

Chapitre 1

Modèle numérique du métabolisme bactérien de la production d'un fromage

Soumis dans metabolic engineering

Les données multi-omiques permettent une compréhension du métabolisme à différents niveaux que sont : métatranscriptomiques, l'étude de l'expression de tous les gènes d'une communauté données ; métabolomiques, l'étude de l'ensemble des métabolites ; génomiques ; les génomes de chaque souche bactérienne. Dans ce chapitre nous verrons comment les données omiques, le paramétrage de chaque souche bactérienne ainsi qu'un partage équilibré des ressources pour chaque membre de la communauté permet d'obtenir un modèle numérique de la fermentation bactérienne d'un fromage modèle précis mettant en avant des interactions bactériennes.

1.1 Introduction

La fermentation des aliments et des boissons par un consortium bactérien compte parmi les plus vieux processus technologiques alimentaires, associés aux bénéfices alimentaires et utilisés aujourd'hui à la fois dans les contextes traditionnels et industriels (?). La transformation alimentaire par les micro-organismes est le résultat de processus métaboliques et d'interactions entre les membres de la communauté bactérienne. Cette dernière peut être contrôlée, comme c'est le cas dans les processus industriels (?), afin de garantir des produits alimentaires sains et prévoir des propriétés organoleptiques (?). Au sein de communautés, les propriétés fonctionnelles des souches bactériennes associées aux produits fermentés sont mieux comprises (?), en revanche, les mécanismes moléculaires précis mis en jeu commencent seulement à être explorés (?). En effet, la compréhension du métabolisme individuel des membres d'un consortium n'est pas suffisante pour capturer ou prédire les comportements globaux de la communauté, qui est impactée par les interactions microbe-microbe.

Un moyen pour caractériser toutes les fonctions et les interactions d'un consortium microbien de produits fermentés est de combiner les approches en biologie des systèmes avec les données omiques (?). Une première approche consiste à procéder à l'inventaire des métabolites et espèces d'une communauté via des méthodes de cultures dépendantes ou indépendantes(?). Dans le cas d'une communauté contrôlée, une comparaison en mono et en co-cultures peut nous informer sur, l'impact des interactions, sur la dynamique et

sur les profils métaboliques de la communauté (?). Du point de vue des méthodes non culture-dépendantes, l'utilisation de modèles métaboliques à l'échelle du génome (GEMs) (?) semble particulièrement opportun comme un moyen d'associer les souches du consortium à la fonction portée par les systèmes et à termes aux propriétés organoleptiques alimentaires (?). Ces modèles de consortium bactérien sont utilisables pour un grand nombre d'application comme dans les biotechnologies alimentaires, allant de l'assemblage de souches dans la création de communautés à l'optimisation de procédés pour les propriétés organoleptiques désirées (?).

L'analyse de consortia bactérien permet difficilement la construction de modèle mécanistique ou d'un jumeau numérique de part la complexité de la communauté bactérienne, ainsi que la grande spécificité souches-dépendantes des voies métaboliques secondaires impliqué dans les processus organoleptiques. Ainsi, travailler à petite échelle avec des systèmes contrôlés où la complexité est réduite semble être particulièrement pertinent (?). Actuellement, même sur des systèmes simples une modélisation temporelle peut être pris en compte, comme c'est le cas pour la fermentation qui est un processus dynamique.

Dans ce chapitre, nous vous présenterons un modèle simplifié de la fabrication du fromage, à partir des éléments nutritifs du lait jusqu'à l'étape d'affinage. Ce modèle se concentre sur le métabolisme d'une communauté bactérienne composés de deux bactéries lactiques (LAB) *Lactococcus lactis*, *Lactiplantibacillus plantarum* et d'une bactéries propionique *Propionibacterium freudenreichii* permettant de la production de fromage industriel. Nous nous sommes basés sur les travaux de (?), décrivant comment le paramétrage d'un processus standardisé de fabrication d'un fromage peut moduler la production de composés aromatiques. Pour cela, les réseaux métaboliques à l'échelle du génome ainsi que le modèle communautaire fut construit et exploité grâce au données omiques dans le but de (i) évaluer la contribution métabolique de chaque espèce microbienne durant la fabrication de fromage, (ii) de simuler les dynamiques des métabolismes microbiens ainsi que l'effet de la communauté sur le comportement des individus et enfin (iii) proposer des interactions possibles entre les espèces bactériennes.

La méthode consiste à identifier les chemin métaboliques responsables de la transformation des nutriments présent initialement dans le lait. Nous avons par la suite optimisé

chacun de nos modèles métaboliques individuels afin d'être capable de reproduire la dynamique des cultures pures. Enfin, ces modèles individuels sont utilisés pour construire un modèle dynamique de la communauté prédisant avec précision la production de composés aromatiques du fromage.

1.2 Résultats

1.2.1 Les modèles métaboliques individuels identifient des voies métaboliques spécifiques dans le lait

Dans cette première sous-section, nous nous focaliserons sur le processus d'affinage des modèles de flux individuels de chaque espèce. En effet, l'hypothèse que nous explorons consiste à suffisamment nettoyer les réseaux métaboliques à l'échelle du génome pour construire un modèle numérique de la communauté capable de caractériser une communauté bactérienne à la fois sur le plan des interactions bactériennes et à la fois sur des biens métaboliques. Pour cela, nous avons commencé par une étude comparative des génomes ainsi que des réseaux métabolique à l'échelle du génome afin de vérifier la qualité topologique des réseaux. Dans un second temps, nous avons examiné le métabolisme carboné central de chaque souche en prêtant une attention particulière au métabolisme secondaire dans le but de vérifier la production et la consommation de certains métabolites se fait par les bonnes voies métaboliques. Enfin, dans le but de obtenir des valeurs de biomasse, de production et de consommation numériquement proches des observations biologiques, une étape d'inférence de paramètres a été faite souche dépendante.

La communauté bactérienne étudiée est composé de *Propionibacterium freudenreichii* CIRM-BIA122, *Lactiplantibacillus plantarum* CIRMB-BIA465 et de *Lactococcus lactis* biovar *diacetyl lactis*. Pour chacun d'entre eux, le longueur total de l'ADN a été récupéré et les génomes bactériens ont été assemblés à partir d'une combinaison de lectures courtes et longues, produisant un assemblage précis et complet. Cet assemblage a révélé un chromosome circulaire et la présence de sept plasmides pour la souche de *L. lactis* et quatre

plasmides pour la souche de *L. plantarum*. Parmi ce consortium bactérien seul la souche de *P. freudenreichii* ne porte pas de plasmide. Une annotation du génome a été faite et déposé sous le numero d'accesion ENA suivant : ENA : PRJEB42478. L'annotation a identifié une liste de 8234 séquences codantes incluant ceux du chromosome (2460 pour *P. freudenreichii*, 2417 for *L. lactis* et 3016 pour *L. plantarum*) et des plasmides (139 pour *L. lactis* et 82 pour *L. plantarum*). D'autres informations telles que le nombre de gènes codant pour des protéines, de séquences ribosomiques, tRNA et le pourcentage de GC sont rassemblés dans la Table 1.1.

Les réseaux métaboliques à l'échelle du génomes ont été reconstruit pour chaque souche individuelle depuis leurs séquences génomiques annotées respectives. Après une première étape automatique de reconstruction, nous avons réalisé un raffinement manuel du métabolisme carboné central basé sur la littérature et la connaissance biologique dans le but de prendre en compte la spécificité de chaque souche bactérienne. La qualité de chaque réseau métabolique a été vérifié et répond aux attendus : une faible fraction de réactions universellement bloqués ($< 3\%$), et une faible fraction de réaction sans association de gène-proteine-réactions (GPR) ($< 30\%$) (?). La Table 1.2 décrit les caractéristiques des trois modèles et nous observons que *L. lactis* a le plus petit nombre de réactions que les deux autres souches, une différence qualitative qui est aussi observée pour les autres souches de la même espèces dans d'autres bases de données (?).

Nous avons construit les réseaux métaboliques à l'échelle de chaque génome (GEMs) les avons analysés par la méthode d'équilibre des flux (FBA) (??) assurant ainsi la production des molécules de la biomasse à partir de l'environnement nutritionnel du lait (voir Materiels et méthodes). Nous avons obtenu des taux de croissances allant de $0.187 \text{ mmol.gDW}^{-1}.\text{hr}^{-1}$ à $2.17 \text{ mmol.gDW}^{-1}.\text{hr}^{-1}$, où le premier est obtenu par *P. freudenreichii* et la plus haute valeur obtenu par *L. lactis*. Comme les conditions dans le fromage correspondent à des conditions microaérobiques (?), une attention particulière a été donné sur le rôle de l'oxygène sur le métabolisme des trois bactéries en modifiant la borne supérieur de la réaction d'importation dans le compartiment cytosolique.

Le métabolisme carboné central produit du NADH and NADPH et ses formes réduites correspondantes, de l'ATP et les précurseurs nécessaires pour la biosynthèse des molécules essentielles pour la croissance ainsi que pour l'activité métabolique des cellules bactériennes, e.g. acides aminés, purine, pyrimidine, glycerol-3-phosphate, acides gras, N-acetyl-glucosamine, vitamines. Parmi les voies métaboliques du métabolisme carboné central, la glycolyse et la voie des pentose phosphate sont partagés entre les trois espèces presque à l'identique. Les sources de carbone telles que le lactose, l'acide citrique et l'acide lactique, sont converties en acide pyruvique, qui sera utilisé *via* d'autres voies spécifiques. La spécificité des voies métaboliques de chaque espèce du métabolisme carboné central est représenté par la Figure 1.1. Dans la suite du chapitre, les caractéristiques de chaque souche métabolique sera étudié afin de justifier le raffinement fait, en se focalisant sur les voies métaboliques secondaires permettant la production de composés organoleptiques.

1.2.1.1 Métabolisme des bactéries lactiques

Les bactéries lactiques *L. lactis* et *L. plantarum* utilisent le lactose et le citrate, les deux principales sources de carbone du lait (?). Le lactose, dégradé par l'enzyme de la beta-galactosidase, se transforme en galactose et glucose, où ce dernier alimente la glycolyse. Concernant le galactose et l'acide citrique, leur métabolisme est souche dépendant (?). D'un côté, le lactose est convertie en glucose-6-phosphate *via* la voie métabolique de Leloir avant d'entrer dans la glycolyse, et de l'autre l'acide citrique est convertie en pyruvate par l'enzyme de la citrate lyase à travers la production des acides lactiques et citriques (?). Du point de vue du métabolisme secondaire, le butanediol et du diacétyle, responsable de la saveur beurré du fromage, est produit chez les deux modèles à partir du pyruvate et est permise grâce à la voie de l'acétolactate. L'ensemble de ces voies métaboliques décrites ci-dessus sont récapitulés dans le Figure 1.1 a, b).

1.2.1.1.1 *L. lactis* Pour permettre la production de certains composé d'intérêt, comme le butanediol, le GEM a du être manuellement curé (Fig. 1.1 a). Le flux de l'acétolaldehydehydrogenase a été bloqué (ACTD2) permettant la production du butanediol, et les

bornes de l'acetolactate decarboxylase (ACLDC), acetolactate synthase (ACLS) ont été modifié permettant l'activation de la voie de l'acetolactate, comme retrouvé dans la littérature (???). Enfin, la consommation de lactose a été régulée en modifiant les bornes de la réaction LACZ qui a permis un flux de production de lactate pour *L. lactis* (Fig. 1.1 a). Pour valider ce modèle, nous avons considéré un exemple d'activation de voie métabolique durant la production du fromage. Il est montré que *L. lactis* consomme le lactose à la fois par la voie du tagatose, encodée par l'opéron plasmidique *lacABCD*, ou par la voie de Leloir, encodée par l'opéron du chromosome *galKTE*. Au regard des données metatranscriptomiques, *L. lactis* expriment les gènes associés à la voie du tagatose au début de la production de fromage quand la concentration de lactose est grande, et ceux de la voie de Leloir durant l'affinage quand la concentration de lactose est faible. Cela suggère fortement que la concentration de lactose détermine quelle voie métabolique est employée, et prédit que une grande quantité de lactose devrait produire un flux important de lactose à travers la voie du tagatose. Le modèle FBA de *L. lactis* mène à un plus grand flux de lactose à travers la voie du tagatose lors de l'inoculation et ainsi, coïncide avec les attentes (Fig. 1.2 a).

1.2.1.1.2 *L. plantarum* L'exemple pris dans le cas de *L. plantarum* concerne la dégradation du glucose et du galactose dans le cas de la fermentation homolactique et hétérolactique. Dans les deux configurations, le pyruvate est produit à partir de la dégradation du glucose et du galactose. Dans la fermentation homolactique, le pyruvate est associé au relargage de l'acide lactique dans le milieu extracellulaire. Concernant la fermentation hétérolactique, on observe en plus une production de l'éthanol et de l'acétate. Nous savons que *L. plantarum* est une bactérie hétérolactique pouvant produire de l'acétate grâce à l'enzyme spécifique de la transketolase (numéro EC 4.1.2.9) décrit dans (?) à travers la voie métabolique de l'acétate kinase. Pour permettre cette production dans notre modèle, la pyruvate formate lyase (PFL) fut bloquée (?). Concernant le butanediol, la même curation fut faite que pour *L. lactis*. Enfin, Pour reproduire la dégradation du galactose et du glucose par la voie de Leloir et de la glycolyse grâce à l'enzyme de la β -galactosidase *LACZ*, la réaction *LCTSt3ipp* fut bloquée et la réaction UTP-glucose-1-phosphate uri-

glyltransferase (*GALUi*) est rendu réversible pour compléter la voie de dégradation du galactose.

1.2.1.2 Le métabolisme de *P. freudenreichii*

La souche CIRM-BIA122 de *P. freudenreichii* consomme à la fois l'acide lactique, sa principale source de carbone (??) ou le lactose (?). L'acide lactique et le lactose sont catabolisés par les voies de la glycolyse et des penthose-phosphate, menant à la production de pyruvate. Une part du pyruvate alimente le cycle des tricarboxylique (TCA) et la voie de Wood-Werkman (Fig. 1.1 c) durant laquelle l'acide propionique peut être produit. Cette voie métabolique est spécifique au genre *Propionibacterium* et est assurée en redirigeant le succinate à partir du TCA. La production de l'acide acétique et du CO₂ (?) à partir du pyruvate est permis grâce à l'activité de deux enzymes : la pyruvate dehydrogenase et la decarboxylative oxidation.

Pour assurer la proportion correct de propionate sur acétate, le GEM de *P. freudenreichii* a requis de plus amples curation. En effet, les bornes des réactions *PPAKr,PTAr,2131pyrpp* ont été modifiées pour diriger la production du propionate dans sa voie métabolique de prédilection, i.e. Wood-Werkman, et la réaction *POX2* fut bloquée pour réguler le flux d'acétate.

Un moyen de valider nos modifications est de calculer le ratio de propionate sur l'acétate afin de vérifier si ce dernier coïncide avec les dernières observations biologiques. Nous avons obtenu la valeur de 2.19, ce qui est confirmé par ratio de 1 :2 trouvé dans les papiers de (????).

Le raffinement des trois modèles FBA a permis la prédiction de plusieurs composés d'arômes tels que les acides lactiques, acétique et propionique, le diacétyl et le butandiol (??) (Fig. 1.1). En plus, les spécificités métaboliques de chaque souche en fonction de leur métabolisme primaire, i.e. la voie de Wood-Werkman pour *P. freudenreichii*, la voie du tagatose pour *L. lactis* et la voie de la transketolase pour *L. plantarum* ont été reproduit fidèlement et avec précision.

