



Modélisation des dynamiques adaptatives de la levure de boulanger *S. cerevisiae* dans un environnement saisonnier

Dorian Collot

► To cite this version:

Dorian Collot. Modélisation des dynamiques adaptatives de la levure de boulanger *S. cerevisiae* dans un environnement saisonnier. Evolution [q-bio.PE]. Université Paris Saclay (COMUE), 2018. Français. NNT : 2018SACLS179 . tel-01838370

HAL Id: tel-01838370

<https://theses.hal.science/tel-01838370>

Submitted on 13 Jul 2018

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

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

Modélisation des dynamiques adaptatives de la levure de boulanger *S. cerevisiae* dans un environnement saisonnier

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°577 Structure et Dynamique des Systèmes Vivants
(SDSV)

Spécialité de doctorat: Science de la vie et de la santé

Thèse présentée et soutenue à Gif sur Yvette, le 19 juin 2018, par

Dorian Collot

Composition du Jury :

Cécile Fairhead
PR, Univ Paris-Sud (– UMR GQE-Le moulon)
Jérôme HARMAND
DR, INRA (– LBE-Inra)
Denis Roze
CR, CNRS (– Station Biologique de Roscoff)
Aymé Spor
CR, INRA (– UMR 1347 Agroécologie)
Olivier C. Martin
DR, INRA (– UMR GQE-Le Moulon)
Judith Legrand
MCF, Univ Paris-Sud (– UMR GQE-Le Moulon)

Présidente du jury
Rapporteur
Rapporteur
Examineur
Directeur de thèse
Co-Directrice de thèse

Table des matières

Introduction.....	7
I. A propos de l'évolution	7
A. Pourquoi s'intéresse-t-on à l'évolution ?	7
B. Comment évoluent les populations ?	7
II. Les variations du phénotype et les relations entre les traits	9
III. L'adaptation et la valeur sélective.....	11
A. Evolution des traits d'histoire de vie	12
B. Paysages adaptatifs	13
IV. Le rôle de l'environnement dans la sélection	13
A. Effet des modifications de l'environnement	14
B. Devenir de la population (généraliste/spécialiste) selon la forme des trade-off	16
C. Modéliser le branchement par la dynamique adaptative	17
V. Problème et limites de la modélisation de l'adaptation à un environnement saisonnier	18
VI. Modèle biologique	21
A. Cycle de vie de la levure en batch	22
B. Relation entre traits d'histoire de vie.....	23
C. Une évolution expérimentale de plusieurs souches, dans 4 environnements	26
VII. Questions abordées au cours de la thèse	28
Bibliographie :.....	29
Chapitre 1 : Modèles et inférence des paramètres.....	37
I. Modèle de la dynamique de population dans une culture en batch (modèle 1).....	37
A. Consommation des ressources et transition diauxique	38
B. Modélisation de la toxicité de l'éthanol.....	39
C. Modélisation de la mortalité	39
D. Equations du modèle.....	40
II. Calibration du modèle 1	40
A. Données de phénotypage	40
B. Données de compétition	41
C. Valeur des paramètres dans la littérature	42
D. Inférence par méthode ABC-SMC	45
III. Simulation de l'évolution expérimentale (Modèle 2)	51
Chapitre 2 : Boucles éco-évolutives entre l'environnement et les traits sous sélection dans un environnement saisonnier : conséquences pour l'évolution expérimentale.....	57

Chapitre 3 : Emergence de corrélation entre traits d’histoire de vie et métabolisme au cours de l’adaptation à un environnement saisonnier	71
I. Article	71
A. Introduction.....	71
B. Models and Methods	73
D. Results	79
E. Discussion	90
Reference:	93
Chapitre 4 : Caractérisation de souches dans un environnement saisonnier par leurs traits d’histoire de vie et leur fitness	101
I. Caractérisation des souches et de la dynamique adaptative selon leur environnement d’évolution	102
A. Caractérisation des stratégies d’histoire de vie et convergence phénotypique.....	102
II. Analyse de la valeur sélective des souches ancêtres et relation avec les traits d’histoire de vie	106
A. Matériel biologique	106
III. Conclusion du chapitre.....	112
Bibliographie :.....	113
Discussion	117
ANNEXES	129
SUPPLEMENTS DU CHAPITRE 2	130
SUPPLEMENTS DU CHAPITRE 3	144

RESUME EN FRANCAIS

L'adaptation des individus à un environnement dépend d'une combinaison de caractères adaptatifs, les traits d'histoire de vie, qui impactent la valeur sélective. Pour comprendre comment les organismes s'adaptent à leur environnement, on peut étudier quelles sont les traits composants la valeur sélective et comment ils dépendent de l'environnement biotique et abiotique. Au cours de cette thèse, je me suis intéressé aux composantes de la valeur sélective dans un environnement saisonnier et à ses conséquences sur la dynamique évolutive des traits quantitatifs.

Pour cela, j'ai utilisé une approche de modélisation mathématique d'une évolution expérimentale de l'espèce modèle *Saccharomyces cerevisiae* en cultures successives en batch. La levure de boulanger *S. cerevisiae* ici étudiée présente un cycle de vie respiro-fermentaire : en présence de glucose, elle le consomme par fermentation tout en produisant de l'éthanol, qui sera consommé dans un deuxième temps par respiration. Les souches de levures évoluent au cours de cycles successifs de fermentation-respiration. A intervalles de temps réguliers, des cellules sont transférées dans un nouveau milieu contenant du glucose où elles effectuent un nouveau cycle. J'ai développé un modèle mathématique d'équations différentielles pour étudier quels sont les traits sélectionnés dans les différentes saisons dans ce dispositif expérimental et comment l'environnement abiotique, l'environnement biotique et les relations entre les traits, impactent leur évolution.

Dans un premier temps, j'ai développé et paramétré un modèle d'équations différentielles décrivant la dynamique d'une population multi-souches au cours d'un batch (chapitre 1). J'ai ensuite proposé une décomposition de la valeur sélective et étudié quels traits sont sous sélection, et comment les pressions de sélection changent avec la composition de la population (chapitre 2). Deux types de traits sélectionnés ont pu être mis en évidence : les traits d'histoire de vie, liés au taux de croissance et à la mortalité, et les traits de transition, qui correspondent à la façon dont les souches réagissent aux changements de l'environnement. J'ai également montré que l'importance de chacune des composantes de la valeur sélective est liée à ces traits et à des traits non sélectionnés, via la longueur des différentes saisons. Au cours de l'évolution, ces composantes sont modifiées ce qui modifie la force de la sélection sur chaque trait. Ce phénomène de boucles de rétroaction éco-évolutives permet de mieux comprendre pourquoi la valeur sélective est fréquence-dépendante.

