., and Bradley, A.S. (2020) Direct Observation of the Dynamics of Single-Cell Metabolic Activity during Microbial Diauxic Growth. *mBio* **11**.
- McDonald, T.P., Walmsley, A.R., and Henderson, P.J.F. (1997) Asparagine 394 in Putative Helix 11 of the Galactose-H⁺ Symport Protein (GalP) from *Escherichia coli* Is Associated with the Internal Binding Site for Cytochalasin B and Sugar. *J Biol Chem* **272**: 15189–15199.
- Million-Weaver, S., and Camps, M. (2014) Mechanisms of plasmid segregation: Have multicopy plasmids been overlooked? *Plasmid* **75**: 27–36.
- Miquel, S., Peyretailade, E., Claret, L., Vallée, A. de, Dossat, C., Vacherie, B., *et al.* (2010) Complete Genome Sequence of Crohn's Disease-Associated Adherent-Invasive *E. coli* Strain LF82. *PLOS ONE* **5**: e12714.

- Mok, W.W.K., Orman, M.A., and Brynildsen, M.P. (2015) Impacts of Global Transcriptional Regulators on Persister Metabolism. *Antimicrob Agents Chemother* **59**: 2713–2719.
- Møller, A.K., Leatham, M.P., Conway, T., Nuijten, P.J.M., Haan, L.A.M. de, Krogfelt, K.A., and Cohen, P.S. (2003) An Escherichia coli MG1655 lipopolysaccharide deep-rough core mutant grows and survives in mouse cecal mucus but fails to colonize the mouse large intestine. *Infect Immun* **71**: 2142–2152.
- Monod, J. (1942) *Recherches sur la croissance des cultures bactériennes*, Hermann & cie, Paris.
- Monod, J. (1949) The Growth of Bacterial Cultures. *Annu Rev Microbiol* **3**: 371–394.
- Morin, M., Ropers, D., Letisse, F., Laguerre, S., Portais, J.-C., Coccagn-Bousquet, M., and Enjalbert, B. (2016) The post-transcriptional regulatory system CSR controls the balance of metabolic pools in upper glycolysis of Escherichia coli. *Mol Microbiol* **100**: 686–700.
- Negrete, A., and Shiloach, J. (2017) Improving E. coli growth performance by manipulating small RNA expression. *Microb Cell Factories* **16**: 198.
- Nester, E.W., and Stocker, B. a. D. (1963) Biosynthetic Latency in Early Stages of Deoxyribonucleic Acid Transformation in Bacillus Subtilis. *J Bacteriol* **86**: 785–796.
- New, A.M., Cerulus, B., Govers, S.K., Perez-Samper, G., Zhu, B., Boogmans, S., *et al.* (2014) Different Levels of Catabolite Repression Optimize Growth in Stable and Variable Environments. *PLOS Biol* **12**: e1001764.
- Nguyen, D., Joshi-Datar, A., Lepine, F., Bauerle, E., Olakanmi, O., Beer, K., *et al.* (2011) Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria. *Science* **334**: 982–986.
- Ni, L., Tonthat, N.K., Chinnam, N., and Schumacher, M.A. (2013) Structures of the Escherichia coli transcription activator and regulator of diauxie, XylR: an AraC DNA-binding family member with a LacI/GalR ligand-binding domain. *Nucleic Acids Res* **41**: 1998–2008.
- Nikolic, N., Schreiber, F., Dal Co, A., Kiviet, D.J., Bergmiller, T., Littmann, S., *et al.* (2017) Cell-to-cell variation and specialization in sugar metabolism in clonal bacterial populations. *PLoS Genet* **13**: e1007122.
- Okeke, B.C., and Obi, S.K.C. (1994) Lignocellulose and sugar compositions of some agro-waste materials. *Bioresour Technol* **47**: 283–284.
- Orman, M.A., and Brynildsen, M.P. (2013) Dormancy Is Not Necessary or Sufficient for Bacterial Persistence. *Antimicrob Agents Chemother* **57**: 3230–3239.
- Ow, D.S.-W., Nissom, P.M., Philp, R., Oh, S.K.-W., and Yap, M.G.-S. (2006) Global transcriptional analysis of metabolic burden due to plasmid maintenance in Escherichia coli DH5α during batch fermentation. *Enzyme Microb Technol* **39**: 391–398.

- Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I., and Oudenaarden, A. van (2004) Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427**: 737–740.
- Paget, M.S.B., and Helmann, J.D. (2003) The sigma70 family of sigma factors. *Genome Biol* **4**: 203.
- Pelet, S., Rudolf, F., Nadal-Ribelles, M., Nadal, E. de, Posas, F., and Peter, M. (2011) Transient Activation of the HOG MAPK Pathway Regulates Bimodal Gene Expression. *Science* **332**: 732–735.
- Pérès, S.Y., Marchès, O., Daigle, F., Nougayrède, J.P., Herault, F., Tasca, C., *et al.* (1997) A new cytolethal distending toxin (CDT) from *Escherichia coli* producing CNF2 blocks HeLa cell division in G2/M phase. *Mol Microbiol* **24**: 1095–1107.
- Pin, C., Rolfe, M.D., Muñoz-Cuevas, M., Hinton, J.C., Peck, M.W., Walton, N.J., and Baranyi, J. (2009) Network analysis of the transcriptional pattern of young and old cells of *Escherichia coli* during lag phase. *BMC Syst Biol* **3**: 108.
- Pontrelli, S., Chiu, T.-Y., Lan, E.I., Chen, F.Y.-H., Chang, P., and Liao, J.C. (2018) *Escherichia coli* as a host for metabolic engineering. *Metab Eng* **50**: 16–46.
- Potrykus, K., and Cashel, M. (2008) (p)ppGpp: Still Magical? *Annu Rev Microbiol* **62**: 35–51.
- Radzikowski, J.L., Vedelaar, S., Siegel, D., Ortega, Á.D., Schmidt, A., and Heinemann, M. (2016) Bacterial persistence is an active σ^S stress response to metabolic flux limitation. *Mol Syst Biol* **12**: 882.
- Reuven, N.B., and Deutscher, M.P. (1993) Multiple exoribonucleases are required for the 3' processing of *Escherichia coli* tRNA precursors in vivo. *FASEB J Off Publ Fed Am Soc Exp Biol* **7**: 143–148.
- Revelles, O., Millard, P., Nougayrède, J.-P., Dobrindt, U., Oswald, E., Létisse, F., and Portais, J.-C. (2013) The carbon storage regulator (Csr) system exerts a nutrient-specific control over central metabolism in *Escherichia coli* strain Nissle 1917. *PloS One* **8**: e66386.
- Robert, L., Paul, G., Chen, Y., Taddei, F., Baigl, D., and Lindner, A.B. (2010) Pre-dispositions and epigenetic inheritance in the *Escherichia coli* lactose operon bistable switch. *Mol Syst Biol* **6**: 357.
- Rosenthal, A.Z., Qi, Y., Hormoz, S., Park, J., Li, S.H.-J., and Elowitz, M.B. (2018) Metabolic interactions between dynamic bacterial subpopulations. *eLife* **7**: e33099.
- Ruiz, L.M.R., Williams, C.L., and Tamayo, R. (2020) Enhancing bacterial survival through phenotypic heterogeneity. *PLOS Pathog* **16**: e1008439.
- Rydenfelt, M., Cox, R.S., Garcia, H., and Phillips, R. (2014) Statistical mechanical model of coupled transcription from multiple promoters due to transcription factor titration. *Phys Rev E* **89**: 012702.

- Saini, M., Lin, L.-J., Chiang, C.-J., and Chao, Y.-P. (2017) Synthetic Consortium of *Escherichia coli* for n-Butanol Production by Fermentation of the Glucose–Xylose Mixture. *J Agric Food Chem* **65**: 10040–10047.
- Sampaio, N.M.V., and Dunlop, M.J. (2020) Functional roles of microbial cell-to-cell heterogeneity and emerging technologies for analysis and control. *Curr Opin Microbiol* **57**: 87–94.
- Sandberg, T.E., Lloyd, C.J., Palsson, B.O., and Feist, A.M. (2017) Laboratory Evolution to Alternating Substrate Environments Yields Distinct Phenotypic and Genetic Adaptive Strategies. *Appl Environ Microbiol* **83**
- Schellenberg, G.D., Sarthy, A., Larson, A.E., Backer, M.P., Crabb, J.W., Lidstrom, M., *et al.* (1984) Xylose isomerase from *Escherichia coli*. Characterization of the protein and the structural gene. *J Biol Chem* **259**: 6826–6832.
- Schleif, R. (2003) AraC protein: a love-hate relationship. *BioEssays News Rev Mol Cell Dev Biol* **25**: 274–282.
- Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., *et al.* (2016) The quantitative and condition-dependent *Escherichia coli* proteome. *Nat Biotechnol* **34**: 104–110.
- Schmutzer, M., and Wagner, A. (2020) Gene expression noise can promote the fixation of beneficial mutations in fluctuating environments. *bioRxiv* 2020.02.13.947275.
- Shah, D., Zhang, Z., Khodursky, A.B., Kaldalu, N., Kurg, K., and Lewis, K. (2006) Persisters: a distinct physiological state of *E. coli*. *BMC Microbiol* **6**: 53.
- Shalel-Levanon, S., San, K.-Y., and Bennett, G.N. (2005) Effect of ArcA and FNR on the expression of genes related to the oxygen regulation and the glycolysis pathway in *Escherichia coli* under microaerobic growth conditions. *Biotechnol Bioeng* **92**: 147–159.
- Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005) A guide to choosing fluorescent proteins. *Nat Methods* **2**: 905–909.
- Shashkova, S., and Leake, M.C. (2017) Single-molecule fluorescence microscopy review: shedding new light on old problems. *Biosci Rep* **37**.
- Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U. (2002) Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* **31**: 64–69.
- Shimizu, K. (2013) Metabolic Regulation of a Bacterial Cell System with Emphasis on *Escherichia coli* Metabolism. *ISRN Biochem* **2013**: 645983.
- Shyp, V., Tankov, S., Ermakov, A., Kudrin, P., English, B.P., Ehrenberg, M., *et al.* (2012) Positive allosteric feedback regulation of the stringent response enzyme RelA by its product. *EMBO Rep* **13**: 835–839.
- Siegal, M.L. (2015) Shifting sugars and shifting paradigms. *PLoS Biol* **13**: e1002068.