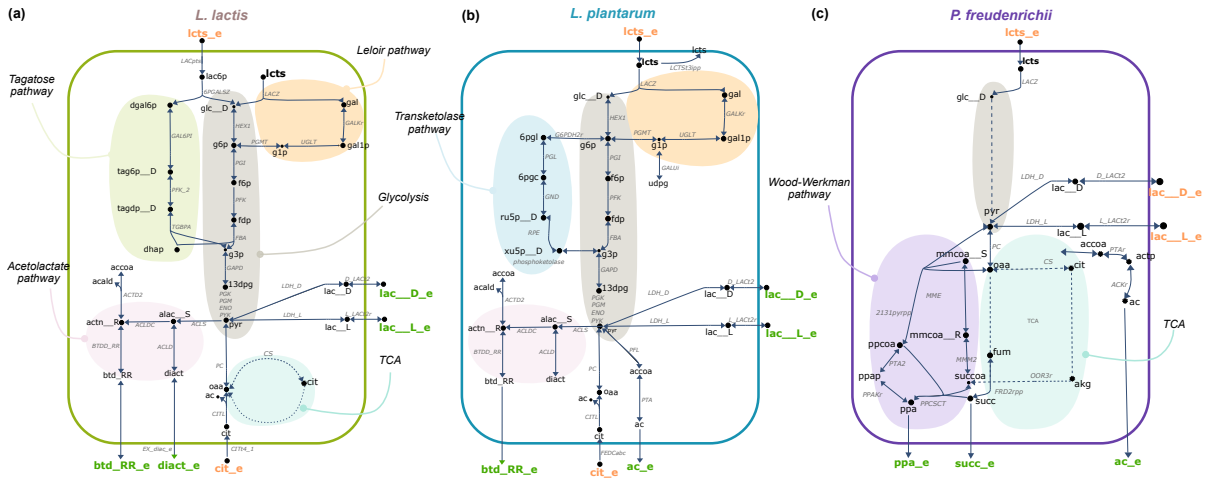


FIGURE 1.1 – Cartes des voies métaboliques spécifiques et partagées parmi *L. lactis*(a), *L. plantarum*(b) et *P. freudenreichii*(c). La voie métabolique de la glycolyse (couleur gris) est partagé par les trois espèces, alors que celle de Leloir (couleur saumon), acétolactate (couleur rose) et du citrate sont partagés par les deux bactéries lactiques. Le cycle du TCA (en bleu clair) est partagé entre *P. freudenreichii* et *L. lactis*. Les voies du tagatose (en vert), de la transketolase (en bleu) et de Wood-Werkman (en violet) sont respectivement spécifiques de *L. lactis*, *L. plantarum* et de *P. freudenreichii*. Enfin, les métabolites en orange et vert correspondent aux entrées et sorties des modèles métaboliques. **Abréviations des composés :** 13dpg : 3-Phospho-D-glycérolyl phosphate; 6pgc : 6-Phospho-D-gluconate; 6pgl : 6-phospho-D-glucono-1,5-lactone; ac : acetate; acald : acétaldéhyde; accoa : acetyl-coa; actn__R : acétoïne; ak : 2-Oxoglutarate; alac__S : (S)-2-Acétolactate; btd__RR : butanediol; cit : citrate; dgald : d-galactose-6-phosphate; dhap : Dihydroxyacétone phosphate; diact : diacetyl; f6p : fructose-6-phosphate; fdp : D-Fructose 1,6-bisphosphate; fum : fumarate; g1p : glucose-1-phosphate; g3p : Glyceraldéhyde 3-phosphate; g6p : glucose-6-phosphate; gal : galactose; gal1p : galactose-1-phosphate; glc__D : glucose; lac__D, lac__L : lactate; lac6p : lactose-6-phosphate; lcts : lactose; mmcoa__R : (R)-Méthylmalonyl-CoA; mmcoa__S : (S)-Méthylmalonyl-CoA; oaa : Oxaloacétate; ppa : propionate; ppap : Propanoyl phosphate; ppcoa : propionyl-coa; pyr : pyruvate; ru5p__D : D-Ribulose 5-phosphate; succ : succinate; succoa : succinyl-coa; tag6p__D : D-tagatose-6-phosphate; tagdp__D : D-Tagatose 1,6-bisphosphate; udpg : UDPglucose; xu5p__D : D-Xylulose 5-phosphate. **Abréviations des réactions :** 2131pyrpp : Méthylmalonyl-CoA carboxyltransferase 5S subunit; 6PGALSZ : 6-phospho-beta-galactosidase; ACLD : Acétolactate decarboxylase; ACTD2 : Acétoïne déshydrogénase; ACLDC : Acétolactate decarboxylase; ACLS : Acétolactate synthase; BTDD__RR : R R butanediol déshydrogénase; CITL : Citrate lyase; CITt4__1 : Citrate transport via sodium symport; CS : Citrate synthase; D_LACT2 : D lactate transport via proton symport; ENO : Enolase; FBA : Fructose-bisphosphate aldolase; FEDCabc : FEDCabc; FRD2rpp : Fumarate reductase / succinate déshydrogénase (irréversible) (periplasmic, membrane potential dissipating); G6PDH2r : Glucose 6-phosphate déshydrogénase; GAL6PI : Galactose-6-phosphate isomérase; GALkr : Galactokinase; GALUi : UTP-glucose-1-phosphate uridylyltransferase (irréversible); GAPD : Glyceraldéhyde-3-phosphate déshydrogénase; GND : Phosphogluconate déshydrogénase; HEX1 : Hexokinase (D-glucose :ATP); L_LACT2r : L lactate réversible transport via proton symport; LACpts : Lactose transport via PEP :Pyr PTS; LACZ : B-galactosidase; LCTSt3ipp : Lactose transport via proton antiport (periplasm); LDH__D : D-lactate déshydrogénase; LDH__L : L-lactate déshydrogénase; MME : Méthylmalonyl-CoA épimérase; MMM2 : Méthylmalonyl-CoA mutase; OOR3r : 2-oxoglutarate synthase (rev); PC : Pyruvate carboxylase; PFK : Phosphofructokinase; PFK__2 : Phosphofructokinase; PFL : Pyruvate formate lyase; PGI : Glucose-6-phosphate isomérase; PGK : Phosphoglycérate kinase; PGL : 6-phosphogluconolactonase; PGM : Phosphoglycérate mutase; PGMT : Phosphoglucomutase; phosphoketolase : phosphoketolase; PPAkr : Propionate kinase; PPCSCT : Propanoyl-CoA : succinate CoA-transferase; PTA : Phosphotransacétyle; PTA2 : Phosphate acétyltransferase; PYK : Pyruvate kinase; RPE : Ribulose 5-phosphate 3-épimérase; TGBPA : Tagatose-bisphosphate aldolase; UGLT : UDPglucose-hexose-1-phosphate uridylyltransferase

Numéro Bioproject	PRJEB42478		
Nom des génomes	<i>Propionibacterium freudenreichii</i>	<i>Lactiplantibacillus plantarum</i>	<i>Lactococcus lactis</i>
Nom des souches	CIRM-BIA122	CIRM-BIA465	<i>lactis</i> biovar <i>diacetylactis</i> CIRM-BIA1206
Identifiant Taxon	1744	1590	1358
Numéro d'accèsion	ERS5564379	ERS5564378	ERS5564377
Nombre de plasmides	∅	4	7
Taille des plasmides (bases)	∅	p1 : 40748, p2 : 30463, p3 : 9152, p4 : 2012	p1 : 40748, p2 : 30463, p3 : 9152, p4 : 2012, p5 : 30463, p6 : 9152, p7 : 2012
Taille des génomes circulaires (bases)	2622405	3121980	2365039
Nombre de ARNr	6	16	19
Nombre d'ARNt	45	67	62
Pourcentage de GC (génom, min & max dans les plasmides)	67	44 (36-41)	35 (30-36)
Nombre de CDS	2472	3112	2563
Nombre de gènes dans chaque réplicon	chromosome : 2557	p1 : 41, p2 : 27, p3 : 13, p4 : 2 chromosome : 3157	p1 : 69, p2 : 43, p3 : 14, p4 : 9, p5 : 5, p6 : 4, p7 : 3, chromosome : 2542

TABLE 1.1 – Vue d'ensemble des génomes de *P. freudenreichii*, *L. lactis* and *L. plantarum* après l'assemblage

	<i>P. freudenreichii</i>	<i>L. plantarum</i>	<i>L. lactis</i>
Nombre de gènes	1473	1433	1272
Nombre de métabolites	1284	1045	939
Nombre de réactions	1790	1523	1337
Pourcentage de réactions associées à au moins un gène	76.6	79.0	82.1
Taux de croissance en mmol.gDW ⁻¹ .hr ⁻¹	0.187	0.645	2.172

TABLE 1.2 – Caractéristiques

1.2.2 Optimisation dynamique des modèles individuels

[[ML – Parler un peu plus haut de la méthode itérative sur le raffinement](#)] La méthode itérative a permis d’imiter, de manière qualitative, la production et la consommation de métabolites connus de la littérature ainsi que de prédire de nouveaux comportements métaboliques, comme la dégradation du citrate chez *L. plantarum* par les enzymes de la famille CITL. A ce stade là, aucune précision numérique n’est modélisée ne permettant pas d’expliquer les observations biologiques avec précision. Dans la continuité de notre hypothèse, i.e. un assemblage de modèle suffisamment bien curé permet d’expliquer les comportement de la communauté, nous avons inféré un faible nombre de paramètres spécifiques pour chaque souche bactérienne. Chaque souche sera ainsi en mesure d’expliquer les observations biologiques de culture pure, c’est à dire, la croissance bactérienne, les courbes de pH pour *L. lactis* et *L. plantarum* ainsi que les métabolites dosés pour *P. freudenreichii*. Dans les paragraphes suivants, les résultats biologiques seront d’abord présentés, puis notre stratégie numérique et enfin nos résultats numériques.

Chacune des souches microbiennes a été mise en culture et des points de croissance pour toutes les souches à différents pas de temps ont été récupéré. Pour les deux bactéries lactiques, nous avons en plus des données des valeurs de pH et pour *P. freudenreichii*, des concentrations de métabolites au temps finale. Chacune des cultures pures ont respecté les conditions microaerophilique en lait (Supp. Table ??), et le milieu de culture de *P. freudenreichii* a été suppléé en lactate pour imiter la croissance en co-culture et d’un

hydrolyzat de protéine du lait simulant sa protéolyse. En prenant en compte toutes ces conditions, chacune des souches a atteint un seuil de cultivabilité au dessus de $9 \log_{10}$ CFU/g. Le pH, d'une valeur de 6.7 lors de l'inoculation des deux bactéries lactiques, a atteint respectivement 5.7 et 5.1 pour *L. plantarum* et *L. lactis*. Pour respectivement *P. freudenreichii*, *L. plantarum* et *L. lactis*, le début de la phase plateau, correspondant à l'entrée dans la phase stationnaire, a mis respectivement quarante huit, quatorze et huit heures. Les données de dosages des métabolites de *P. freudenreichii* montrent une grande production de propionate (huit grammes par litre de lait), une forte consommation de lactate (neuf grammes par litre de lait) et une faible production d'acétate et de succinate, respectivement trois et zéro sept grammes par litre de lait et zéro et trois cents soixante et onze gramme par litre de lait ??).

La prochaine étape de curation de chaque modèle individuel consiste à obtenir une précision numérique, c'est à dire, pourvoir expliquer les données de culture pure décrit précédemment. Dans un premier temps, nous avons ajusté les flux de métabolites présent dans le milieu de façon à être cohérent avec les composés du lait : concernant *L. lactis* les bornes inférieure de *EX_glc__D_e* (glucose), *EX_coa_e* (coenzyme A) and *EX_starch1200_e* (amidon) sont mis à 0. Pour *L. plantarum*, les bornes inférieure de *EX_glc__D_e* (glucose), *EX_gal__e* (galactose), *EX_coa_e* (coenzyme A) and *EX_dha_e* (dihydroxyacetone) sont mis à 0. Enfin, pour *P. freudenreichii*, les bornes inférieure de *EX_starch1200_e* (amidon) and *EX_dha_e* (dihydroxyacetone) ont été mise à 0. Ces composés n'apparaissent pas dans la composition du lait ce qui justifie ces contraintes.

Dans un second temps, nous avons établis des contraintes de modélisation dynamiques en se basant sur la connaissance biologique de l'espèce. Expérimentalement, nous avons observé que le métabolisme des bactéries lactiques est toujours actif pendant la phase stationnaire, que l'acide lactique est produit par le lactose et que ce dernier n'est pas considéré comme un substrat limitant dans nos conditions expérimentales. Nous avons ainsi régulé le flux de production de l'acétate et de l'acide lactique par le flux de consommation du lactose, source de carbone principale.

[ML – a retravailler la présentation, enlever sous forme d’item]

L. lactis et *L. plantarum*

— EX_lac__L_e, EX_lac__D_e:

- inférieur: 0

- supérieur:
$$\min\left(\frac{m_{lcts_e}}{\Delta t * \sum_{i \in \mathcal{M}(lcts_e)} b_i}, \mu_{max,lcts} * 10^{(-k_{lac} * \phi_{undiss})} + \mu_{min,lcts}\right) * 4 \quad (1.1)$$

Justification: Comme la production de lactate est liée à la disponibilité du lactose, une modulation dynamique pour l’export de lactate imitant l’import de lactose (voir eq. (1.3)), avec un coefficient stoechiométrique de 4.

— EX_ac_e:

- inférieur: 0

- supérieur:
$$\max\left(-\frac{m_{lcts_e}}{\Delta t * \sum_{i \in \mathcal{M}(lcts_e)} b_i}, v_{i,ac_e}^{exp}\right)$$

Justification: La production d’acetate est régulée par la disponibilité de lactose quand celui ci est épuisé, et est exporté autrement selon la limite d’export physiologique intrinsèque $v_{i,ac_e}^{exp} = 0.5$ pour *L. lactis* et $v_{i,ac_e}^{exp} = 1.0$ pour *L. plantarum*.

Lorsque les concentrations extra-cellulaires de lactose sont suffisamment importante, le flux de consommation du lactose par les bactéries lactiques est également régulé par l’acidité libéré par ces bactéries. En effet, le métabolisme des bactéries lactiques continue même lorsqu’elles sont a l’état stationnaire. En revanche, lorsque la concentration de lactose devient insuffisante, un partage uniforme entre les bactéries consommatrices, composés des deux bactéries lactiques et de la bactérie propionique est mise en place. L’approche modélisant ce processus est inspiré des travaux de (?) où l’acidité libéré par les bactéries lactiques est représentée par le forme non dissocié de l’acide lactique (voir matériel supplémentaire §??).

L. lactis

— EX_lcts_e:

- inférieur:
$$\max\left(-\frac{m_{lcts_e}}{\Delta t * \sum_{i \in \mathcal{M}(lcts_e)} b_i}, -\mu_{max,lcts} * 10^{(-k_{lac} * \phi_{undiss})} - \mu_{min,lcts}\right) \quad (1.2)$$

- supérieur: 1000

Justification: Quand le lactose est épuisé, le premier terme est activé, modélisant le partage uniforme du lactose disponible parmi l'ensemble de bactérie $\mathcal{M}(lac_L)$ métabolisant le lactose. Autrement, une régulation par le pH est ajoutée : la quantité de molécules d'acide lactique non dissocié est calculé à partir de la concentration de lactate au moyen de la fonction ϕ_{undiss} (voir (1.11) pour l'expression de cette fonction) ; puis, avec une croissance exponentielle à partir de la valeur $-\mu_{max,lcts}^{lact} - \mu_{min,lcts}^{lact}$ à la valeur $-\mu_{min,lcts}^{lact}$ quand la concentration d'acide lactique non dissocié augmente, avec un taux exponentiel k_{lac} . Les valeurs de $\mu_{max,lcts}^{lact}$, $\mu_{min,lcts}^{lact}$ et k_{lac} sont inférées des expériences de co-cultures.