Dans un second temps, j'ai utilisé des simulations d'un modèle de dynamique adaptative pour montrer que l'existence d'un trade-off entre deux traits dans la population ancêtre pouvaient entraîner l'émergence d'autres relations entre un trait sélectionné et un trait non-sélectionné au cours de l'évolution (chapitre 3).

Enfin, pour mettre en regard les prédictions issues de modèles théoriques et des observations expérimentales, j'ai analysé deux jeux de données à travers le prisme de mon modèle mathématique (chapitre 4). Le premier jeu de données concerne le phénotypage de souches évoluées en batch successifs et leurs ancêtres. L'estimation des paramètres du modèle pour chacune des souches du jeu de données et leur analyse montrent que les traits liés à l'éthanol, sa consommation et sa production ont été principalement sélectionnés. Le second jeu de données, obtenu à partir de compétitions entre plusieurs couples de souches aillant des traits d'histoire de vie contrastés, a permis de mettre en évidence des différences de valeur sélective entre souches et de les relier avec des différences de traits phénotypiques, en cohérence avec les prédictions théoriques.

Mots clés : adaptation, environnement saisonnier, traits d'histoire de vie, composantes de la valeur sélective, modélisation mathématique, levure, *Saccharomyces cerevisiae*, boucles éco-évolutives

INTRODUCTION

Introduction

I. A propos de l'évolution

A. Pourquoi s'intéresse-t-on à l'évolution ?

Quel va être l'impact du réchauffement climatique sur le monde vivant en général ? Sur les rendements agricoles ou l'émergence de nouveaux pathogènes ? Comment préserver la diversité du monde vivant ? Comment lutter contre l'émergence de bactéries résistantes ? Comment sauvegarder des espèces en voie de disparition lorsque leur habitat est dégradé ? Comment utiliser et jouer sur la flore intestinale pour soigner des maladies chroniques ? L'évolution des espèces est au cœur de toutes ces questions largement médiatisées.

L'objectif de la biologie évolutive est de comprendre la diversité du monde vivant et son histoire. Comment expliquer l'existence d'un grand nombre d'espèces ? Comment, au sein d'une espèce, la diversité des individus joue sur le maintien de cette espèce ? Quel est le rôle de l'environnement sur l'évolution des espèces ? Pour expliquer ces grands mécanismes, la biologie évolutive fait appel à des facteurs reliés à la génétique des populations, l'écologie, l'environnement, la biologie du développement ou l'effet du hasard.

Au cours de ma thèse, je me suis intéressé au processus de l'adaptation à un environnement saisonnier et plus spécifiquement à la dynamique évolutive des traits quantitatifs caractérisant des souches de l'espèce modèle *Saccharomyces cerevisiae*. Cependant, avant d'exposer mes questions de recherche, je propose ci-dessous un panorama bref des questions et connaissances concernant l'adaptation à un environnement, et plus particulièrement à un environnement saisonnier.

B. Comment évoluent les populations ?

Au sein d'une population les individus sont différents, il existe des variations du phénotype. La génétique évolutive a pour but l'étude de l'évolution des fréquences de ces phénotypes dans la population. Le phénotype varie en fonction de deux facteurs : l'environnement et le

génotype. Le génotype est transmis de génération en génération, donc la part de la variance produite par le génotype est héritable, contrairement à celle produite par l'environnement. Si ces variations sont héritables, la génétique des populations nous indique que la composition phénotypique de la population peut évoluer au cours des générations successives sous l'action de 4 forces évolutives : d'une part la mutation et la migration qui peuvent créer ou faire arriver de nouveaux phénotypes dans une population, et d'autre part la sélection et la dérive qui peuvent faire évoluer les fréquences de ceux-ci (Henry and Gouyon 1999).

Les variations du génotype sont produites par mutation ou recombinaison. De nouveaux allèles ou de nouvelles combinaisons d'allèles peuvent apparaître dans la population, et produire de nouveaux phénotypes. De plus une population n'est pas un système clos et des allèles peuvent apparaître (ou disparaître) lorsque des individus d'une autre population y arrivent (ou lorsqu'une partie de ses individus quitte celle-ci). On appelle ce processus la migration. Ces deux processus font aussi varier les fréquences des génotypes, mais expliquent rarement la fixation d'un génotype.

L'évolution des fréquences et la fixation des génotypes peuvent s'expliquer par les deux autres forces : la sélection et la dérive.

La sélection est le processus selon lequel si, dans un environnement donné, un individu produit plus de descendants que les autres du fait de son phénotype, alors la fréquence de son génotype augmentera au fur et à mesure des générations. Cette sélection conduit à une adaptation de la population aux conditions environnementales dans laquelle elle évolue. L'adaptation est définie par (Stearns and Hoekstra 2000) comme « Un état qui a évolué car cela a permis d'augmenter la capacité relative à se reproduire ; l'adaptation désigne également le processus qui a produit cet état. »

Cependant tous les individus ne vont pas se reproduire d'une génération à l'autre. Les parents d'une génération sont échantillonnés aléatoirement dans le pool d'individus de la génération précédente. Dans une population de petite taille où les fluctuations d'échantillonnage peuvent avoir un impact important sur les fréquences, des variations aléatoires de la fréquence des gènes d'une génération à une autre peuvent avoir lieu, du seul fait du hasard.

Ainsi, par dérive, certains allèles peuvent se fixer ou disparaître de la population. Dans ce cas, il n'y a pas d'adaptation, et les phénotypes sont le fait du hasard.

L'importance relative de la sélection et de la dérive sur l'évolution des populations a fait l'objet de controverses au cours du 20^{ème} siècle. Dans la théorie proposée par Darwin en 1859 (Darwin 1859), la sélection naturelle est centrale. Mais, la théorie neutraliste de l'évolution, formalisée par Kimura dans les années 1960-70, explique une grande partie du polymorphisme observé dans les populations par l'émergence de mutations « neutres » (ayant peu d'effet sur la capacité des individus à se reproduire) et leur perte ou fixation par dérive (Stearns and Hoekstra 2000).