- Silva, F., Queiroz, J.A., and Domingues, F.C. (2012) Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*. *Biotechnol Adv* **30**: 691–708.
- Şimşek, E., and Kim, M. (2018) The emergence of metabolic heterogeneity and diverse growth responses in isogenic bacterial cells. *ISME J* **12**: 1199–1209.
- Smith, A., Kaczmar, A., Bamford, R.A., Smith, C., Frustaci, S., Kovacs-Simon, A., *et al.* (2018) The Culture Environment Influences Both Gene Regulation and Phenotypic Heterogeneity in *Escherichia coli*. *Front Microbiol* **9**
- Solopova, A., Gestel, J. van, Weissing, F.J., Bachmann, H., Teusink, B., Kok, J., and Kuipers, O.P. (2014) Bet-hedging during bacterial diauxic shift. *Proc Natl Acad Sci* **111**: 7427–7432.
- Song, S., and Park, C. (1997) Organization and regulation of the D-xylose operons in *Escherichia coli* K-12: XylR acts as a transcriptional activator. *J Bacteriol* **179**: 7025–7032.
- Sooriyaarachchi, S., Ubhayasekera, W., Park, C., and Mowbray, S.L. (2010) Conformational Changes and Ligand Recognition of *Escherichia coli* d-Xylose Binding Protein Revealed. *J Mol Biol* **402**: 657–668.
- Sozhamannan, S., and Stitt, B.L. (1997) Effects on mRNA degradation by *Escherichia coli* transcription termination factor Rho and pBR322 copy number control protein rop11 Edited by M. Gottesman. *J Mol Biol* **268**: 689–703.
- Storz, G., Opdyke, J.A., and Zhang, A. (2004) Controlling mRNA stability and translation with small, noncoding RNAs. *Curr Opin Microbiol* **7**: 140–144.
- Strugeon, E., V, T., Mc, P., and S, D.R. (2016) The Stringent Response Promotes Antibiotic Resistance Dissemination by Regulating Integron Integrase Expression in Biofilms. *Mbio* **7**
- Stülke, J., and Hillen, W. (1999) Carbon catabolite repression in bacteria. *Curr Opin Microbiol* **2**: 195–201.
- Subach, O.M., Cranfill, P.J., Davidson, M.W., and Verkhusha, V.V. (2011) An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore. *PLOS ONE* **6**: e28674.
- Sumiya, M., Davis, E.O., Packman, L.C., McDonald, T.P., and Henderson, P.J. (1995) Molecular genetics of a receptor protein for D-xylose, encoded by the gene xylF, in *Escherichia coli*. *Receptors Channels* **3**: 117–128.
- Swinnen, I.A.M., Bernaerts, K., Dens, E.J.J., Geeraerd, A.H., and Van Impe, J.F. (2004) Predictive modelling of the microbial lag phase: a review. *Int J Food Microbiol* **94**: 137–159.
- Takhaveev, V., and Heinemann, M. (2018) Metabolic heterogeneity in clonal microbial populations. *Curr Opin Microbiol* **45**: 30–38.