L. plantarum

— EX_lcts_e:

- inférieur:
$$\max\left(-\frac{m_{lcts_e}}{\Delta t * \sum_{i \in \mathcal{M}(lcts_e)} b_i}, -\mu_{max,lcts} * 10^{(-k_{lac} * \phi_{undiss})} - \mu_{min,lcts}\right) \quad (1.3)$$

- supérieur: 1000

Justification: Quand le lactose est épuisé, le premier terme est activé, modélisant le partage uniforme du lactose disponible parmi l'ensemble de bactérie $\mathcal{M}(lac_L)$ métabolisant le lactose. Autrement, une régulation par le pH est ajoutée : la quantité de molécules d'acide lactique non dissocié est calculé à partir de la concentration de lactate au moyen de la fonction ϕ_{undiss} (voir (1.11) pour l'expression de cette fonction) ; puis, avec une croissance exponentielle à partir de la valeur $-\mu_{max,lcts}^{plant} - \mu_{min,lcts}^{plant}$ à la valeur $-\mu_{min,lcts}^{plant}$ quand la concentration d'acide lactique non dissocié augmente, avec un taux exponentiel k_{lac} . Les valeurs de $\mu_{max,lcts}^{plant}$, $\mu_{min,lcts}^{plant}$ et k_{lac} sont inférées des expériences de co-cultures.

P. freudenreichii

— EX_lcts_e:

- inférieur:
$$\max\left(-\frac{m_{lcts_e}}{\Delta t * \sum_{i \in \mathcal{M}(lcts_e)} b_i}, F_{lcts}\right)$$

- supérieur: 1000

Justification: Quand le lactose est épuisé, le premier terme est activé, modélisant le partage uniforme parmi l'ensemble des bactéries $\mathcal{M}(lcts_e)$, métabolisant le lactose. Autrement, un flux constant F_{lcts} estimé des données de métabolomiques en culture pure (voir la section ??) est appliqué.

Nous avons utilisé jusqu'ici uniquement la connaissance apporté par la littérature afin de contraindre chaque modèle individuel en y associant une équation de régulation. Dans les paragraphes suivant, une communication entre les données et les modèle sera utilisé

afin d'ajouter des contraintes et pouvoir expliquer les données observées pour chaque souche.

Pour rappel, le lactate, l'acétate, le succinate et le propionate sont dosés pour *P. freudenreichii* et les concentrations initiales et finales sont récupérées. Ainsi, les bornes des flux de production de ces composés sont estimées des données dans le but de rapprocher la valeur de flux de production prédit par le modèle avec celle des données. Concernant le lactate, ce dernier étant consommé, une répartition uniforme de la quantité de lactate disponible dans le milieu a été faite puisque les nutriments ne sont pas considérés comme limitant ici. Enfin, une croissance logistique a été supposé pour modéliser la phase plateau pour les mêmes raisons.

P. freudenreichii

— EX_lac__L_e, EX_lac__D_e:

- inférieur:
$$\max\left(-\frac{m_{lacL,e}}{\Delta t * \sum_{i \in \mathcal{M}(lacL)} b_i}, F_{lac} * K_{freud} * \frac{b_{freud}}{B_{p,freud}}\right) \quad (1.4)$$

- supérieur: 1000

Justification: Quand le lactate est épuisé, le premier terme est activé, modélisant le partage uniforme parmi l'ensemble des bactéries $\mathcal{M}(lacL)$, métabolisant L_lactate (resp. D_lactate). Autrement, un flux F_{lac} estimé des données de métabolomiques en culture pure (voir la section ??) est appliqué après une régulation par la fraction $\frac{b_{freud}}{B_{p,freud}}$, où $B_{p,freud}$ est la valeur plateau de la concentration de la biomasse en pure culture, et b_{freud} et la densité de biomasse actuelle, et le facteur K_{freud} , qui est inféré des données de croissance en culture pure. Cette modulation modélise un ordre de grandeur différent du rendement du métabolisme en culture pure ou en co-culture.

— EX_ac_e, EX_ppa_e, EX_succ_e:

- inférieure: 0

- supérieur:
$$F_i * K_{freud} * \frac{b_{freud}}{B_{p,freud}}, \quad \text{pour } i \in \{ \text{EX_ac_e}, \text{EX_ppa_e}, \text{EX_succ_e} \}$$

Justification: Le taux maximal de production est définie par F_i , pour $i \in \{ \text{EX_ac_e}, \text{EX_ppa_e}, \text{EX_succ_e} \}$, qui est estimé des données de métabolomique de culture pure (voir la section ??). Le même terme de modulation que 1.4 est appliqué, modélisant un un ordre de grandeur différent du rendement du métabolisme en culture pure ou en co-culture.

Toutes ces modifications faites, soit à partir de la connaissance biologique soit à partir

des données biochimiques, rendent la prédiction des modèles plus précise et plus fidèle aux données. Cependant, un processus itératif dynamique d’ajustement de ces valeurs est nécessaire pour permettre une prédiction numérique : l’étape de calibration des modèles.

L’étape de calibration des modèles consiste à estimer des paramètres des modèles dynamiques afin de minimiser l’écart entre les valeurs des modèles et celles des données. Cela revient à résoudre le problème de minimisation suivant pour les bactéries lactiques (équations (1.12)) et pour *P. freudenreichii* (équation (1.13)). Brièvement, nous cherchons à minimiser l’écart aux données de croissances, pour les bactéries lactiques comme pour *P. freudenreichii*, l’écart aux données de pH pour les deux bactéries lactiques et l’écart aux données de métabolites dosés pour *P. freudenreichii*. Avec ce type de méthode, le sur apprentissage, consistant à sur-prédire les valeurs, est une dimension à prendre en compte. Afin de l’éviter, nous avons optimisé un faible nombre de paramètre pour chaque espèce : seulement deux paramètres ont été utilisé pour calibrer la croissance de *L. lactis* et la régulation du pH, trois pour *L. plantarum*, qui sont impliqués dans la régulation du pH ainsi que pour la croissance, et enfin, un paramètre pour *P. freudenreichii*, ajustant sa croissance (voir la section matériels et méthode, §??, pour de plus amples détails).

Cette étape itérative de calibration dynamique est coûteuse en temps. En effet, pour chacune des nouvelles valeurs de paramètres, nous avons lancés une simulation dynamique de chaque souche optimisés. En fonction de la courbe simulé et de la connaissance *a priori* du métabolisme de chaque souche, nous avons raffiné certains processus métaboliques, en ajustant dynamiquement les bornes d’import et d’export des métabolites d’intérêt pour arriver aux ajustements montré ci-dessus. Pour tous les autres métabolites suivi dynamiquement, c’est à dire le butanediol, le diacétyle et le citrate, la borne inférieure est modulée selon la disponibilité de ce substrat selon le principe de partage équitable décrit plus haut lorsque la quantité de substrat devient limitante :

L. lactis

— EX_diact_e, EX_btd_RR_e, EX_cit_e:

$$\max\left(-\frac{m_j}{\Delta t * \sum_{i \in \mathcal{M}(j)} b_i}, v_{i,j}^{int}\right)$$

- inférieur: pour $i = L. lactis$ et $j \in \{EX_diact_e, EX_btd_RR_e, EX_cit_e, EX_ac_e\}$

- supérieur: 1000

Justification: Limitation dynamique de consommation usuelle (voir eq. (1.10)) Quand le substrat est pas limité, l'import is contraint par la limite physiologique intrinsèque modélisé par $v_{i,j}^{int} = -8$ pour $j = L. lactis$ et $j \in \{EX_diact_e, EX_btd_RR_e\}$ et $v_{i,j}^{int} = -1$ pour $j = EX_cit_e$, et autrement, par la disponibilité du nutriment. La disponibilité du substrat est uniformement partagé par les micro-organismes consommateurs.

L. plantarum

— EX_btd_RR_e:

$$\max\left(-\frac{m_j}{\Delta t * \sum_{i \in \mathcal{M}(j)} b_i}, v_{i,j}^{int}\right),$$

- inférieur: pour $j = L. plantarum$ et $j \in \{EX_btd_RR_e\}$

- supérieur: 1000

Justification:

Limitation dynamique de consommation usuelle (voir eq. (1.10)) Quand le substrat est pas limité, l'import est contraint par la limite physiologique intrinsèque modélisé par $v_{i,j}^{int} = -5$ pour $i = L. plantarum$ et $j \in \{EX_ac_e, EX_btd_RR_e\}$ et autrement, par la disponibilité du nutriment. La disponibilité du substrat est uniformement partagé par les micro-organismes consommateurs.

En plus, de ces régulations régulations dynamiques, nous avons observé, via la méthode de FVA, que le flux de consommation du citrate, de la serine, du glutamate et de l'alanine étaient extrêmement fort empêchant la prédiction des composés dosés en culture pure. Ainsi, pour réorienter les flux dans les bonnes voies métaboliques nous avons réduits chacune des bornes d'importation de ces composés à -12, améliorant ainsi les prédictions.

Nous avons ainsi curé manuellement et calibré dynamiquement chacun de nos modèles FBA individuels dans le but premier de retrouver les valeurs de concentration des expériences biologiques. Chacune de nos simulations statiques et dynamiques de nos modèles calibrés sont représentés avec les figures 1.2, 1.3 et 1.4. Chaque sous figure (a) représente

les flux des voies métaboliques d'intérêt calculés par le FBA lors de l'inoculation où aucun métabolite est limitant.

Le modèle FBA de *L. lactis* nous montre une préférence d'utilisation de la voie métabolique du tagatose plutôt que de Leloir pour dégrader le lactose. Nous remarquons également que les deux isomères de l'acide lactique sont produits *via* la voie de la glycolyse avec un plus grand flux de production pour le D-lactate, expliquant ainsi l'acidification du lait et de la caillé durant le processus de fabrication de fromage.(Fig. 1.2a). Concernant la voie de l'acétolactate, nous remarquons qu'elle est bien active grâce à la production du diacétyl et du butanediol par cette voie. Enfin, nous observons également un flux de consommation du citrate, *via* la réaction *CITL* encodé par le gène *citL*.

Contrairement à *L. lactis*, le modèle métabolique de *L. plantarum* semble avoir une préférence de production pour le L-acide lactique, montrant ainsi sa contribution à l'acidification du fromage. En revanche, la voie de la glycolyse est bien active de la consommation du lactose jusqu'à la production des deux formes du lactate. Nous remarquons également que un flux de production de l'acide acétique par la voie de l'acétatekinase, contribuant plus légèrement à l'acidification (Fig. 1.1b). Au regard de la simulation dynamique de *L. plantarum*, nous avons remarqué une activation de la voie métabolique hétérolactique et homolactique. Cependant à notre connaissance, la voie métabolique de la transkéolase, représentant la fermentation hétérolactique, est observé pour la première fois dans un contexte laitier. En revanche, la consommation du citrate est souche dépendante pour *L. plantarum* (?) et a été observé dans notre cas à travers la réaction de transport *FEDCabc*, conduisant à une augmentation de la quantité de pyruvate intracellulaire.

Enfin, le modèle FBA de *P. freudenreichii* consomme préférentiellement du lactose malgré la présence de l'acide lactique (Fig. 1.4a). Cependant, nous retrouvons une activation simultanée des cycles du TCA et de Wood-Werkman (?) où le premier permet la régénération des sources de carbones nécessaires pour l'activation du cycle de Wood-Werkman et ainsi, la production de propionate.

Les sous figures (b) montrent que les courbes de croissances simulées (en vert) sont fidèles aux données de croissances expérimentales de chaque souche microbienne. Concernant le pH, pour les sous figure 1.2, 1.3 (c), nous avons observé que l'utilisation du lactate comme proxy du pH permet d'expliquer le pH observé lors des cultures pures de *L. lactis* et *L. plantarum*. Enfin, notre modèle métabolique de *P. freudenreichii* retrouve les concentrations finales prédits par les expériences biologiques 1.4(c). Afin de vérifier statistiquement que l'apprentissage de nos modèles est bon, nous avons représenté sur la figure Fig. 1.5, fenêtre de gauche, les valeurs de croissance, de pH et de métabolites calculé par notre modèle et celles des données expérimentales. Nous observons une très bonne corrélation de 0.99, montrant ainsi que nos modèles métaboliques individuels sont suffisamment bien ajustés.

1.2.3 Le suivie de la production des métabolites dans les conditions de co-cultures par le modèle de communauté

Nous venons de voir tout le processus de curation et de calibration des modèles métaboliques individuels dans le but de prédire les concentrations et les croissances de co-culture ainsi que d'émettre de nouvelles interactions métaboliques au sein de ce consortium. Dans cette section, les données de co-cultures qui nous servirons de données tests à nos prédictions seront d'abord présentées brièvement, puis nous nous concentrons sur comment le modèle communautaire fut créé et son exploitation.

Un suivi de la croissance des bactéries ainsi que de la production des métabolites ont été réalisé durant le fabrication du fromage. Chaque souche de ce consortium que sont *L. lactis*, *L. plantarum* et *P. freudenreichii* ont atteint respectivement un seuil de cultivabilité de $8.45 \log_{10}\text{CFU/g}$, $8.47 \log_{10}\text{CFU/g}$ et $8.59 \log_{10}\text{CFU/g}$ en approximativement 1200 heures.