Mon travail de thèse porte plus spécifiquement sur la sélection, en particulier l'adaptation des populations à un environnement changeant. Pour cela, il faut notamment comprendre les variations du phénotype et l'effet des mutations.

II. Les variations du phénotype et les relations entre les traits

L'étude de ces variations est l'apanage de la génétique quantitative. Sans être exhaustif, je résume très brièvement ici quelques résultats utiles pour la suite.

Le taux d'apparition des mutations et leur effet peut être mesuré sur un ensemble de souches collectés dans différents environnements. On peut alors mesurer la diversité génotypiques et phénotypiques (Liti et al. 2009; Spor et al. 2009). Dans ce cas, on étudie des populations évoluées, dans lesquelles la sélection et la dérive ont éliminé une partie de la diversité que l'on voudrait aussi étudier. En laboratoire, on peut cependant réaliser des expériences d'évolution dans lesquels on peut augmenter le taux de mutation (en analysant des mutateurs ou en ajoutant des agents qui augmentent la fréquence des mutations) et limiter la sélection en ne repiquant qu'un seul individu à chaque transfert sur un nouveau milieu (Halligan and Keightley 2009). Dans ce cas, la nouvelle population est composée de clones de cet individu, même s'il est mal adapté. A l'exception des mutations délétères drastiques (létalité, stérilité), il est alors possible de mesurer le nombre et l'effet de toutes les mutations qui apparaissent, et non plus celui des mutations qui se fixent. Des études (Bataillon and Bailey 2014; Eyre-Walker and Keightley 2007; Kibota and Lynch 1996; Mukai 1964) ont montré que la majorité

des mutations sont délétères car elles réduisent le taux de reproduction des individus (par exemple (Wloch et al. 2001)). Les estimations sur les variations du taux de mutation et leur effet restent limitées par le nombre de lignées analysées et la difficulté de distinguer le nombre de mutations accumulées et leurs effets propres. Récemment, l'analyse par microfluidique de l'accumulation de 20 000 mutations sur des centaines de générations au niveau d'une cellule d'*Escherichia coli* a montré que la dynamique d'apparition des mutations suivait une loi de Poisson, que la majorité des mutations était quasi neutre, et la fraction des mutations délétères moins importante que précédemment estimé (Robert et al. 2018).

L'analyse de l'effet des mutations nous permet d'avoir une idée de la variance des différents traits, sous l'effet de la mutation. Cependant le phénotype est constitué de plusieurs traits, et des relations entre ses traits peuvent exister. Des études chez de nombreux microorganismes comme le plancton (Andersen et al. 2014) ou la levure (Spor et al. 2009; Warringer et al. 2011) (Tableau 1-2) ont mis en évidence des corrélations entre certains traits.

Les relations entre ces traits peuvent s'expliquer par différents mécanismes. Il est possible qu'elles soient dues à des mécanismes « intracellulaires » régis par des lois chimiques et physiques, à l'architecture génétique, ou à des processus évolutifs.

Différents mécanismes intracellulaires peuvent engendrer des relations entre traits. Le plus courant est l'allocation de ressources : les ressources dans la cellule étant limitées (énergie, métabolites, place...), si une cellule alloue une plus grande partie de ses ressources à une fonction, cela se fera au détriment d'autres fonctions. Par exemple, les ressources peuvent être allouées soit à la croissance en produisant les composantes de la biomasse, soit à la résistance aux toxines en produisant des lipides pour la membrane par exemple. Certaines relations entre traits s'expliquent par des contraintes physiques comme le premier principe de la thermodynamique qui relie négativement la vitesse d'une réaction et son rendement (Pfeiffer, Schuster, and Bonhoeffer 2001). Ces contraintes au niveau cellulaires pourraient expliquer des relations entre la vitesse de consommation des sucres et le rendement (Beardmore et al. 2011; Maharjan et al. 2013; Spor et al. 2009).

L'architecture génétique peut aussi engendrer des relations entre les traits. Si des gènes sont pléiotropiques, i.e. impliqués dans plusieurs fonctions, cela va entraîner une relation entre les traits affectés par ces fonctions. Si des traits sont gouvernés par des gènes en déséquilibre de liaison, tous les phénotypes ne pourront être produits s'il n'y a pas de reproduction sexuée et une relation entre les traits sera observée au niveau de la population. Ces relations sont plus

variables que celles causées par des mécanismes « intracellulaires ». En effet, il est possible que les relations disparaissent suite à des recombinaisons.

Enfin lorsque l'on observe des phénotypes, on observe uniquement ceux que la sélection n'a pas éliminés. Certains phénotypes sont toujours désavantagés et ne seront donc jamais observés. Par exemple, (Brodie 1992) une étude chez les *Thamnophis* a montré que le pattern rayé de la peau du serpent et sa façon de fuir sont reliés par une relation de ce type. Les serpents rayés qui ondulent beaucoup et les serpents peu rayés qui s'enfuient en ligne droite ont une meilleure fitness que les serpents rayés qui s'enfuient en ligne droite, ou les non-rayés qui ondulent. On observe donc une relation entre ces deux traits mais uniquement car la sélection « biaise » l'observation.

Afin de comprendre les causes des relations entre les traits, différentes méthodes peuvent être utilisées. La première est de supprimer les effets du déséquilibre de liaison et de la sélection. C'est ce qui est fait dans l'expérience d'accumulation de mutation présentées dans le premier paragraphe (Wloch et al. 2001). La deuxième façon de faire est de regarder les relations en prenant en compte la structure des populations et des espèces. En effet, les corrélations liées à la sélection ne seront observées que pour des populations ayant évolué dans des environnements proches, celles liées à la génétique seront observées pour des populations pour lesquels le déséquilibre de liaison n'a pas été rompu par de la recombinaison. Enfin les corrélations liées à des contraintes cellulaires seront toujours observées (Immler, Gonzalez-Voyer, and Birkhead 2012).

Une autre façon de comprendre ces relations est de modéliser l'évolution des traits pour comprendre quelles relations peuvent apparaître par sélection. C'est cette approche que j'ai utilisée au cours de ma thèse, pour comprendre notamment l'impact de ces relations sur l'évolution des traits et l'émergence de relations au cours de l'évolution.