- Tang, F., Barbacioru, C., Nordman, E., Li, B., Xu, N., Bashkirov, V.I., *et al.* (2010) RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat Protoc* **5**: 516–535.
- Toulokhonov, I.I., Shulgina, I., and Hernandez, V.J. (2001) Binding of the Transcription Effector ppGpp to Escherichia coli RNA Polymerase Is Allosteric, Modular, and Occurs Near the N Terminus of the β' -Subunit. *J Biol Chem* **276**: 1220–1225.
- Traxler, M.F., Chang, D.-E., and Conway, T. (2006) Guanosine 3',5'-bispyrophosphate coordinates global gene expression during glucose-lactose diauxie in Escherichia coli. *Proc Natl Acad Sci* **103**: 2374–2379.
- Traxler, M.F., Summers, S.M., Nguyen, H.-T., Zacharia, V.M., Hightower, G.A., Smith, J.T., and Conway, T. (2008) The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. *Mol Microbiol* **68**: 1128–1148.
- Tripathi, A., Dewan, P.C., Siddique, S.A., and Varadarajan, R. (2014) MazF-induced growth inhibition and persister generation in Escherichia coli. *J Biol Chem* **289**: 4191–4205.
- Valgepea, K., Adamberg, K., Nahku, R., Lahtvee, P.-J., Arike, L., and Vilu, R. (2010) Systems biology approach reveals that overflow metabolism of acetate in Escherichia coli is triggered by carbon catabolite repression of acetyl-CoA synthetase. *BMC Syst Biol* **4**: 166.
- Veening, J.-W., Hamoen, L.W., and Kuipers, O.P. (2005) Phosphatases modulate the bistable sporulation gene expression pattern in Bacillus subtilis. *Mol Microbiol* **56**: 1481–1494.
- Veening, J.-W., Smits, W.K., and Kuipers, O.P. (2008) Bistability, Epigenetics, and Bet-Hedging in Bacteria. *Annu Rev Microbiol* **62**: 193–210.
- Verstraeten, N., Knapen, W.J., Kint, C.I., Liebens, V., Van den Bergh, B., Dewachter, L., *et al.* (2015) Obg and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. *Mol Cell* **59**: 9–21.
- Wang, J., Atolia, E., Hua, B., Savir, Y., Escalante-Chong, R., and Springer, M. (2015) Natural Variation in Preparation for Nutrient Depletion Reveals a Cost–Benefit Tradeoff. *PLoS Biol* **13**: e1002041.
- Wang, X., Xia, K., Yang, X., and Tang, C. (2019) Growth strategy of microbes on mixed carbon sources. *Nat Commun* **10**: 1–7.
- Wang, Y., Bojer, M.S., George, S.E., Wang, Z., Jensen, P.R., Wolz, C., and Ingmer, H. (2018) Inactivation of TCA cycle enhances Staphylococcus aureus persister cell formation in stationary phase. *Sci Rep* **8**: 1–13.
- Wolfe, A.J. (2005) The Acetate Switch. *Microbiol Mol Biol Rev* **69**: 12–50.
- Wolfe, A.J. (2015) Glycolysis for Microbiome Generation. *Microbiol Spectr* **3**.
- Wolfsberg, E., Long, C.P., and Antoniewicz, M.R. (2018) Metabolism in dense microbial colonies: ¹³C metabolic flux analysis of E. coli grown on agar identifies two distinct cell populations with acetate cross-feeding. *Metab Eng* **49**: 242–247.

Wovcha, M.G., Steuerwald, D.L., and Brooks, K.E. (1983) Amplification of D-xylose and D-glucose isomerase activities in *Escherichia coli* by gene cloning. *Appl Environ Microbiol* **45**: 1402–1404.

Yaginuma, H., Kawai, S., Tabata, K.V., Tomiyama, K., Kakizuka, A., Komatsuzaki, T., *et al.* (2014) Diversity in ATP concentrations in a single bacterial cell population revealed by quantitative single-cell imaging. *Sci Rep* **4**: 6522.

Yoon, S.H., Han, M.-J., Jeong, H., Lee, C.H., Xia, X.-X., Lee, D.-H., *et al.* (2012) Comparative multi-omics systems analysis of *Escherichia coli* strains B and K-12. *Genome Biol* **13**: R37.

Zambrano, M.M., Siegele, D.A., Almiron, M., Tormo, A., and Kolter, R. (1993) Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. *Science* **259**: 1757–1760.

Zheng, D., Constantinidou, C., Hobman, J.L., and Minchin, S.D. (2004) Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Res* **32**: 5874–5893.

Zhou, K., Zhou, L., Lim, Q. 'En, Zou, R., Stephanopoulos, G., and Too, H.-P. (2011) Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by quantitative PCR. *BMC Mol Biol* **12**: 18.

Zhou, Y., and Gottesman, S. (1998) Regulation of Proteolysis of the Stationary-Phase Sigma Factor RpoS. *J Bacteriol* **180**: 1154–1158.

Zukin, R.S., Strange, P.G., Heavey, R., and Koshland, D.E. (1977) Properties of the galactose binding protein of *Salmonella typhimurium* and *Escherichia coli*. *Biochemistry* **16**: 381–386.