Pour vérifier si nous prédisons ces données, nous avons tout d'abord construit un modèle de communauté regroupant les modèles individuels calibrés où chaque individu maximise sa propre biomasse. Par la suite, nous avons ajusté la valeur du "quorum sen-

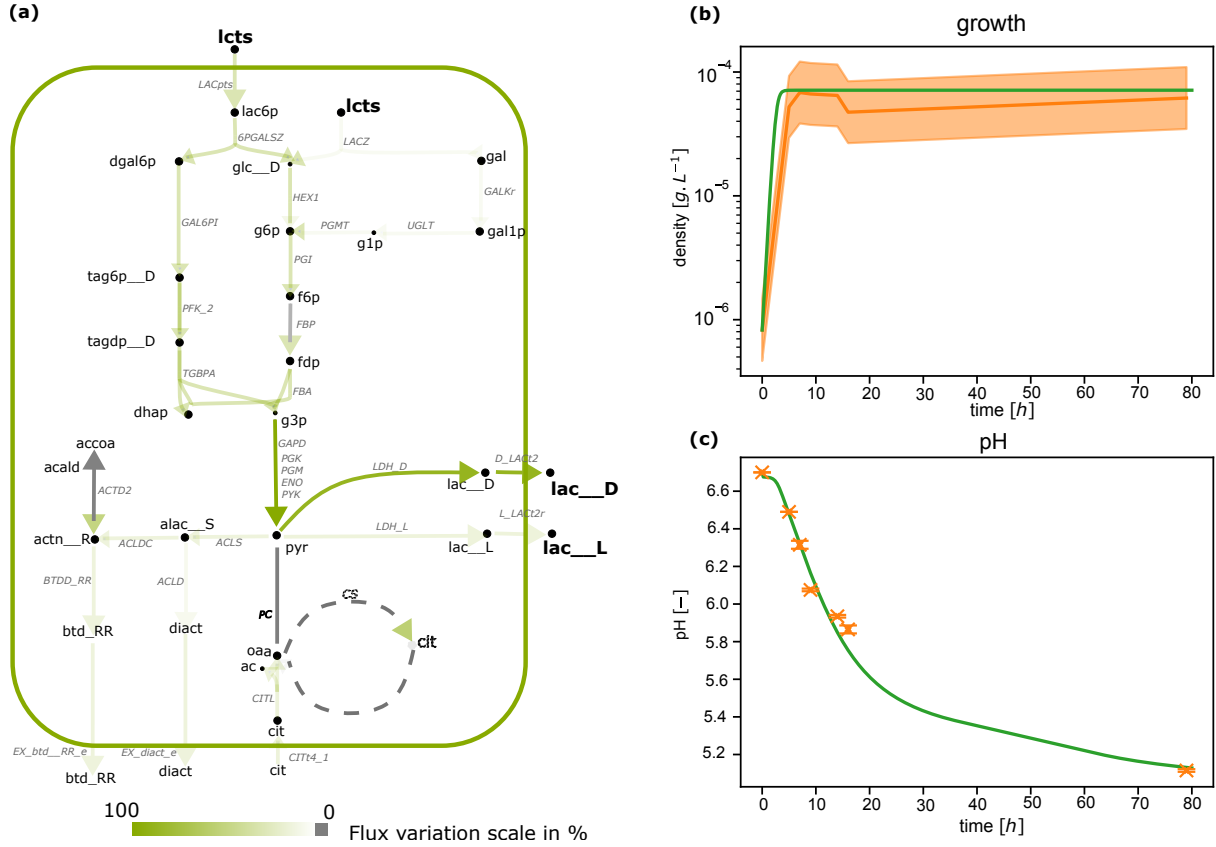


FIGURE 1.2 – *L. lactis* GEM fitted on pure culture data. (a) FBA optimization with optimal parameters applied to the central carbon metabolism at the inoculation step. The color scale represents the reaction flux values predicted by the FBA and normalized by the highest flux value of the illustrated pathways. (b) Dynamics of *L. lactis* population in pure culture after parameter inference in the model (green line) and in the experimental data (orange line, with $\pm 1/4$ log, which is the admitted range of precision for plating numbering). (c) pH in pure-culture of *L. lactis* in the model (green line) and in the experimental data (orange crosses, with standard deviation). *Compound abbreviations* : ac : acetate; actn__R : acetoin; alac__S : (S)-2-Acetolactate; btd__RR : butanediol; cit : citrate; dgal6p : d-galactose-6-phosphate; dhap : Dihydroxyacetone phosphate; diact : diacetyl; f6p : fructose-6-phosphate; fdp : D-Fructose 1,6-bisphosphate; g1p : glucose-1-phosphate; g3p : Glyceraldehyde 3-phosphate; g6p : glucose-6-phosphate; gal : galactose; gal1p : galactose-1-phosphate; glc__D : glucose; lac__D, lac__L : lactate; lac6p : lactose-6-phosphate; lcts : lactose; oaa : Oxaloacetate; pyr : pyruvate; tag6p__D : D-tagatose-6-phosphate; tagdp__D : D-Tagatose 1,6-bisphosphate. *Reaction abbreviations* : 6PGALSZ : 6-phospho-beta-galactosidase; ACLD : Acetolactate decarboxylase; ACTD2 : Acetoin dehydrogenase; ACLDC : Acetolactate decarboxylase; ACLS : Acetolactate synthase; BTDD__RR : R R butanediol dehydrogenase; CITL : Citrate lyase; CITt4__1 : Citrate transport via sodium symport; CS : Citrate synthase; D__LACT2 : D lactate transport via proton symport; ENO : Enolase; FBA : Fructose-bisphosphate aldolase; GAL6PI : Galactose-6-phosphate isomerase; GALkr : Galactokinase; GAPD : Glyceraldehyde-3-phosphate dehydrogenase; HEX1 : Hexokinase (D-glucose :ATP); L__LACT2r : L lactate reversible transport via proton symport; LACpts : Lactose transport via PEP :Pyr PTS; LACZ : B-galactosidase; LDH__D : D-lactate dehydrogenase; LDH__L : L-lactate dehydrogenase; PC : Pyruvate carboxylase; PFK : Phosphofructokinase; PFK__2 : Phosphofructokinase; PGI : Glucose-6-phosphate isomerase; PGK : Phosphoglycerate kinase; PGM : Phosphoglycerate mutase; PYK : Pyruvate kinase; TGBPA : Tagatose-bisphosphate aldolase; UGLT : UDPglucose–hexose-1-phosphate uridylyltransferase

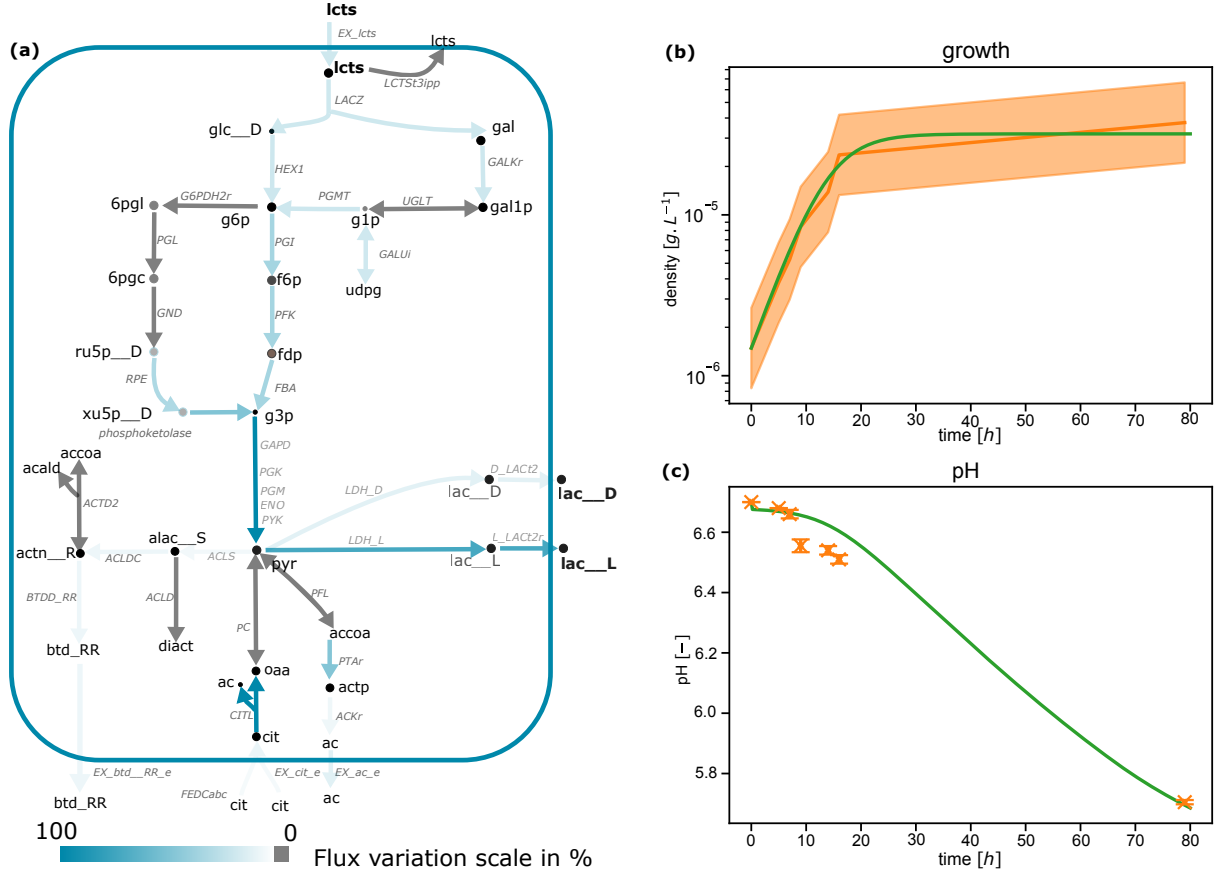


FIGURE 1.3 – *L. plantarum* GEM fitted on pure culture data. (a) FBA optimization with optimal parameters applied to the central carbon metabolism at the inoculation step. The color scale represents the reaction flux values predicted by the FBA and normalized by the highest flux value of the illustrated pathways. (b) Dynamics of *L. plantarum* population in pure culture after parameter inference in the model (green line) and in the experimental data (orange line, with $\pm 1/4$ log, which is the admitted range of precision for plating numbering). (c) pH in pure-culture of *L. plantarum* in the model (green line) and in the experimental data (orange crosses, with standard deviation). 6pgc : 6-Phospho-D-gluconate; 6pgl : 6-phospho-D-glucono-1,5-lactone; ac : acetate; acald : acetaldehyde; accoa : acetyl-coa; actn__R : acetoin; alac__S : (S)-2-Acetolactate; btd__RR : butanediol; cit : citrate; dhap : Dihydroxyacetone phosphate; f6p : fructose-6-phosphate; fdp : D-Fructose 1,6-bisphosphate; g1p : glucose-1-phosphate; g3p : Glyceraldehyde 3-phosphate; g6p : glucose-6-phosphate; gal : galactose; gal1p : galactose-1-phosphate; glc__D : glucose; lac__D, lac__L : lactate; lcts : lactose; oaa : Oxaloacetate; pyr : pyruvate; ru5p__D : D-Ribulose 5-phosphate; tag6p__D : D-tagatose-6-phosphate; tagdp__D : D-Tagatose 1,6-bisphosphate; udpg : UDPglucose; xu5p__D : D-Xylulose 5-phosphate. *Reaction abbreviations* : ACLD : Acetolactate decarboxylase; ACTD2 : Acetoin dehydrogenase; ACLDC : Acetolactate decarboxylase; ACLS : Acetolactate synthase; BTDD__RR : R R butanediol dehydrogenase; D_LACt2 : D lactate transport via proton symport; ENO : Enolase; FBA : Fructose-bisphosphate aldolase; FEDCabc : FEDCabc; G6PDH2r : Glucose 6-phosphate dehydrogenase; GALKr : Galactokinase; GALUi : UTP-glucose-1-phosphate uridylyltransferase (irreversible); GAPD : Glyceraldehyde-3-phosphate dehydrogenase; GND : Phosphogluconate dehydrogenase; HEX1 : Hexokinase (D-glucose :ATP); L_LACT2r : L lactate reversible transport via proton symport; LACZ : B-galactosidase; LCTSt3ipp : Lactose transport via proton anionport (periplasm); LDH_D : D-lactate dehydrogenase; LDH_L : L-lactate dehydrogenase; PC : Pyruvate carboxylase; PFL : Pyruvate formate lyase; PTA : Phosphotransacetylase; PGL : 6-phosphogluconolactonase; PFK : Phosphofructokinase; PGK : Phosphoglycerate kinase; PGI : Glucose-6-phosphate isomerase; PGM : Phosphoglycerate mutase; PGMT : Phosphoglucomutase; phosphoketolase : phosphoketolase; PYK : Pyruvate kinase; RPE : Ribulose 5-phosphate 3-epimerase; UGLT : UDPglucose-hexose-1-phosphate uridylyltransferase.

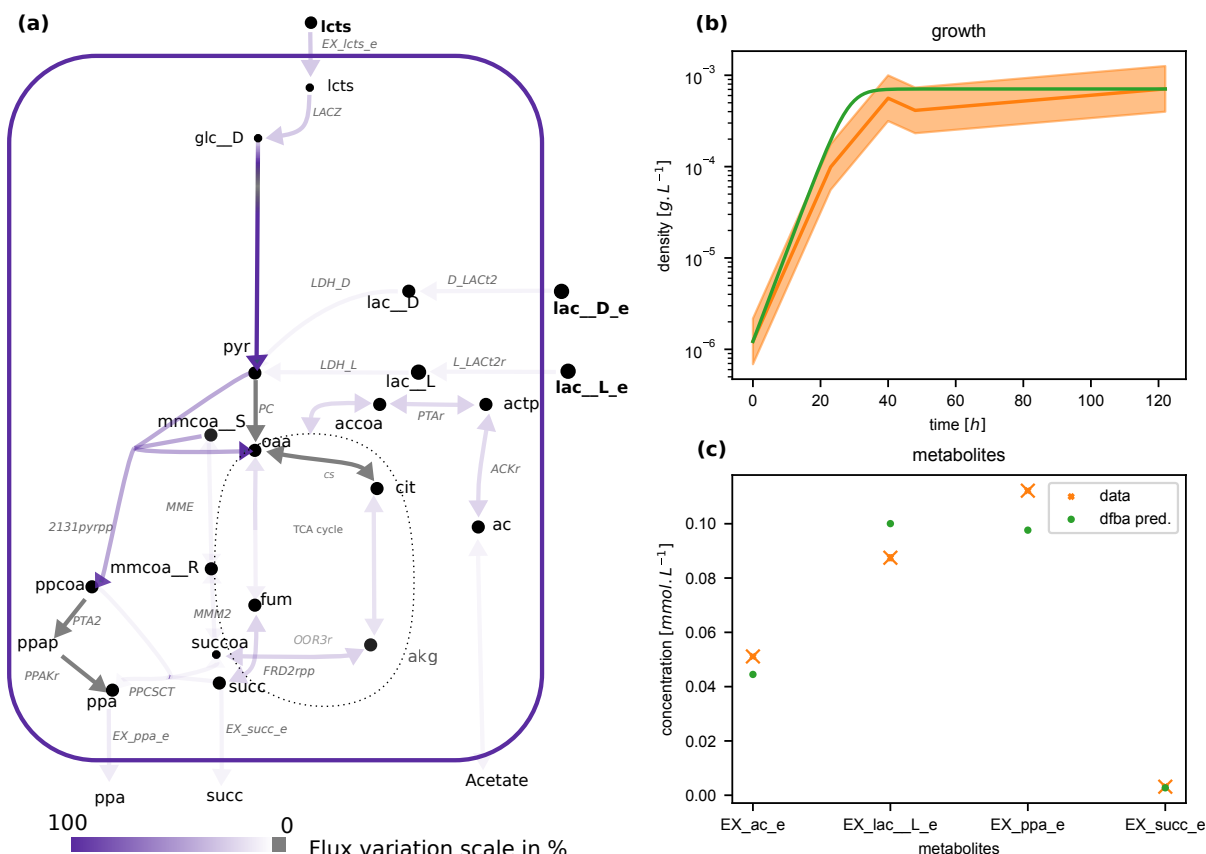


FIGURE 1.4 – *P. freudenreichii* GEM fitted on pure culture data. (a) FBA optimization with optimal parameters applied to the central carbon metabolism at the inoculation step. The color scale represents the reaction flux values predicted by the FBA and normalized by the highest flux value of the illustrated pathways. (b) Dynamics of *P. freudenreichii* population size in pure culture after parameter inference in the model (green line) and in the experimental data (orange line, with $\pm 1/4$ log, which is the admitted range of precision for plating numbering). (c) Lactate, acetate, propionate and succinate concentrations simulated (green points) and observed (orange crosses, with standard deviation) at the final time point of a second experiment with different lactate inoculation (see Material and method) *Compound abbreviations* : ac : acetate; akg : 2-Oxoglutarate; fum : fumarate; glc__D : glucose; lac__D, lac__L : lactate; lcts : lactose; mmcoa__R : (R)-Methylmalonyl-CoA; mmcoa__S : (S)-Methylmalonyl-CoA; oaa : Oxaloacetate; ppa : propionate.; ppap : Propanoyl phosphate; ppcoa : propionyl-coa; pyr : pyruvate; succ : succinate; succoa : succinyl-coa. *Reaction abbreviations* : 2131pyrpp : Methylmalonyl-CoA carboxyltransferase 5S subunit; CS : Citrate synthase; D_LACT2 : D lactate transport via proton symport; FRD2rpp : Fumarate reductase / succinate dehydrogenase (irreversible) (periplasmic, membrane potential dissipating); L_LACT2r : L lactate reversible transport via proton symport; LACZ : B-galactosidase; LDH_D : D-lactate dehydrogenase; LDH_L : L-lactate dehydrogenase; MME : Methylmalonyl-CoA epimerase; MMM2 : Methylmalonyl-CoA mutase; OOR3r : 2-oxoglutarate synthase (rev); PC : Pyruvate carboxylase; PPAKr : Propionate kinase; PPCSCT : Propanoyl-CoA : succinate CoA-transferase; PTA2 : Phosphate acetyltransferase.

sing", c'est à dire la valeur à laquelle nous observons une phase plateau, de chaque modèle individuel en intégrant la valeur des données expérimentales au sein de l'équation 1.9 dans le but de reproduire le comportement métabolique de la communauté *in silico* (voir eq. (1.7)-(1.8) et matériels et méthodes §??). Enfin, notons que, comme aucun ajustement du modèle de communauté n'a été fait, l'expérience de co-culture peuvent être considéré comme des données tests impliquant de nouveaux points de validations.