III. L'adaptation et la valeur sélective

Différentes définitions de l'adaptation ont été proposées dans la littérature, on peut notamment reprendre la définition proposée par (Stearns and Hoekstra 2000) : « Un état qui a évolué car cela a permis d'augmenter la capacité relative à se reproduire ; l'adaptation désigne également le processus qui a produit cet état. ». L'adaptation est donc le résultat de

l'évolution, alors que la sélection est le mécanisme qui conduit à ce résultat. La « capacité relative à se reproduire » est souvent appelé fitness ou valeur sélective d'un individu.

Différentes mesures de la valeur sélective ont été proposées. De très nombreux débats existent autour de la fitness, notamment sur le fait que ce soit une mesure absolue ou relative, une espérance ou une mesure réalisée, si elle se définit au niveau de la population ou de l'individu (Thomas, Lefèvre, and Raymond 2016). On ne discutera pas ces définitions ici. Une des définitions la plus classique de la valeur sélective R_0 est le nombre moyen de descendants produits par un individu. On peut alors l'écrire comme la somme à chaque stade de développement i , de la probabilité de survivre jusqu'à ce stade et du nombre de descendants produits à ce stade de développement. Mathématiquement, cela se définit comme $R_0 = \sum_{i=0}^S \text{survie}_i \times \text{reproduction}_i$. Il existe d'autres définitions proches (Metcalf and Pavard 2007), selon le contexte et l'organisme étudié, mais l'idée générale reste identique. Cette définition est malheureusement peu utile pour les microorganismes. En effet, il est difficile de compter le nombre de descendant d'un seul individu. On utilise alors la définition de Fisher, qui est $W = (r - m)$ où r est le taux de croissance et m le taux de mortalité (Fisher 1958). Dans ces deux définitions, on retrouve la survie et la reproduction.

A. Evolution des traits d'histoire de vie

On appelle traits d'histoire de vie les traits associés directement à la reproduction et la survie comme la taille à la naissance, le taux de croissance, le nombre de descendants ou la durée de vie (Stearns and Hoekstra 2000). Ces traits vont donc influencer directement la fitness.

Si on ne considère que les traits d'histoire de vie, la sélection naturelle devrait amener à un accroissement de ces traits à l'infini : des individus qui ne meurent jamais, et se reproduisent très rapidement. De tels organismes, appelés « Démons Darwiniens », ne sont jamais observés car les relations entre les différents traits contraignent leur évolution (Roff and Fairbairn 2007). On observe alors différentes stratégies d'histoire de vie. Ces stratégies permettent de classer les individus, les deux plus utilisées étant « r » et « K » (Macarthur and Wilson 1967): la stratégie « r » privilégie la reproduction. Les individus ont donc un fort taux de reproduction mais une durée de vie moindre. A l'inverse la stratégie « K » privilégie la survie, les individus ont un taux de croissance r plus faible mais une durée de vie plus importante. Les individus « r » sont généralement des espèces colonisatrices, alors que les espèces « K » sont plutôt des

espèces compétitrices. Cette classification est aujourd'hui moins utilisée mais reste utile pour décrire le vivant (Reznick, Bryant, and Bashey 2002).

B. Paysages adaptatifs

Un paysage adaptatif est une représentation graphique ou mathématique qui associe à un phénotype donné une valeur de la valeur sélective. Les paysages adaptatifs sont souvent utilisés afin de prédire quelle stratégie est optimale. Dans un paysage adaptatif, le phénotype moyen de la population va se déplacer vers les « pics » du paysage, là où la valeur sélective est maximale. Ce type de modèle permet de visualiser simplement, l'effet de la variance des traits, et de prédire l'effet des mutations. Par exemple, on peut prédire que plus la population est adaptée, plus le nombre de mutations bénéfiques disponibles diminue. En effet plus la population se rapproche du sommet, plus la zone de mutation favorable diminue. De même, les effets des mutations diminuent également (Tenaillon 2014). Ce type de paysage a été proposé par Fisher. Il est utilisé pour prédire et réfléchir l'adaptation et inversement les données expérimentales viennent le confirmer et préciser. Par exemple, la longue évolution expérimentale mise en place par Richard Lenski (LTEE) a montré que le taux de croissance d'*Escherichia coli* augmente au cours du temps. Cette augmentation diminue au cours de l'évolution, comme prédit par le paysage adaptatif (Lenski et al. 1991).

Si le paysage possède plusieurs « pics », il est possible que la population ne soit pas « la plus adaptée possible », mais se retrouve bloquée à un maximum local. Dans ce cas, la dérive et les migrations jouent un rôle important, en permettant de passer d'un pic à l'autre. Ce type de paysage a été proposé par S. Wright. Le débat sur la forme générale des paysages adaptatifs reste encore aujourd'hui largement répandu parmi les généticiens.

Cependant, ce type de modèle suppose que l'on peut associer une valeur sélective fixe à chaque phénotype. Or ce n'est pas toujours possible, principalement si l'environnement varie.

IV. Le rôle de l'environnement dans la sélection

La valeur sélective d'un phénotype va dépendre de l'environnement. En effet, l'environnement définit quels traits influencent la survie et la reproduction et donc quels traits sont sélectionnés. Par exemple, si une toxine est présente dans l'environnement, la survie dépendra des traits liés à la résistance à cette toxine.

Des pressions de sélections différentes vont donc conduire à des phénotypes différents. Un des cas emblématiques est celui de la taille du bec chez les pinsons de Darwin (Grant and Grant 2006).

A. Effet des modifications de l'environnement

Si l'environnement change, les pressions qui vont s'exercer sur les individus changent, conduisant à l'adaptation de la population (ou à sa disparition). Dans la première évolution expérimentale, mise en place par Dallinger en 1890, la température de culture varie. Une bactérie est cultivée dans un environnement où la température augmente graduellement. En 5 ans, la population est passée d'une température maximale supportée de 21°C à 70°C. De même la salinité (Dhar et al. 2011) ou la présence de métaux lourds (Gorter et al. 2016) modifient les pressions de sélection.

Ces trois exemples sont des cas où une composante abiotique a été modifiée, mais les pressions de sélection qui s'exercent au sein d'un système peuvent aussi dépendre de la nature des interactions entre les individus d'une même espèce, avec d'autres espèces ainsi que de paramètres globaux comme la densité et les fréquences des différentes espèces. On parle alors de composante biotique de l'environnement.