Les prédictions à l'échelle de la communauté se font sur trois niveaux : croissance des bactéries, pH des bactéries lactiques, et métabolomique. Nous remarquons que le modèle de communauté prédit correctement (Fig. 1.7(a)-(c), Table supplémentaire ??). Nous remarquons cependant une légère surestimation de la croissance de *P. freudenreichii* durant la phase exponentielle. Concernant la prédiction des métabolites issue de la métabolomique, nous observons une consommation plus lente du lactose tandis que celle du citrate s'ajuste bien avec les données expérimentales. La production de lactate est légèrement surprédit durant la phase exponentielle résultant d'une légère suradification du fromage durant les étapes de moulage et de démoulage. En revanche, la production d'acétate est sous prédit durant les premières phases de la fabrication du fromage, mais le taux de production pendant la maturation est correctement modélisé. Les courbes du propionate, et dans une moindre mesure, celle du succinate sont prédit avec précision par le modèle suggérant que la production de composés organoleptiques par *P. freudenreichii* est correctement capturée. De plus, les voies métaboliques activées en mono-cultures le sont également en co-culture (Figure 1.6). Enfin, afin de valider statistiquement les prédictions du modèle communauté, nous avons fait une régression entre les valeurs prédites par le modèle et celles des données, et nous observons un R^2 globale de 0.98 pour cette étape de validation (Figure 1.5, fenêtre de droite). Cela démontre que le temps passé sur la curation et le raffinement des modèles individuels, en combinant méthodes numériques et la connaissance biologique *a priori* permet d'obtenir des prédictions à l'échelle d'une communauté avec une bonne précision.

[ML – TODO : traduire les figure + retravailler paragraphe interactions]

img/tango/qqplot.pdf

FIGURE 1.5 – **Model goodness of fit.** All the model outputs (\hat{y}) are plotted versus their respective data value (y) in the training (i.e. the mono-culture data, left panel) and in the testing set (i.e. the co-culture data, right panel). Each point is colored according to the data type and the bissector is plotted (red dotted line). A global coefficient of determination is computed with the formula $R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$, where \bar{y} is the average value of y . We can observe a very good agreement of the model in the training set, and a globally good agreement in the testing set, with 3 outliers on lactose and citrate.

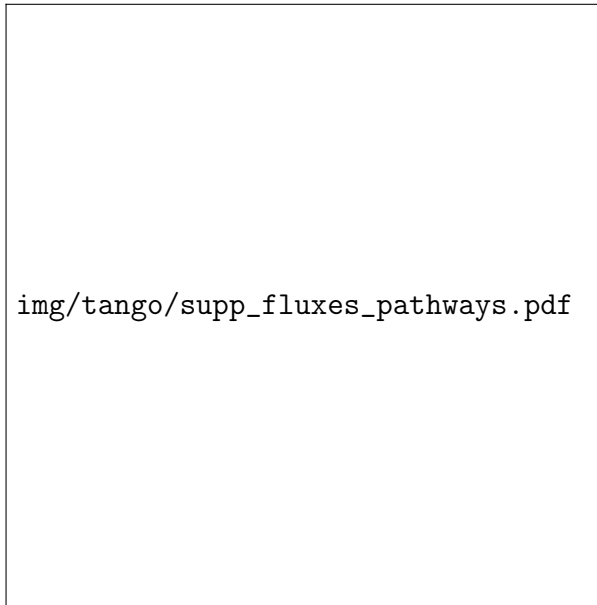


FIGURE 1.6 – **Metabolic pathway switches during cheese production.** For each bacterium we ran a dFBA simulation and reaction fluxes per each pathways of interest in Figure 1.1 were retrieved at the inoculum of LAB (Linoc, t=0 hours), the inoculum of *P. freudenreichii* (Finoc, t=18 hours the molding (Cm, t=19.5 hours), demolding (Cdm, t=40hours), start of ripening (C0w, t=60 hours), fourth week ripening (C4w, t=732 hours) and seventh week ripening state (C7w, t=1236 hours). We normalized fluxes by the inoculum flux value and the average flux for each pathways is therefore represented.

1.2.4 Les interactions bactériennes dynamiques au sein de la communauté

Nous avons montré jusqu'à présent qu'associer la connaissance biologique, issu de la littérature et/ou d'un ou une biologiste, avec un modèle numérique est essentielle pour comprendre ce consortium bactérien du point du vue du métabolisme. De plus, nous avons montré que même un réseau métabolique de bonne qualité initialement nécessite une curation supplémentaire rigoureuse. En effet, l'utilisation de modèle suffisamment bien curé et raffiné a permis de reproduire les expériences biologique lors des simulations en communauté avec une bonne confiance dans les résultats. Grâce à cette assurance, nous pouvons exploiter d'avantage le modèle de communauté du point de vue des interactions bactériennes possibles au sein de cette communauté. Dans les paragraphes suivant, nous déduirons dans un premier temps des interactions que notre approche numérique a pu mettre en évidence en analysant les flux relatifs et absolu de production et de consommation de chaque métabolite pour chaque espèce. Dans un second temps, nous verrons ce que des outils sans connaissance *a priori* peuvent découvrir.

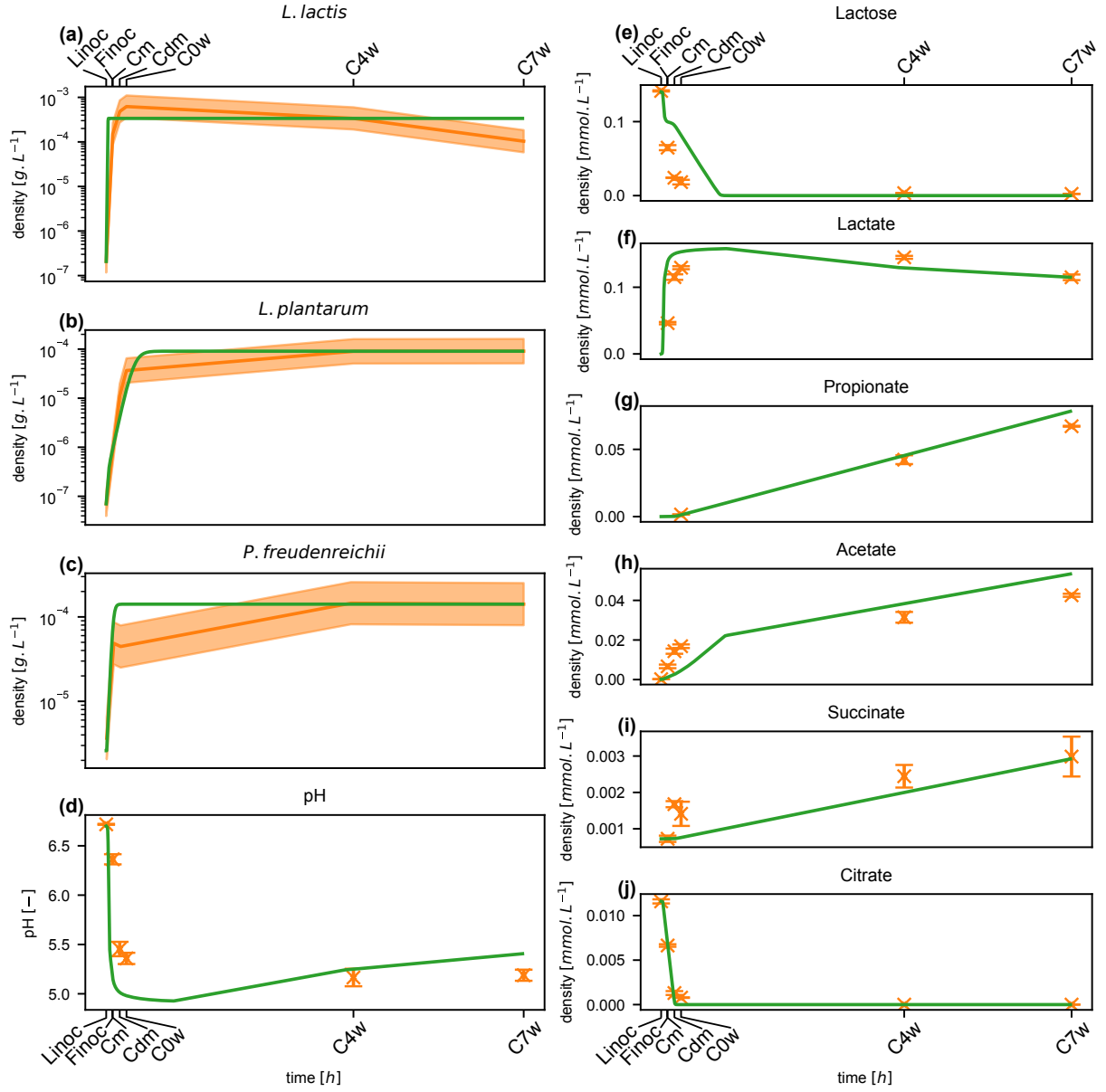


FIGURE 1.7 – (a-c) Simulated growth of respectively *L. lactis*, *P. freudenreichii* and *L. plantarum* species computed by the community model (green lines) and experimentally-observed (orange lines, confidence interval of $\pm 1/4$ log). (d-j) Computational and experimental co-culture pH and metabolic profiles. Each data point is represented with its corresponding standard deviation (error bars). Abbreviations : Linoc, Lactic acid bacteria inoculation ($t = 0$); Finoc, *P. freudenreichii* inoculation ($t = 18$ hours); Cm, molding ($t = 19.5$ hours); Cdm, demolding ($t = 40$ hours); C0w, start of ripening ($t = 60$ hours); C4w, Fourth week ripening ($t = 732$ hours); C7w, seventh week ripening ($t = 1236$ hours).

Notre modèle numérique de communauté peut révéler des interactions microbiennes telles que l'alimentation croisée ou la compétition pour un nutriment. Dans ce but, nous nous sommes focalisés sur la dynamique des flux des métabolites d'intérêts pour chaque espèce (Fig. 1.9, (a)) en calculant, à chaque pas de temps de la dynamique, et à l'échelle de la population, les flux d'échange $\mu_{i,j}b_i$ pour chaque métabolite j et bactérie i et les concentrations des métabolites prédites associées. Concernant le lactose et le citrate, nous observons que *L. lactis* les consomme très rapidement durant leur phase exponentielle avant d'être contrôlé par la régulation du pH pour le lactose et par la disponibilité du citrate. Nous pouvons observer que *L. plantarum* ne contribue pas ou peu dans la consommation totale de ces métabolites suggérant ainsi que d'autres métabolites non suivie en dynamique sont utilisés pour la croissance de ce dernier, comme par exemple, les acides aminés. Une hypothèse pouvant expliquer ce phénomène peut être due à la croissance plus lente de *L. plantarum* dans les premières heures du processus de production du fromage. En effet, on remarque que les processus de régulations sur la consommation du lactose et du citrate sont mis en jeu lorsque la densité bactérienne de *L. plantarum* est encore faible. De manière surprenante, nous observons que *P. freudenreichii* termine la déplétion du lactose à la place de *L. plantarum* au tout début de la phase d'affinage, malgré que le lactose soit déjà produit.

En effet, la production de lactate étant corrélée avec la consommation de lactose, nous observons une très forte production conduite par *L. lactis* pendant deux cent cinquante heures, correspondant à l'affinage dans le processus de fabrication du fromage. En dépit d'une plus faible croissance de *L. plantarum*, ce dernier contribue faiblement à la production de lactate également. Durant la phase de l'affinage, le lactate est consommé par *P. freudenreichii* contribuant ainsi à l'augmentation du pH dans la communauté comme observé dans la Figure 1.7 d. Originellement, nous remarquons que les trois souches bactériennes sont capables de produire de l'acétate dans le milieu extracellulaire où *L. lactis* est encore le contributeur majeur suivi de *L. plantarum* jusqu'au temps $t =$ deux cents cinquantes heures, temps à laquelle leur phase plateau commence. Ce schéma est retrouvé chez les bactéries lactique intervenant également dans la production du butanediol via la voie de l'acétolactate. Cependant, *P. freudenreichii* continue de garder un taux de produc-

tion constant d'acétate durant la maturation du fromage malgré la phase plateau atteinte. Enfin, le propionate et le succinate est produit à taux constant par *P. freudenreichii* durant tout le processus d'affinage et seul la souche de *L. lactis* est capable de produire le diacétyle jusqu'au debut de la période de l'affinage.

Afin d'obtenir la contribution en flux net de chaque souche au sein de ce consortium, nous avons intégré dans le temps la prédiction du FBA $\int_0^T \mu_{i,j}(t)b_i(t)dt$. Après normalisation de cet échange net par les échanges totaux au sein de la communauté (i.e. la somme des échanges net de chaque individu), nous avons obtenu les contributions de chaque souche pour chaque métabolite (Fig. 1.9, (b)). Ces contributions confirment que le succinate et le propionate sont uniquement produits par *P. freudenreichii*, alors que le diacétyle (production) et le citrate (consommation) sont métabolisé par *L. lactis*. *L. plantarum* contribue à la production du butanediol, bien que le principale producteur est prédit comme pour être *L. lactis*. Dans une moindre mesure, *L. lactis* contribue plus que *L. plantarum* à la production de l'acétate où cette dernière est dominée par *P. freudenreichii*. En terme d'interaction bactérienne, on observe une alimentation coisée pour le lactate entre *L. lactis* (producteur) et *P. freudenreichii* (consommateur). La consommation du lactose est partagé entre *L. lactis* et *P. freudenreichii* et curieusement, *P. freudenreichii* commence la consommation après que *L. lactis* arrête de le métaboliser, indiquant une segregation temporelle de l'utilisation de lactose, et pas de competition direct pour ce substrat.

Nous avons démontré ci-dessus que capturer la complexité des processus métabolique dans une communauté microbienne durant la fabrication de fromage requiert des considérations de la dynamique sous jacente au sein du système. Nous nous demandions si l'application d'approches computationelle *a prioio* se reposant sur les modèles métaboliques à l'échelle du génome peuvent mettre en avant de nouvelles interactions putatives entre les souches. A cette fin, nous avons utilisé deux méthodes basé sur les flux, SME-TANA (?) et MiCOM (?), et une approche basé sur le raisonnement, Metage2Metabo (?) dans le but de suggérer de nouvelle complémentarité métabolique au sein du consortium.