Les interactions peuvent être de différentes natures, selon le bénéfice ou la perte obtenue par chacun des acteurs de la relation. La figure 1-1, tirée de (Faust and Raes 2012), montre les différentes interactions possibles.

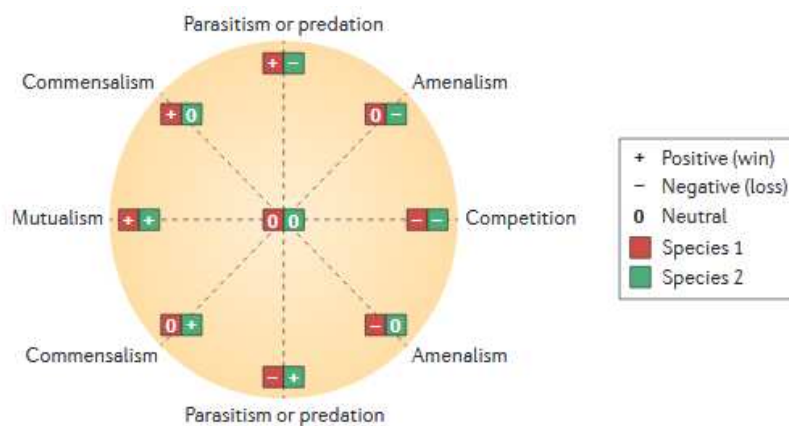


Figure 1 | **Summary of ecological interactions between members of different species.** The wheel display introduced by Lidicker⁴ has been adapted to summarize all possible pairwise interactions. For each interaction partner, there are three possible outcomes: positive (+), negative (-) and neutral (0). For instance, in parasitism, the parasite benefits from the relationship (+), whereas the host is harmed (-); this relationship is thus represented by the symbol pair +-.

Figure 1-1 : Différents type d'interaction trophiques (issue de (Faust and Raes 2012))

En modifiant ces relations, on modifie les pressions de sélections. Le cas le plus connu est celui des guppys (Auer et al. 2010). Des guppys n'ayant pas connu de prédation ont été introduits dans des mares de l'île de Trinidad contenant un prédateur. Ce prédateur se nourrit de jeunes guppys. En quelques générations leur taille à la reproduction a diminué et l'âge de la reproduction a été réduit. Le cycle de vie a été accéléré, et les guppys sont passés d'une stratégie plutôt « K » à une stratégie plutôt « r ». De même, après une introduction volontaire de lézards sur une île en Croatie, la population s'est adaptée et son régime alimentaire a été modifié : elle est en effet passé d'un régime insectivore à un régime herbivore. Ce changement s'est accompagné de nombreux changement morphologique, au niveau de la mâchoire ou du tube digestif (Herrel et al. 2008).

Dans ces exemples, la population s'adapte à une seule pression. Mais s'il y plusieurs ressources, les individus ne pourront pas nécessairement s'adapter à toute les conditions en même temps. Cela peut conduire à deux stratégies : généraliste ou spécialiste.

B. Devenir de la population (généraliste/spécialiste) selon la forme des trade-off

S'il existe plusieurs ressources, il y a deux possibilités d'évolution : soit un phénotype utilise les deux ressources sans préférence (un généraliste), soit un phénotype utilise préférentiellement une ressource (un spécialiste). Il a été montré théoriquement et expérimentalement que l'évolution vers l'un ou l'autre de ces comportements dépend de la forme trade-off entre l'utilisation de ses deux ressources. Si le trade-off est convexe, on observera un spécialiste, si le trade-off est concave, on observera un généraliste (Levins 1962; Maharjan et al. 2013). Ce résultat se généralise à d'autres types de trade-off entre traits d'histoire de vie et permet de prédire la forme de ceux-ci. C'est par exemple le cas du trade-off transmission-virulence chez les virus, dont la forme concave a été prédite par (Anderson and May 1979).

Il est aussi possible que la population passe d'un généraliste à deux spécialistes. Ce phénomène a été observé par (Rainey and Travisano 1998). Si l'on fait évoluer *Pseudomonas* dans un environnement non agité, un gradient de dioxygène va se créer. La bactérie va d'abord évoluer vers un phénotype généraliste, en se plaçant au centre du tube. Ainsi elle se développe dans les strates du tube riche en dioxygène (proche de la surface), et celles plus pauvre en dioxygène (au fond du tube). Ensuite deux sous-populations se séparent : une vers la surface, et donc se spécialisant vers le milieu aérobique, et l'autre vers le fond, donc le milieu sans oxygène. Ce phénomène a aussi été observé par (Frenkel et al. 2015), où les populations se spécialisent en fonction de la forme du tube : celle qui peuvent adhérer à la paroi, et celles qui ne peuvent pas mais qui résistent mieux à la compétition. Ce phénomène de passage d'un généraliste à deux spécialistes qui coexistent est appelé le branchement évolutif. Dans ces deux exemples, la séparation des niches est spatiale : chaque population est localisée à un endroit différent car les ressources sont réparties de manière hétérogène dans le milieu.

Cette séparation des niches peut aussi être temporelle. Dans ce cas les ressources s'alternent dans le temps, et l'environnement est dit saisonnier. Par exemple c'est le cas avec les planctons étudiés dans (Kremer and Klausmeier 2013). Dans cette expérience, des planctons sont placés dans un environnement qui alterne entre des phases de lumière où la croissance est possible et des zones d'ombre où la croissance n'est plus possible. La possibilité de

branchement dépend de la longueur des différentes saisons, ainsi que de la forme des relations entre l'utilisation des ressources (lumière/obscurité). Un autre exemple est celui observé par (Spencer et al. 2007). Dans ce cas, au terme de l'évolution de *E. Coli* dans un environnement dont le glucose est renouvelé toutes les 24 heures, deux souches coexistent : une souche L qui consomme le glucose et produit de l'acétate, et une souche S qui consomme l'acétate.