Les approches quantitative, MiCOM et SMETANA, ont mis en avant respectivement quatorze et vingt-cinq métabolites échangés (Table1.3 et 1.4), alors que l'approche par raisonnement, Metage2Metabo, identifie onze métabolites qui ne peuvent pas être produits

par une espèce sans interactions dans la communauté (Table supplémentaire 1.5). Une première observation est qu'un relativement grand nombre de métabolites sont en communs dans les prédictions des composés échangeables de SMETANA et MiCOM : lactate, qui est aussi prédit dans le modèle dynamique, phénylalanine, serine, malate, succinate, xanthine, H_2S , 2-ketoglutarate, glycollate et acétaldéhyde. Les autres composés échangés prédit incluent principalement des acides aminés (isoleucine, proline, glycine, alanine). Pour chaque composé mis en avant par SMETANA nous avons utilisé les données métagénomiques dans le but d'évaluer la validité des interactions les plus plausibles selon le score SMETANA (voir matériels et méthodes). Le H_2S est prédit comme produit de *L. lactis* et *L. plantarum* et donné à *P. freudenreichii*, le ribose est produit par *L. lactis* et *L. plantarum* et consommé par *P. freudenreichii*, le glycérol produit par les bactéries lactiques et consommé par *P. freudenreichii* et enfin la phénylalanine produite uniquement par *L. plantarum* et consommée par *P. freudenreichii*. Pour chacun de ces métabolites prédits comme échangés, nous avons vérifié l'expression des gènes associés à la production de ces métabolites chez les espèces donneurs, et à la consommation de ces métabolites chez les espèces receveuses (voir méthodes et Figure supplémentaire 1.8). Les résultats suggèrent que les interactions pour H_2S , ribose et le glycérol sont plausibles à plusieurs étapes de la fabrication du fromage. À l'inverse, alors que les données d'expression de gènes montrent que les voies métaboliques de la consommation de la phénylalanine sont fortement exprimées chez *P. freudenreichii*, leur production par *L. plantarum* ne l'est pas, suggérant que cette interaction semble être moins recevable. Enfin, l'ensemble des métabolites dont la production est prédit par Metage2Metabo comme nécessitant des interactions dans la communauté inclut divers acides gras, comme galactose-1-phosphate, benzyl-CoA, Glycéraldéhyde et xanthosine, indiquant une complémentarité métabolique entre les souches.

Bigg ID	Metacyc ID	Name	Metacyc Ontology	Export fluxes	Import
acald_e	ACETALD	acetaldéhyde	Aldehydes-Or-Ketones, Aldehydes	Pf; Ll	Lp
akg_e	2-KETOGLUTARATE	2-oxoglutarate	Acids, Organic-Acids	Pf	Lp
ala__D_e	D-ALANINE	D-ALANINE	All-Amino-Acids, Acids, Amino-Acids, Organic-Acids	Pf; Lp	Ll
ala__L_e	L-ALPHA-ALANINE	L-ALPHA-ALANINE	All-Amino-Acids, Acids, Amino-Acids, Organic-Acids	Pf; Lp	Ll
co2_e	CARBON-DIOXIDE	CARBON-DIOXIDE	Others, Others	Pf; Ll	Lp
coa_e	CO-A	co enzyme A	Groups, Others	Pf; Ll	Lp
fe2_e	FE+2	Fe2+	Ions, Inorganic-Ions, Cations	Lp	Pf;
gly_e	GLY	Glycine	All-Amino-Acids, Acids, Amino-Acids, Organic-Acids	Lp	Pf;
glyclt_e	GLYCOLATE	GLYCOLATE	Acids, Organic-Acids	Pf	Lp
gthox_e	OXIDIZED-GLUTATHIONE	glutathione disulfide	ORGANOSULFUR, All-Glutathiones	Lp	Ll
gthrd_e	GLUTATHIONE	glutathione	ORGANOSULFUR, All-Glutathiones, Thiols	Ll	Lp
h2o2_e	HYDROGEN-PEROXIDE	hydrogen peroxide	Peroxides, Others	Ll	Lp
h2o_e	WATER	H2O	Pseudo-Compounds, Others	Pf; Ll	Lp
h2s_e	HS	hydrogen sulfide	Ions, Anions, Inorganic-Ions	Ll	Lp;
ile__L_e	ILE	isoleucine	All-Amino-Acids, Acids, Amino-Acids, Organic-Acids	Lp	Pf;
lac__D_e	D-LACTATE	d-lactate	Acids, Organic-Acids	Pf	Ll;
lac__L_e	L-LACTATE	l-lactate	Acids, Organic-Acids	Ll	Lp;
mal__L_e	MAL	malate	Acids, Organic-Acids	Pf; Lp	Ll
nh4_e	AMMONIUM	ammonium	Ions, Inorganic-Ions, Cations	Lp	Pf;
phe__L_e	PHE	phenylalanine	Aromatics, All-Amino-Acids, Acids, Amino-Acids, Organic-Acids, Organic-aromatic-compounds	Pf	Lp
pro__L_e	PRO	proline	All-Amino-Acids, Acids, Amino-Acids, Organic-Acids	Ll	Pf;
ser__D_e	D-SERINE	serine	All-Amino-Acids, Acids, Amino-Acids, Organic-Acids	Ll	Lp;
ser__L_e	SER	serine	All-Amino-Acids, Acids, Amino-Acids, Organic-Acids	Ll	Lp;
succ_e	SUC	succinate	Acids, Organic-Acids	Ll	Pf;
xan_e	XANTHINE	xanthine	Others, Others	Ll	Lp;

TABLE 1.3 – **Exchangeable metabolites highlighted by MICOM on the cheese bacterial community.** Abbreviations : Ll, *L. lactis* ; Lp, *L. plantarum*, Pf, *P. freudenreichii*

Bigg ID	Metacyc ID	Name	Metacyc Ontology	Export fluxes	Import fluxes
acald_e	ACETALD	acetaldéhyde	Aldehydes-Or-Ketones, Aldehydes	Pf Pf Lp Ll	Ll Lp Ll Lp
akg_e	2-KETOGLUTARATE	2-oxoglutarate	Acids, Organic-Acids	Lp Pf	Pf Lp
fum_e	FUM	fumarate	Acids, Organic-Acids	Pf	Lp
glyc_e	GLYCEROL	GLYCEROL	Alcohols, All-Carbohydrates, Sugar-alcohols, Carbohydrates	Lp Ll Ll	Pf Pf Lp
glyclt_e	GLYCOLATE	GLYCOLATE	Acids, Organic-Acids	Lp	Pf
h2s_e	HS	hydrogen sulfide	Ions, Anions, Inorganic-Ions	Lp Pf Ll	Pf Lp Lp
lac__D_e	D-LACTATE	d-lactate	Acids, Organic-Acids	Pf Ll	Lp Lp
lac__L_e	L-LACTATE	l-lactate	Acids, Organic-Acids	Ll	Lp
mal__L_e	MAL	malate	Acids, Organic-Acids	Pf Lp Pf Ll	Ll Ll Lp Lp
phe__L_e	PHE	phenylalanine	Aromatics, All-Amino-Acids, Acids, Amino-Acids Organic-Acids, Organic-aromatic-compounds	Lp	Pf
rib__D_e	D-Ribopyranose	D-Ribopyranose	Aldehydes-Or-Ketones, All-Carbohydrates, Aldehydes, Carbohydrates	Ll Lp Pf Ll	Pf Pf Lp Lp
ser__D_e	D-SERINE	serine	All-Amino-Acids, Acids, Amino-Acids, Organic-Acids	Ll Lp Pf Lp Pf Ll	Pf Pf Ll Ll Lp Lp
succ_e	SUC	succinate	Acids, Organic-Acids	Ll Lp	Pf Pf
xan_e	XANTHINE	xanthine	Others, Others	Ll Lp	Pf Pf

TABLE 1.4 – **Exchangeable metabolites highlighted by SMETANA on the cheese bacterial community.** Abbreviations : Ll, *L. lactis* ; Lp, *L. plantarum*, Pf, *P. freudenreichii*

Bigg ID	Metacyc ID	Name	Metacyc Ontology
galt1p	D-galactose-1-phosphate	a D-galactopyranose 1-phosphate	All-Carbohydrates, Carbohydrates
benzcoa	BENZOYLCOA	benzoyl-CoA	Esters, Thioesters
xtsn	XANTHOSINE	XANTHOSINE	Organic-heterocyclic-compound, All-Nucleosides, Nitrogen-Molecular-Entities, Organic-heteromonocyclic-compounds, Organic-Heteropolycyclic-Compounds, Organonitrogen-Compounds, Nucleosides
glyald	GLYCERALD	glyceraldehyde	Aldehydes-Or-Ketones, All-Carbohydrates, Aldehydes, Carbohydrates
3hocoa	∅	∅	∅
3hhdcoa	CPD0-2232	(S)-3-hydroxyhexadecanoyl-CoA	Thioesters, Esters
3hdcoa	CPD0-2244	(S)-3-hydroxydecanoyl-CoA	Thioesters, Esters
3odcoa	CPD0-2123	3-oxodecanoyl-CoA	Thioesters, Esters
3hhcoa	OH-HEXANOYL-COA	(S)-3-hydroxyhexanoyl-CoA	Thioesters, Esters
udcpp	UNDECAPRENYL-P	all-trans-undecaprenyl phosphate	Lipids, Polyisoprenoids
3oocoa	CPD0-2106	3-oxooctanoyl-CoA	Thioesters, Esters

TABLE 1.5 – Metabolites predicted to be producible through metabolic cross-feeding only by Metage2Metabo



FIGURE 1.8 – **Gene expression with metatranscriptomic data.** This heatmap displays the decile of the RPKM of gene transcript counts at a given time point (columns) and in a given micro-organism (the colored areas indicate to which bacteria the gene belongs). Then, the decile were color coded from highest deciles (dark red, highest expression in this micro-organism at this time point) to lowest decile (light red, low expression). The displayed genes were manually selected from the list of possible interactions raised by SMETANA : for a selected interaction metabolite, genes involved in production (in the donor), and in consumption (in the receiver) were retained for analysis. Selected interaction metabolites were *H2S*, ribose and phenylalanine. Additionally, genes involved in lactose consumption were added to this list. Abbreviations : Cm, molding (t = 19.5 hours); Cdm, demolding (t = 40 hours); C0w, start of ripening (t = 60 hours); C4w, Fourth week ripening (t = 732 hours); C7w, seventh week ripening (t = 1236 hours).

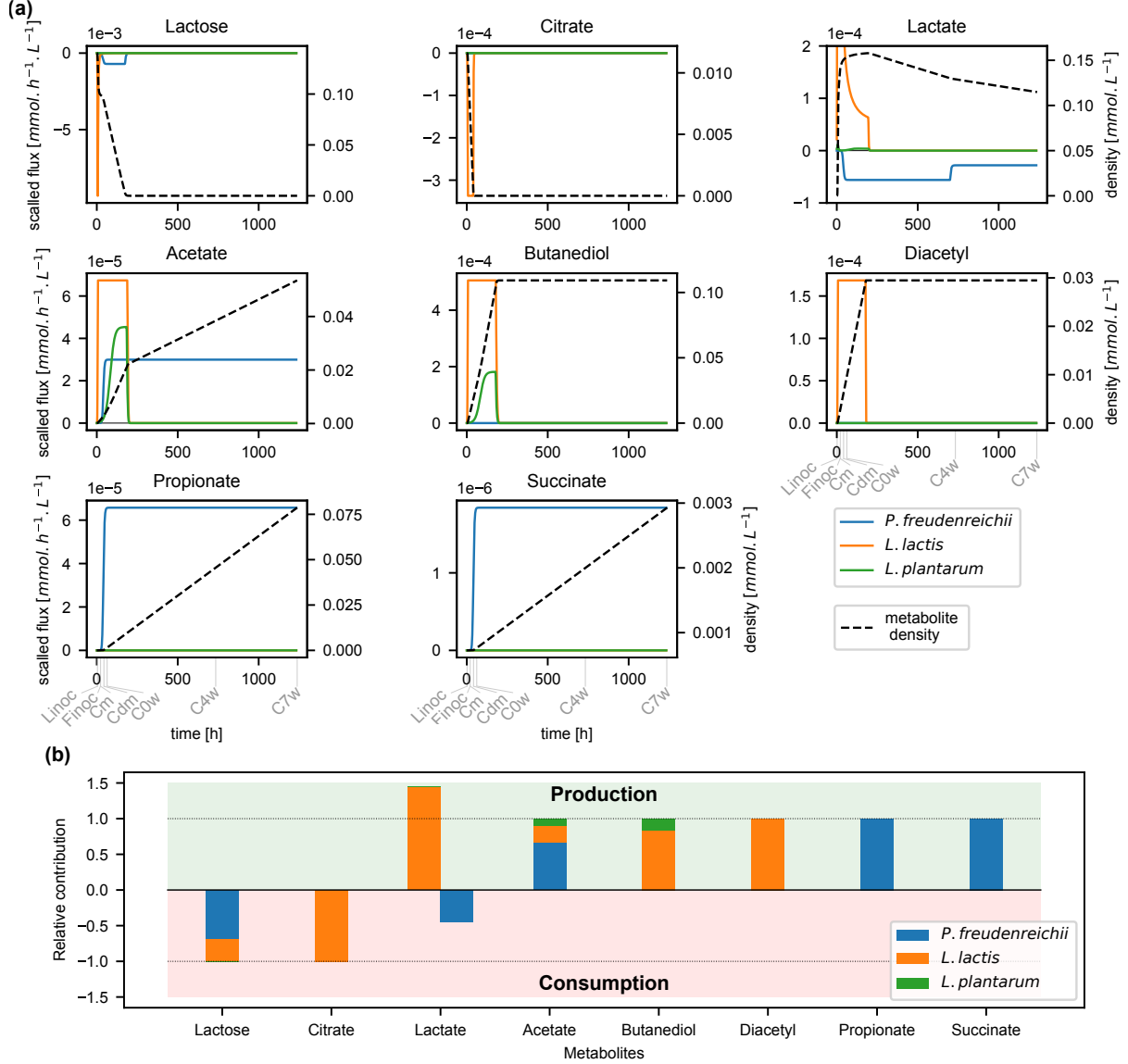


FIGURE 1.9 – (a) **Consumption and production flux dynamics.** We represent for each bacterium the dynamics of the consumption and production fluxes of the followed-up metabolites (colored plain lines). We also represent the dynamics of the concentration of the corresponding metabolite in the cheese (dashed black line, right axis for scale). (b) **Bacterial relative contribution to the metabolite fate.** We represent the relative contribution of each bacteria to the final metabolite concentration. Namely, we computed each microbial overall contribution by integrating in time its consumption or production flux ($\int_0^t \mu_{i,j} b_i$ for metabolite j and bacteria i , see eq.1.8), and normalized the result by the sum over the three bacteria. The value 1 (or -1) then represents the net production (or consumption), i.e. the difference between the final and initial concentration of the metabolite (horizontal dashed lines). For each metabolite, the bacterial contribution is represented by a colored bar, with positive value for production, and negative value for consumption. Abbreviations : Linoc, Lactic acid bacteria inoculation ($t = 0$); Finoc, *P. freudenreichii* inoculation ($t = 18$ hours); Cm, molding ($t = 19.5$ hours); Cdm, demolding ($t = 40$ hours); C0w, start of ripening ($t = 60$ hours); C4w, Fourth week ripening ($t = 732$ hours); C7w, seventh week ripening ($t = 1236$ hours).

1.3 Discussion

Dans ce travail, nous fournissons un modèle de jumeau numérique du métabolisme bactérien pendant le processus de fabrication de fromage. En utilisant les modèles métaboliques de trois souches bactériennes inoculées et des données omiques, nous étions capable de reproduire, avec précision, les motifs de production de métabolites ainsi que la dynamique des populations bactériennes. Nous nous sommes reposés sur les conditions expérimentales de cultures pures pour calibrer les modèles métaboliques individuels avec peu de paramètres pour chaque, ainsi limitant la surévaluation du modèle durant les simulations communautaires. Une originalité du modèle est sa capacité de prédire les dynamiques sur tout le procédé de fabrication du fromage (sept semaines). Notre travail soulève des hypothèses sur les voies métaboliques mis en jeu dans le lait, ainsi que la contribution de chaque souche bactérienne dans la consommation de nutriments et la productibilité de composé organoleptiques.