C. Modéliser le branchement par la dynamique adaptative

Dans un environnement qui n'est pas constant, la fitness instantanée d'un phénotype va varier au cours du temps. Le paysage adaptatif n'est donc plus suffisant pour prédire l'adaptation car il est mouvant. Il faut prendre en compte l'environnement, mais aussi les effets de l'évolution elle-même. En effet les composantes biotiques et abiotiques décrites ci-dessus ne sont pas indépendantes. L'environnement abiotique impose des pressions de sélection qui vont entraîner une adaptation des différentes espèces à l'environnement et une modification de la composition et de la densité de la population. Ces modifications peuvent à leur tour entraîner une modification de l'environnement abiotique (par exemple en consommant des ressources ou en modifiant l'acidité du milieu) ce qui modifiera les pressions de sélection. Ce phénomène est appelé « boucles de rétroaction éco-évolutive » (Post and Palkovacs 2009). Par exemple, ce sont les souches L qui produisent l'acétate qui est la deuxième ressource dans le système de (Spencer et al. 2007). Donc si ces souches évoluent, elles ne vont pas produire les ressources de la même façon, l'environnement va être modifié et donc la fitness également.

Afin de prendre en compte ces boucles éco-évolutives, des modèles de dynamiques adaptatives ont été introduits (Geritz et al. 1998).

La fitness est calculée pour chaque couple de souches possible, chacune devenant tour à tour mutante (avec une fréquence proche de 0) et résidente. La souche résidente impose l'environnement qui est supposé à l'état stationnaire et permet le calcul de la fitness. On représente alors la fitness pour chaque couple de souche et on obtient un graphique d'invasion par paire. Ils ont été classés par (Geritz et al. 1998).

Ce type de modèle distingue deux comportements pour chaque souche : lorsqu'elle est mutante : peut-elle envahir la population ? Et lorsqu'elle est résidente : peut-elle être envahie ?

Comme la fitness est modifiée, à cause des boucles de rétroaction éco-évolutives, il est possible qu'une souche puisse envahir la population et puisse être envahie à son tour. Dans ce cas, on peut observer un branchement évolutif. De même, il est possible que des stratégies qui ne peuvent pas être envahies, dites Stratégie Evolutivement Stable (ESS), ne puissent pas envahir et n'apparaissent donc jamais au cours de l'évolution. Cette situation est appelé « Jardin d'Eden ».

Ces modèles sont très utilisés et ont des comportements qui correspondent à un grand nombre de situations (Débarre, Nuismer, and Doebeli 2014; Kisdi 2015). Cependant, les résultats obtenus sont souvent numériques. Il est alors difficile de comprendre entièrement d'où vient cette valeur et de quoi elle dépend. Avoir une fonction de fitness ou des Stratégies Evolutivement Stable permettrait de mieux comprendre ce qui les influence. Ceci est notamment intéressant pour interpréter les résultats issus d'évolutions expérimentales, où l'on cherche à contrôler au maximum l'environnement. Ces difficultés sont explicitées dans le paragraphe suivant.

V. Problème et limites de la modélisation de l'adaptation à un environnement saisonnier

L'évolution expérimentale consiste à faire évoluer des organismes dans un environnement contrôlé afin d'étudier les processus à l'œuvre dans l'évolution de populations expérimentales soumises à des conditions choisies par l'expérimentateur (Kawecki et al. 2012). Il est donc nécessaire de décrire l'environnement correctement et d'identifier les pressions de sélection et les boucles de rétroaction éco-évolutives. Bien que des évolutions expérimentales aient été réalisées pour de nombreuses espèces, les microorganismes sont les plus utilisés pour plusieurs raisons comme la facilité de conservation, la vitesse de reproduction, l'importance des connaissances scientifiques. En effet, le stockage permet de comparer les phénotypes ancestraux et évolués en les mettant ensemble en compétition et la vitesse de reproduction permet d'accumuler plusieurs centaines de générations en quelques mois. C'est dans le contexte de l'évolution expérimentale de microorganismes que je place mes travaux de thèse et je vais donc me focaliser sur cela dans la suite de cette introduction et de mon manuscrit.

D'un point de vue pratique, pour les microorganismes, deux grands types de dispositifs expérimentaux sont utilisés principalement : les chemostats et les cultures en batch successifs (Barrick and Lenski 2013).

- Les batchs successifs

Un batch est un milieu de culture clos. Les ressources vont donc être consommées et s'épuiser. Au bout d'un certain temps, une fraction de la population est transférée dans un nouveau batch, initialement identique au précédent. La dynamique de la population est alors composée de deux phases : une phase de croissance sur le batch suivie d'une phase d'échantillonnage forte (un « bottleneck ») quand seulement une partie de la population est transférée au batch suivant. Les ressources et la population n'atteignent donc jamais un état stationnaire. Il existe des variantes de ces dispositifs, comme les fed-batch (il n'y a pas de transfert et le volume de culture augmente régulièrement avec l'ajout de milieu frais), mais nous ne les considérons pas ici.

- Le chemostat

Le chemostat, inventé par (Monod 1949), permet de garder une population en croissance à un taux constant. Les ressources sont renouvelées en continu par un système de circulation du milieu : un flux D de milieu neuf entre, et un flux D de milieu usagé sort. Comme les organismes sont aussi évacués, la population n'atteint jamais sa taille maximale, un équilibre s'établit rapidement, et la population atteint un état stationnaire (cet état stationnaire pouvant être la disparition de la population quand le flux D est trop important, on parle de « lavage »).

L'état stationnaire atteint par le chemostat permet d'utiliser les modèles de dynamiques adaptatives pour calculer le point d'équilibre (Doebeli 2002; Gresham and Hong 2015; Kremer and Klausmeier 2013). Cependant dans le cas du batch, l'absence d'état stationnaire limite les calculs et les résultats sont uniquement numériques (Doebeli 2002; Kremer and Klausmeier 2013). Il est donc plus difficile de comprendre exactement comment se passe l'adaptation au cours de batch successive. Ceci conduit à des prédictions différentes. Par exemple, dans le modèle de (Doebeli 2002), le branchement est plus fréquent en chemostat qu'en batch, alors que dans le modèle de (Rocabert et al. 2017), c'est l'inverse. Les hypothèses des deux modèles sont différentes, mais il est difficile de comprendre les différences sans pouvoir calculer

l'équilibre dans un batch. De même, il semble que le branchement ne soit possible que pour certaines formes de trade-off (Pluain et al. 2014), mais là encore ceci n'a été montré que par des simulations numériques.