La connaissance des microbiologistes, de la littérature scientifique et des données multi-omiques étaient nécessaire pour accomplir des modèles métaboliques à l'échelle du génome de grandes qualités et les simulations dynamiques qui en découlent. Le métabolisme carbonné central des bactéries lactiques produisent principalement l'acide lactique à partir de la dégradation du lactose *via* les voies métaboliques de la glycolyse, du tagatose et de Leloir (???) ; pour la première fois, l'utilisation du citrate et de la fermentation hétérolactique ont été observé dans le modèle curé de *L. plantarum* en lait. La vérification que chaque modèle métabolique individuel peut activer les différentes voies métaboliques décrites plus haut était ainsi une étape important pour valider les modèles, permettant la reproduction d'activités enzymatiques observé dans (??). La bactérie propionique *P. freudenreichii* convertie le lactose, et préférentiellement le lactate en propionate selon (??). En suivant (?), nous avons ajouté l'enzyme propionyl-CoA :succinate CoA transferase (2.8.3.-) pour compléter la cycle de Wood-Werkman, et ainsi assurer la production de propionate.

Dans (?), les paramètres individuels de calibration pour tous les nutriments du milieu de culture sont imposés dans le but d'expliquer la métabolomique à l'échelle de l'espèce,

et de les prédire à l'échelle de la communauté. En comparaison de ce travail, nous avons implémenté un modèle FBA dynamique en se basant sur (?), qui a utilisé des paramètres de calibration souche spécifique pour prédire la croissance, le pH et les concentrations métaboliques. Pour éviter la surévaluation du modèle, nous avons réduit le nombre de paramètres inférés pour chaque bactérie, où seulement deux paramètres sont gardés pour la régulations du pH pour les bactéries lactiques et un paramètre se focalisant sur les croissances de *L. plantarum* et *P. freudenreichii*. Des calibrations additionnelles ont été mené pour *P. freudenreichii* sans aucune optimisations mathématiques supplémentaires, en calculant la valeur de la borne supérieure des métabolites dosés en monoculture.

Le modèle de co-culture a donné des indices sur le comportement de la communauté pendant la fabrication du fromage et a suggéré temporellement la contribution de chaque espèce impliqué dans la production des composés aromatiques. Il est montré que la production de propionate et d'acetate peut être attribué uniquement à *P. freudenreichii*, comme indique dans (?), et que le diacetyl semble être seulement produit par *L. lactis*. D'après le modèle de communauté, le butanediol a été produit pendant l'étape du moulage et au début de l'affinage par à la fois, *L. lactis* et *L. plantarum*, comme observé dans les modèles de mono-cultures respectifs. En regardant la production d'acétate, le modèle déduit une production précoce de l'étape du moulage jusqu'à l'affinage par *L. lactis* et *L. plantarum* suivi par *P. freudenreichii* pendant l'affinage, responsable de la majeure partie de production de l'acétate à cette étape. Concernant l'utilisation des sources de carbonnes, alors que *L. plantarum* possède les voies métaboliques de dégradation du lactose et du citrate, qui sont complètement activées dans le modèle en monoculture, il apparaît que ces voies métaboliques sont fortement sous régulés en co-culture. Lors de la mise en culture avec *L. lactis*, et comme la vitesse de croissance de *L. plantarum* est la plus faible des deux, il accomplit son efficacité métabolique maximale après la phase exponentielle de *L. lactis*. Comme *L. lactis* pousse principalement sur le lactose produisant le lactate, sa phase exponentielle est associées à la forte acidification de milieu, ce qui active la régulation du pH par la répression du lactose sur le métabolisme de *L. plantarum*. *L. plantarum* peut utiliser des acides aminés ou d'autres métabolites non suivi par le modèle dynamique permettant sa croissance. Pour explorer cette piste, nous avons utilisé des outils dans le but

d'identifier des interactions bactériennes en dehors de l'ensemble des composés suivi par le modèle dynamique en co-culture, révélant d'autres interactions métaboliques putatives, où les acides aminés représentent une grande proportion.

Dans le modèle de co-culture, *P. freudenreichii* garde activé son métabolisme dans la dernière phase de production d'un fromage : l'affinage. Ce métabolisme est reflété par la consommation de lactate durant tout l'affinage, et le relargage de composés organoleptiques à flux constant (Figure 1.9). Ce comportement est cohérent avec les capacités métaboliques connues de *P. freudenreichii* et avec les données métabolomiques obtenues durant le procédé de production du fromage. Dans le modèle, *P. freudenreichii* consomme aussi le lactose restant après la croissance de *L. lactis*. En effet, durant sa phase exponentielle, *L. lactis* métabolise principalement le lactose, produisant fortement du lactate et réduisant le pH. La réduction du pH inhibe l'absorption du lactose régulée par le pH dans les deux bactéries lactiques, et donc stoppant la consommation de lactose par les *Lactobacilli*. Comme le lactose reste dans le milieu de culture et que *P. freudenreichii* est l'unique bactérie capable de l'utiliser, il épuise donc le lactose encore présent. Dans les données métatranscriptomiques, le gène LACZ est hautement activé par *P. freudenreichii* durant l'affinage confirmant la dégradation du lactose par cette souche après l'acidification (Figure 1.8)

Ce travail révèle comment la combinaison d'expert biologique, des données omiques et de la modélisation métabolique permet d'obtenir des prédictions précises sur les mécanismes responsables du comportement dynamique de la communauté microbienne. D'autre part, cela souligne également que la quantité de données et les efforts nécessaires pour créer de tels modèles de haute qualité restent un prix élevé à payer malgré l'amélioration des approches de simulation. Des développements méthodologiques doivent encore être proposés afin d'automatiser l'étalonnage des modèles avec des données et de garantir à la fois l'exactitude des prédictions et l'extensibilité à des communautés plus importantes ou à des communautés de composition empirique.

1.4 Matériels et méthodes

1.4.1 Données biologiques

1.4.1.1 Composition communauté.

Les souches qui ont été utilisées dans le projet sont décrites dans (?). Brièvement, la communauté bactérienne contrôlée est composée de deux bactéries lactiques (LAB), *L. lactis* subsp. *lactis* biovar *diacetylactis* CIRM-BIA1206 (nommé *L. lactis*), *L. plantarum* CIRM-BIA465 (nommé *L. plantarum*), et une bactérie propionique *P. freudenreichii* CIRM-BIA122 (nommé *P. freudenreichii*) provenant du Centre international pour les ressources microbiennes - Bactéries associées aux aliments (CIRM-BIA ¹)

1.4.1.2 Fabrication de fromage et collecte des échantillons.

Le protocole standard de production d'un fromage décrit dans (?) a été utilisé. Brièvement, après la collecte de lait frais, pasteurisé (76, 20 secondes), écrémer, standardisé (jusqu'à 30 grammes de gras et 36 mg de calcium par kg de lait, *L. lactis* et *L. plantarum* sont inoculés à 5.7 et 5.2 log₁₀ CFU (Unité de Formant Colonie). Après pré-maturation du lait inoculé (à 14 pour 18 h), *P. freudenreichii* a été inoculé à 6.1 log₁₀ CFU et remué pendant 20 min. Le lait est versé dans des cuves et chauffé à 33 pendant 30 min. De la présure commerciale a été ajoutée et le coagulum fut coupé, mélangé, lavé, drainé, pré-pressé et transféré dans des moules pour le présurage. Le démoulage apparaît le troisième jour de la fabrication du fromage, et les fromages ont été salés (immersion à 12 pendant 10h avec de la saumure en excès), puis séchés toute la nuit et emballés sous vide dans des sacs en plastique le quatrième jour de la fabrication du fromage pour un affinage de 7 semaines à 13 .

Après l'inoculation des bactéries lactiques (Linoc, t=h) et de *P. freudenreichii* (Finoc, t=18h), les échantillons sont collectés à cinq étapes de production pour de plus amples analyses comme décrit dans (?) : l'étape de moulage (Cm, t=19.5h), l'étape de démoulage (Cdm, t=40h), le commencement de l'affinage (C0w, t=60h) après quatre et sept semaines

1. <https://collection-cirm-bia.fr/>

d'affinage (C4w, t=732h et C7w, t=1236h). Les fromages sont répliqués quatre fois (des répétitions biologiques ont été réalisés à différents jours avec des incolua et laits différents), et échantillonnés chaque fois pour la numérations et la biochimie et deux fois (répétitions biologiques) for le séquençage ARN.

1.4.1.3 Expériences de culture souche spécifique.

Dans le but de calibrer les modèles métaboliques, un jeu de données d'entraînement a été réalisés en utilisant les données expérimentales de chaque souche individuel. Ces croissances individuelles des trois bactéries fut réalisé dans un lait gras UHT (Delisse) en tube dans des conditions microaerobic, c'est à dire, sans secouer dans l'incubateur a 30, avec deux réplicats. Pour *P. freudenreichii*, le lait fut supplémenté en lactate (20 g/L) et avec de la peptone caseine (10 g/L). Une seconde expérience de croissance fut réalisé pour *P. freudenreichii* dans le but de mesurer la concentration finale de metabolites (acetate, lactate, propionate, et succinate). *L. lactis* fut inoculé a 1×10^6 et 4×10^6 CFU, *L. plantarum* a 3×10^6 and 6×10^6 CFU et *P. freudenreichii* a 1.4×10^7 et 9×10^6 CFU. Les données de croissance en culture pure sont disponibles en Table ??.

1.4.1.4 Données métatranscriptomiques.

Dix grammes de fromage ont été dispersé dans 90 grammes d'une solution de citrate de sodium à 2% durant 1 minute à vitesse maximal suivie de 1 min à faible vitesse dans un stomacher (Humeau, Treillière, France) à température ambiante. Dix millilitres de la dispersion ont été centrifugés à 10000 grammes à 4 pendant cinq minutes. Les culots cellulaires ont été conservés à -80 pour l'extraction de l'ARN. Après décongélation, les échantillons ont été centrifugés à nouveau pendant 5 minutes à 4, 10 000 grammes pour retirer le surnageant. Les culots cellulaires ont été suspendus dans 200 L de tampon de lyse (50 mM Tris-HCl, 1 mM EDTA ; pH 8,0) 20 mg/g de lysozyme (MP Biomedicals, Illkirch, France) et 50 U/g de mutanolysine (Sigma, Saint Quentin Fallavier, France) et incubés pendant 30 min à 24. Les suspensions ont ensuite été transférées dans des tubes de 2 mL contenant 50 mg de billes de zirconium (diamètre : 0,1 mm ; BioSpec Products, Bartlesville) et 50 L de SDS (20 %). Les échantillons ont ensuite été agités dans un Pre-

cellys Evolution (Bertin, Montigny-le-Bretonneux, France) pendant deux cycles de 40 s à 6500 rpm. L'ARN a ensuite été extrait des lysats cellulaires à l'aide de Qiazol et du kit RNeasy Mini (Qiagen) selon (?) avec deux ou trois étapes de traitement ultérieur à la DNase (Dnase Rnase free, Ambion) selon les instructions du fabricant. L'intégrité de l'ARN (RIN) a été évaluée à l'aide d'un bioanalyseur Agilent 2100 (Agilent Technologies, Santa Clara, CA). Les concentrations d'ARN ont été quantifiées à l'aide de la quantification fluorométrique Qubit (ThermoFisher, France). L'absence de contamination par l'ADN a également été vérifiée à l'aide du Qbit et du protocole de quantification de l'ADN conformément aux instructions du fabricant. Les échantillons d'ARN ayant une valeur RIN supérieure à 8, indiquant une bonne intégrité de l'ARN, ont été conservés pour la suite de l'analyse.

Les données métatranscriptomiques ont été obtenues à cinq étapes de la fabrication du fromage. Les librairies Illumina et le séquençage ont été réalisés par GeneWiz (Leipzig, Allemagne). Les lectures sont disponibles à l'ENA (Cambridge) sous le numéro d'accès ENA : PRJEB42478. Les lectures séquencées ont été découpées à l'aide de Trim Galore et mises en correspondance avec les génomes de référence à l'aide de Bowtie 2 (?). Aucune erreur d'appariement n'a été autorisée et seules les lectures correspondant à des séquences uniques ont fait l'objet d'une analyse plus poussée. Les lectures correspondant à chaque séquence codante ont été comptées à l'aide de HTSeq-count (-stranded=reverse, -a 0, -t CDS, -i db_xref, -m union) (?).

Les analyses statistiques ont été réalisées à l'aide du logiciel statistique R version 4.0.3 (?). Les données brutes de comptage de l'ARN-Seq ont été normalisées en deux étapes. Tout d'abord, un facteur d'échelle spécifique à l'espèce a été appliqué pour éliminer les biais de composition entre les bibliothèques. Pour chacune des trois espèces, elles ont été calculées en utilisant tous les gènes de l'espèce considérée et deux autres rangées pour l'autre espèce, calculées comme la somme des comptages attribués à chacune de ces deux autres espèces, en utilisant la méthode TMM (Trimmed Mean of M-Values) telle qu'implémentée dans le package edgeR version 3.32.1 (?). Une étape supplémentaire de normalisation au sein de l'échantillon a été réalisée pour corriger la longueur des gènes et permettre la comparaison. Enfin, la moyenne des réplicats a été calculée.

1.4.1.5 Métabolomiques ciblées.

Les méthodes sont décrites dans (?). Brièvement, les sucres et les acides organiques dans les échantillons sont quantifiés en utilisant la chromatographie liquide à haute performance (HPLC). Les composés volatils sont analysés en utilisant l'extraction par sorption de l'espace de la tête (HS) et couplés à la chromatographie en phase gazeuse-spectrométrie de masse (GC-MS).

1.4.2 Modelisation métabolique

1.4.2.1 Séquençage du génome.

Des culots cellulaires (équivalent à 10^{10} CFU) ont été obtenus par centrifugation pendant 10 minutes à 5000 g à partir de cultures de souches dans un milieu de bouillon. L'ADN fut extrait en utilisant le kit DNeasy midi (Qiagen) d'après le protocole décrit dans (?). Les génomes sont séquencés par Illumina NovaSeq 2 x 1500 pb et PacBio Sequel (Eurofins Genomics, Constance, Allemagne). Les génomes sont assemblés en utilisant Unicycler (??) sur la plateforme bioinformatique Galaxy Genotoul (Toulouse, France). Les séquences génomiques sont intégrés sur la plateforme MicroScope hébergé au Genoscope (CEA, Evry, France) pour une annotation automatique selon (?). Les génomes annotés sont disponibles à l'ENA (Cambridge) sous le numréo d'accèsion ENA : PRJEB54980.

1.4.2.2 Reconstruction de modèle métaboliques à l'échelle du génome.

Les génomes bactériens individuels sont annotés fonctionnellement en utilisant la plateforme MicroScope (?) Les réseaux métaboliques à l'échelle du génomes sont par la suite reconstruits à partir de ces génomes annotés avec CarveMe (?) version 1.4.1. Chaque modèle métaboliques à l'échelle du génome est poli en utilisant ModelPolisher (?) qui améliore l'annotation en ajoutant des informations liés aux identifiants BIGG depuis la base de connaissances BIGG (?). La consistance de chaque modèle est vérifié avec MEMOTE (?) à la fin de l'étape de polissage (<https://memote.io>). Le milieu nutritionnel inclus le milieu minimal trouvé par CarveMe et les composé du lait (<https://>

//ciqual.anses.fr/#/aliments/19024/lait-entier-pasteurise, où les acides aminés sont considérés en excès (fichier supplémentaire 1).

1.4.2.3 Analyse par équilibre des flux.

Un modèle métabolique à l'échelle du génomes (GEM) est reconstruit pour chaque souche en utilisant le réseaux métabolique et la composition nutritionnelle, et analysé avec la méthode de l'analyse par équilibre des flux (FBA) (?). Le FBA calcule une distribution de flux optimal satisfaisant la stoechiométrie et la faisabilité biophysique donnant un maximum de biomass, en solvant le programme linéaire suivant :

$$\begin{aligned} \text{trouver } v^* \in \mathbb{R}^{N_r} \text{ tel que } v^* = & \arg \max v_{growth} \\ & S.v = 0 \\ & v_{min}^{in} \leq v \leq v_{max}^{in} \\ & v_{min}^{ex} \leq v \leq v_{max}^{ex} \end{aligned} \quad (1.5)$$

ou v_{growth} est le flux de la reaction de biomasse de modèle, S est la matrice stoechiométrique dérivé du modèle métabolique, v_{min}^{in} et v_{max}^{in} (resp. v_{min}^{ex} et v_{max}^{ex}) sont les valeurs minimales et maximales de flux intracellulaires (resp. échangés) définissant les bornes de flux pour le vecteur de flux v contenant \mathbb{R}^{N_r} dimension. Les ensembles de valeurs $(v_{min}^{in}, v_{max}^{in})$ et $(v_{min}^{ex}, v_{max}^{ex})$ dans les GEMs sont disponibles dans le matériel supplémentaire.

L'équation (1.5) définit alors, pour chaque bactérie i , la correspondance μ_i entre les vecteurs de contraintes sur les réactions d'échange définissant l'environnement nutritionnel et le flux optimal v^* obtenu dans ce contexte nutritionnel.

$$\mu_i(v_{min}^{ex}, v_{max}^{ex}) = v^* \quad (1.6)$$

Le problème d'optimisation FBA (1.5) fut résolu en utilisant CobraPy (?)version 0.17.1 utilisant l'ordre lexicographique pour garantir une solution unique (?).

1.4.2.4 Raffinement manuel des GEMs

Une étape de raffinement manuel fut réalisé pour tous les modèle pour assurer la bonne production de composés organoleptiques et l'activité des fonctions métaboliques d'intérêt au regard de la littérature. Les modifications sont récapitulées dans le matériels supplémentaire Table 1.6. Comme les trois espèces sont micro-aerobique, la borne d'import de l'oxygène est défini comme la valeur minimal permettant la production de composés aromatiques.

GEM	Reaction	Bound	Value	Justification
<i>L. lactis</i>	ACTD2	lower upper	0 0	Allows the activation of the acetolactate pathway
	ACLDC	lower upper	0 2	Allows the activation of the acetolactate pathway
	ACLS	lower upper	2 1000	Allows the activation of the acetolactate pathway
	LACZ	lower upper	0.0006 2	Force the consumption of lactose
<i>L. plantarum</i>	PFL	lower upper	0 0	Permit the production of both isomers of lactate from lactose.
	LCTSt3ipp	lower upper	0 0	Leads the flux of lactose into the Glycolysis pathway
	ACTD2	lower upper	0 0	Allows the activation of the acetolactate pathway
	GALUi	lower upper	-30 10	Permit the production of both isomers of lactate from lactose.
	ACLS	lower upper	2 1000	Allows the activation of the acetolactate pathway
<i>P. freudenreichii</i>	6PGALSZ	lower upper	0 0	Tagatose pathway was used
	XYLI2	lower upper	0 0	Production of fructose from glucose instead of glucose-6-phosphate. Glycolysis was not activated
	POX2	lower upper	0 0	Regulate flux of acetate
	PPAKr	lower upper	0 0	Inhibit acetate over production
	LACZ	lower upper	0 5	Allow consumption of lactose
	PTAr	lower upper	-1000 8	Allow the appropriate ratio of acetate/propionate
	2131pyrpp	lower upper	8 1000	Allow the appropriate ratio of acetate/propionate

TABLE 1.6 – Manual refinement of the GEMs and corresponding justifications.

1.4.2.5 Modélisation dynamique

Le comportement dynamique de chaque système de culture est calculé en utilisant la l'analyse par équilibre des flux dynamique (dFBA) (?) en ajoutant des mécanismes dynamiques de la population. Notons $b = (b_i)$ pour $i \in \mathcal{B} = \{L. lactis, L. plantarum, P. freudenreichii\}$ les densités de la population bactérienne, et $m = (m_j)_{1 \leq j \leq N_m}$ la densité de composés métaboliques N_m qui sont expérimentalement suivi, nous avons ainsi le système, pour $i \in \mathcal{B}$ et $1 \leq j \leq N_m$

$$\partial_t b_i = \mathcal{R}_i(b_i) \mu_{i,i} ((c_{min,i}^{ex}, c_{max,i}^{ex})(b, m)) b_i \quad (1.7)$$

$$\partial_t m_j = \sum_{i \in \mathcal{B}} \mu_{i,j} ((c_{min,i}^{ex}, c_{max,i}^{ex})(b, m)) b_i \quad (1.8)$$

où le terme $\mu_{i,i}$ est la composante correspondant au métabolite j (ou à la biomasse i) de μ_i , calculé à partir de 1.6 compte tenu de l'ensemble des contraintes $(c_{min,i}^{ex}, c_{max,i}^{ex})(b, m)$. Ces contraintes dependent dynamiquement de l'état de la variable b et m .

Le vecteur \mathcal{R}_i modélise un processus de régulation de la population sur la croissance de la population de manière phenomenologique avec le terme

$$\mathcal{R}_i = \lambda_i (1 - \frac{b_i}{\beta_i}), \text{ for } i \in \mathcal{B} \quad (1.9)$$

où λ_i est un poids et β_i est la capacité de charge de la population, correspondant à la valeur de la phase plateau de chaque bacterie dans l'expérience.

Pour un substrat j et une modèle métabolique i , la bornes usuelles sont

$$c_{min,i,j}^{ex} = \max(-\frac{m_j}{\Delta t \sum_{i \in \mathcal{M}_j} b_i}, v_{i,j}^{int}) \quad c_{max,i,j}^{ex} = 0, \quad (1.10)$$

où \mathcal{M}_j est le sous ensemble de bactérie métabolisant j and Δt est une caractéristique temporelle du temps d'importation. Ces équations reflète le partage uniforme des ressources parmi les bactéries durant la fenêtre temporelle Δt . Quand les substrats sont en excès, le flux est limité par la limite d'import intrinsèque $v_{i,j}^{int}$.

Certains substrats et produits sont également régulés. En suivant l'approche de (?),

comme le métabolisme du lactose par les bactéries lactiques est régulé négativement par la forme non dissocié de l'acide lactique avec une décroissance exponentielle, pour $j = lcts_e$ et $i \in \{L. lactis, plantarum\}$, nous avons

$$c_{min,i,j}^{ex} = \max\left(-\frac{m_{lcts_e}}{\Delta t * \sum_{i \in \mathcal{M}(lcts_e)} b_i}, -\mu_{max,lcts} * 10^{(-k_{lac} * \phi_{undiss})} - \mu_{min,lcts}\right)$$

où k_{lac} correspond à la décroissance exponentielle, $\mu_{max,lcts}$ et $\mu_{min,lcts}$ sont les valeurs maximales et minimales et ϕ_{undiss} est une fonction calculant la concentration de la forme non dissocié de l'acide lactate à partir du lactate. Cette fonction ϕ_{undiss} est dérivée de l'équation de Henserson-Hasselbalch comme dans (?) et lit

$$\phi_{undiss}(m_{lac_L_e}, m_{lac_D_e}) = \frac{m_{lac_L_e} + m_{lac_D_e}}{1 + 10^{c_1 * (m_{lac_L_e} + m_{lac_D_e}) + c_2}}. \quad (1.11)$$

Dans cette équation le terme $c_1 * (m_{lac_L_e} + m_{lac_D_e}) + c_2$ est une approximation linéaire de $(pH - pKa)$, alors que les paramètres c_1 et c_2 sont inférés par une régression linéaire des données de co-culture. La production de lactate est régulée de la même manière, alors que la production d'acétate est régulé en fonction de la disponibilité en lactose. Voir matériel supplémentaire section ?? pour une description détaillée et justifications des bornes dynamiques configurées sur les réactions d'échange.

Pour *P. freudenreichii*, les bornes supérieures de productions sont évalué en calculant par le flux de production à partir des dosages de métabolites et des données de croissance dans les expériences de culture pure, en supposant un flux constant (voir le matériel supplémentaire ??).

La condition initiale, l'inoculum et l'environnement nutritionnel constitué des composés du lait, acides aminés et co-facteurs, a été défini comme spécifié dans (?), après conversion des unités (voir matériel supplémentaire pour la liste des métabolites). Une masse cellulaire de 0.33×10^{-12} CFU a été appliqué pour la conversion d'unité de biomasse.

1.4.2.6 Implementation numérique

Le système dynamique est résolu avec le schéma semi-implicite d'Euler. A chaque étape n , après le calcul de F_j le flux total du composé j comme défini par l'ensemble des GEMs,

$$F_j = \sum_{i \in \mathcal{B}} \mu_{i,j} \left((c_{min,i}^{ex}, c_{max,i}^{ex})(b^n, m^n) \right) b_i,$$

nous calculons,

$$\begin{aligned} b_i^{n+1} &= b_i^n + \Delta t * F_{b_i} \\ m_j^{n+1} &= \begin{cases} m_j^n + \Delta t * F_j & \text{si } F_j > 0 \text{ (cas explicite)} \\ m_j^n / (1 - \Delta t * F_j / m_j^n) & \text{sinon (cas implicite)} \end{cases} \end{aligned}$$

Ce schéma garantie la positivité de la solution pour tous les temps.

Les problèmes dFBA (1.7)–(1.8) ainsi que les paramètres inférés sont résolu en créant un script Python et en utilisant les bibliothèques `numpy` v.1.19.5(?), `pandas` v.0.25.3 (?), and `scipy` v.1.5.3 (?). Durant le dFBA, le FBA est résolu en utilisant l'ordre lexicographique défini (?).

1.4.3 Ajustement du modèle

Pour ajuster les paramètres du modèle, nous avons adapté les équations (1.7)–(1.8) aux expériences de monocultures. Nous avons ensuite inféré les le vecteur de paramètre du modèle θ du résultat du modèle de dFBA en maximisant le logarithme de la ressemblance des données observé en monoculture. Formellement, pour $i \in \{L. lactis, L. plantarum\}$, nous avons minimisé la fonction coût suivante

$$J(b_i, pH | \theta_i, b_{i,exp}, pH_{exp}) = \left\| \frac{\log_{10}(b_i) - \log_{10}(b_{i,exp})}{\sigma_{log,i,exp}} \right\|^2 + \alpha \left\| \frac{pH - pH_{exp}}{\sigma_{pH,exp}} \right\|^2 \quad (1.12)$$

où σ (resp. $\sigma_{log,i}$) est l'écart-type (resp. l'écart-type log transformé) des données

correspondantes, et $\theta_i = (k_{lac}, \mu_{max,lcts})$ si $i = L. lactis$ et $\theta_i = (k_{lac}, \mu_{max,lcts}, \lambda_i)$ si $i = L. plantarum$. Pour $P. freudenreichii$, la fonction coût est

$$J(b, m | \theta_i, b_{exp}, m_{exp}) = \left\| \frac{\log_{10}(b) - \log_{10}(b_{exp})}{\sigma_{log,b,exp}} \right\|^2 + \alpha \left\| \frac{m - m_{exp}}{\sigma_{m,exp}} \right\|^2 \quad (1.13)$$

où m_{exp} est le dosage final d'acétate, lactate, propionate et succinate, et $\theta_i = (\lambda_i)$. Nous avons gardé le nombre de paramètres d'ajustement petit (deux paramètres pour *L. lactis*, trois pour *L. plantarum* et un pour *P. freudenreichii*) pour éviter la sur prédiction. Les résultats d'optimisation sont présentés dans la Table ??.

1.4.4 Inférence des interactions métaboliques

Dans le but de prédire des échanges métaboliques comme ceux qui apparaissent dans le consortium microbien, MICOM v 0.32.0 (?) et SMETANA v1.2.0 (?) ont été utilisé sur l'environnement nutritionnel du lait. En plus, la prédiction d'une complémentarité métabolique entre les trois modèles métaboliques poussant dans le lait fut révélé avec Metage2Metabo v1.5.0 (?). Tous les métabolites prédit par les outils sont classés dans des catégories supérieures : les identifiants de la base de données Metacyc v25.5 (?) des composés BIGG ont été retrouvé et l'ontologie de ces composés ont été utilisé. La liste des composés prédit comme intervenant dans ue interaction d'alimentation croisé est disponible à la Table supplémentaire 1.5.

Un sous ensemble d'interaction bactérienne putative fut sélectionné parmi ceux prédit par SMETANA, selon son propre score avec un seuil choisi à 0.5. La pertinence des métabolites échangés sélectionné fut vérifié en utilisant les données de métatranscriptomiques (RPKM et moyenne des réplicats). Pour chaque métabolite, les réactions de production (dans l'espèce productrice) et de consommation (dans l'espèce consommatrice) des métabolites fut extraites en utilisant l'association gene-proteine-reactions. Dans le but d'évaluer l'expression putative de ces gènes, les valeurs d'expressions ont été comparé à l'expression globale des gènes dans chaque espèce, séparé en quartiles.

1.4.5 Visualisation et analyses statistiques

Les figures ont été générées en utilisant `matplotlib` version 3.3.4 (?). Les cartes métaboliques des voies métaboliques ont été générées avec Escher (?) pour Python v3.6.9 (?)², après avoir lancé un FBA pour chaque modèle. Les flux des réactions d'intérêt pour chaque espèce ont été normalisés par la valeur de flux maximal des voies métaboliques ciblées de l'espèce correspondante.

1.4.6 Code et disponibilité des données

Le code pour les simulations métaboliques et les modèles métaboliques sont disponibles dans https://forgemia.inra.fr/tango/tango_models. Les génomes des trois souches bactériennes et leurs annotations sont disponibles dans l'ENA sous le numéro d'accès PRJEB54980. Les données métatranscriptomiques sont disponibles dans l'ENA sous le numéro d'accès PRJEB42478. Les génomes annotés, les lectures ARN et RPKM sont déposés dans <https://entrepot.recherche.data.gouv.fr/privateurl.xhtml?token=4bf01466-29dd-4d71-b25e-b775d9cc39dc>. Le dosage biochimique et métabolique sont donc disponibles (?).

1.5 Conclusion

Nous avons décrit dans ce chapitre une méthodologie numérique à l'échelle de chaque individu qui permet de garantir des bons résultats lors du passage à l'échelle de la communauté. Pour rappel elle consistait tout d'abord en un raffinement de chaque modèle FBA individuel dans le but de vérifier si les flux métaboliques passent par les bonnes voies métaboliques. Puis, une combinaison entre le processus de partage équilibré des ressources ainsi qu'un faible nombre de paramètres optimisés ont permis de reproduire les observations biologiques et d'émettre des nouvelles hypothèses sur des possibles interactions bactériennes.

Tout ce travail demande de la patience, notamment pour l'étape de curation des modèles

2. <https://escher.readthedocs.io/en/latest/escher-python.html>

et requiert une grande quantité de données, deux points difficilement applicables pour le criblage au débit de communautés microbiennes. Le chapitre suivant présente ainsi une méthode existante utilisable pour le criblage de communautés, permettant de s'abstenir des deux principales contraintes mis en avant dans ce chapitre, et son extension pseudo-quantitative.