Si l'évolution est difficile à modéliser, il en va de même pour la fitness. En pratique, pour les microorganismes on utilise donc une mesure de fitness relative, qui est $W = \frac{\lambda_m}{\lambda_s}$. Cette définition correspond au cas de population à croissance géométrique. Ainsi, au cours d'une génération, le nombre de mutants est multiplié par λ_m , et le nombre d'individus sauvages par λ_s . Ces deux paramètres sont les taux de croissance de chacune des populations dans ce modèle. On peut calculer l'augmentation relative de la fréquence de la souche mutante par rapport à la souche sauvage par $W = \frac{\lambda_m}{\lambda_s}$, qu'on appelle la valeur sélective relative. Si la valeur sélective est inférieure à 1, cela signifie que le mutant est moins adapté que le résident et la sélection conduira à son extinction. A l'inverse, si elle est supérieure à 1, c'est le résident qui s'éteindra. λ correspond au taux de croissance de la population, et la mortalité est souvent considéré comme étant négligeable. Cette mesure se rapproche donc de celle de Fisher, définit dans le paragraphe « Adaptation et valeur sélective ».

Cette mesure étant simple, elle est la plus utilisée pour mesurer la fitness dans les expérimentations en laboratoire et notamment en évolution expérimentale et cela même si la croissance de la population ne suit pas une suite géométrique. Elle reste cependant une approximation utile si l'environnement est simple. En testant cette mesure en modifiant le compétiteur ou la fréquence initiale des souches en compétition, (Wiser and Lenski 2015) a pu montrer que les résultats étaient cohérents. Cependant dans cet environnement la fréquence dépendance est faible. Si la fitness dépend fortement de la fréquence, comme on peut s'y attendre s'il y a des boucles de rétroaction éco-évolutives, cette mesure est problématique d'un point de vue théorique (Chevin 2011) et expérimental (Yi and Dean 2013). Dans l'expérience de (Yi and Dean 2013), *E. coli* est cultivée en batch, contenant alternativement deux antibiotiques. Dans cet environnement, deux souches coexistent chacune résistant à une des deux antibiotiques. Or la fitness n'est pas fréquence-dépendante, il est donc théoriquement impossible que les deux souches coexistent. La mesure de fitness n'est donc pas adaptée. De même, la fitness mesurée comme le rapport des taux de croissance est transitive : si la souche A bat la souche B et la souche B bat la souche C, alors la souche A

bat la souche C. Or il existe des cas expérimentaux où ce n'est pas le cas. Par exemple dans l'évolution expérimentale réalisée par (Paquin and Adams 1983) sur la levure, certaines souches évoluées C ont une fitness plus faible que les ancêtres A. Cependant, on observe bien pour une souche intermédiaire B que la souche C bat la souche B et la souche B bat la souche A. Il y a donc bien eu adaptation (chaque souche est plus adaptée que la précédente), mais non transitive.

Il semble donc que la fitness mesurée comme $W = \frac{\lambda_m}{\lambda_s}$ ne permette pas de comprendre tous les phénomènes observés en évolution expérimentale, notamment ceux liés à la fréquence-dépendance et aux boucles de rétroaction éco-évolutive. Le but de ma thèse est donc de proposer une mesure différente de la fitness, qui prenne en compte ces effets chez la levure *Saccharomyces cerevisiae*. A partir de cette mesure, je regarderai les conséquences sur la modélisation des dynamiques adaptatives et si cette mesure permet de mieux comprendre l'effet des trade-off, ainsi que la forme des équilibres évolutifs lors de batch successifs.

VI. Modèle biologique

Au cours de ma thèse, je me suis intéressée à l'évolution des traits d'histoire de vie de la levure de boulanger *Saccharomyces cerevisiae*, un champignon ascomycète unicellulaire. *S. cerevisiae* est utilisée dans de nombreux processus industriels, notamment pour sa capacité à produire de l'éthanol (production de vin, brasserie) ou du CO₂ (boulangerie). De plus, c'est un organisme modèle, son génome est séquencé et son cycle de vie est également connu (Kurtzman, Fell, and Boekhout 2011; Liti et al. 2009). Elle est capable de se reproduire de façon asexuée, ou lorsque les conditions environnementales sont défavorables, de façon sexuée. Lors du cycle sexué, la cellule diploïde produit des spores haploïdes par méiose. Ces spores peuvent ensuite fusionner pour former une nouvelle cellule diploïde par fécondation. Les cellules diploïdes peuvent également se reproduire de façon clonale par bourgeonnement. Dans le contexte expérimental que j'ai étudié au cours de ma thèse, il n'y avait *a priori* que des cycles asexués; les cellules filles sont donc identiques à la cellule mère, aux mutations près, et il n'y a pas de recombinaison possible.

A. Cycle de vie de la levure en batch

J'ai basé mon travail sur la dynamique évolutive de populations de levures dans des cultures en batch successives, en aérobie, où le glucose est la ressource limitante, tous les autres nutriments nécessaires à la croissance des levures (les sources d'azote assimilable, les lipides et les vitamines) étant en quantité suffisante tout au long de la culture. Dans ces conditions de culture, on observe au cours d'un batch plusieurs phases métaboliques qui se répètent à chaque batch formant un environnement saisonnier (voir paragraphe précédent IV.b.).

Tant que le glucose est suffisamment abondant, les cellules vont consommer le glucose par fermentation, même en présence d'oxygène, c'est l'effet Crabtree (de Deken 1966). Durant cette phase métabolique la production d'ATP utilisé pour la croissance des cellules s'accompagne principalement de la production d'éthanol et du CO₂. Une fois le glucose quasiment épuisé, les cellules passent en respiration et consomment l'éthanol pour continuer à croître. Cette réaction est plus lente que la fermentation, mais produit plus d'énergie (32 ATP pour la respiration contre 2 ATP pour la fermentation). Cette dynamique de la population et des ressources au cours d'une culture en batch est représentée sur la Figure 2-2.

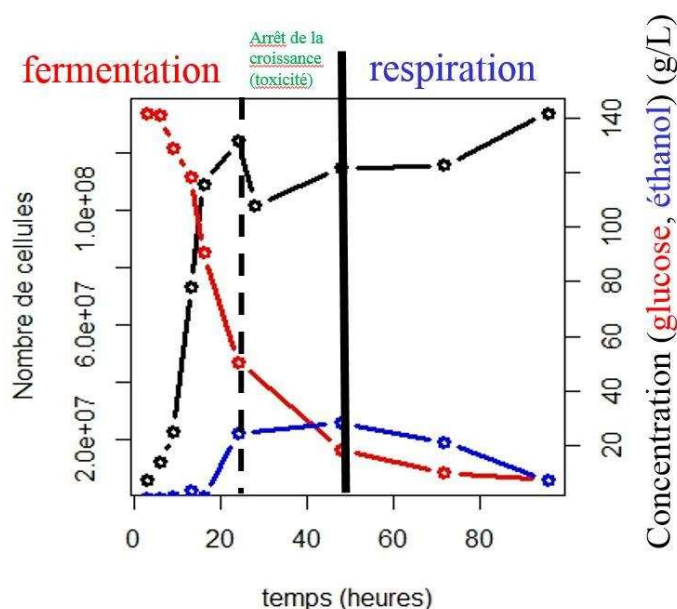


Figure 2-2 : dynamiques d'une souche de *S. cerevisiae* lors d'une culture en batch.

Si la concentration initiale en glucose est très élevée, la fermentation peut s'interrompre avant que le glucose ne soit entièrement consommé. Cet arrêt pourrait être lié à la toxicité des

coproduits de la fermentation comme l'éthanol ou l'acétate (Frohman, Widmer, and Orduña 2014).

B. Relation entre traits d'histoire de vie

La levure étant un organisme modèle, les relations entre plusieurs traits ont été étudiées. Ces relations sont résumées dans le tableau 1. On retrouve les relations classiques, comme celles entre le taux de mortalité et le taux de croissance, ou celle entre le taux de croissance et la tolérance à l'éthanol. On peut cependant remarquer que la relation entre le taux de croissance et le rendement change de sens, selon la manière dont elle est mesurée : Lorsqu'elle est mesurée sur plusieurs souches issues d'une même évolution expérimentale (Frohman, Widmer, and Orduña 2014; Jasmin, Dillon, and Zeyl 2012; Novak et al. 2006), la relation est négative. En revanche, si elle est mesurée sur plusieurs souches isolées d'environnements différents (Spor et al. 2009), la relation est positive.

En mesurant la taille des cellules et la taille de la population, on peut définir plusieurs stratégies d'histoire de vie (voir Figure 1-3). Ces stratégies ont été identifiées par (Spor et al. 2008). Deux stratégies extrêmes ont été définies : de grosses cellules qui atteignent des tailles de population faibles ou à l'inverse, des cellules plus petites atteignant des tailles de population plus importantes. Ces deux stratégies sont trouvées dans différents environnements, les « grosses cellules » sont plutôt trouvées dans les environnements riches en glucose plutôt anthropisés (industrie, clinique), alors que les « petites cellules » sont trouvées dans les environnements pauvres en glucose plutôt sauvages comme la forêt.

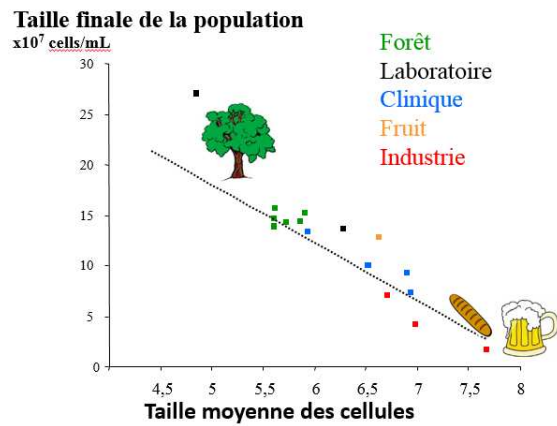


Figure 1-3 : relation entre la tailles de cellules et la taille de populations de 16 souches issues de différents milieu (données issues (Spor et al. 2008)). Les souches ont été phénotypées dans un même environnement batch.

article	traits...	...corrélé avec	sens de la corrélation	test trade-off
(Jasmin, Dillon, and Zeyl 2012)	taux de croissance en fermentation	rendement en biomasse de la fermentation	négative	souches évoluées à partir d'une souche
(Zakhartsev et al. 2015)			positif	une souche dans différents environnements (température)
(Novak et al. 2006)			négative	souches évoluées à partir de plusieurs souches (LTEE)
(Spor et al. 2009)			positif	souches différentes
(Elbing et al. 2004)	vitesse de consommation du glucose	taux de croissance en fermentation	positif	17 mutants
(Spor et al. 2009)	rendement en biomasse de la fermentation	vitesse de consommation du glucose	négative	souches différentes
(Elbing et al. 2004)			négative	17 mutants
(Zakrzewska et al. 2011)	taux de croissance en fermentation	tolérance à l'éthanol	négative	4825 mutants
(Antoce, Nămoșanu, and Radoi 2011)			négative	8 souches œnologiques

(Wright et al. 2011)	rendement en biomasse de la fermentation	production d'éthanol	négative	souches évoluées
(Wright et al. 2011)	vitesse de consommation du glucose	production d'éthanol	positif	souches évoluées
(Elbing et al. 2004)			positif	17 mutants
(Roop, Chang, and Brem 2016)	taux de croissance en respiration	transition entre fermentation et respiration	négative	souches différentes
(Roop, Chang, and Brem 2016)	rendement en biomasse de la respiration	transition entre fermentation et respiration	négative	souches différentes
(Maharjan et al. 2013)	rendement en biomasse de la fermentation	taux de mortalité	positif	une souche dans différents environnements
(Spor et al. 2009)	rendement en biomasse de la fermentation	taux de croissance en respiration	négative	souches différentes

Tableau 1-1 : Relations entre traits chez la levure documentées dans la littérature

C. Une évolution expérimentale de plusieurs souches, dans 4 environnements

Comme indiqué ci-dessus, Spor et al. ont montré que l'on peut identifier différentes stratégies d'histoire de vie chez la levure et que ces stratégies dépendent du milieu d'origine (Spor et al. 2009). Afin de tester si le contenu en glucose du milieu et le temps de survie pouvaient conduire à l'évolution vers l'une ou l'autre de ces stratégies et plus généralement afin de mieux comprendre l'évolution des corrélations entre les différents traits, une évolution expérimentale a été menée par A. Spor lors de sa thèse au laboratoire.

Dans cette expérience, six souches de *S. cerevisiae*, échantillonnées dans des environnements différents, ont été cultivées dans quatre conditions de batch par transferts successifs pendant 6 mois. Les quatre conditions diffèrent par la concentration initiale de glucose (10 g/L ou 150g/L) ainsi que la durée entre deux transferts (48 ou 96 heures). Pour chaque souche et chaque condition de culture, l'évolution est répétée trois fois indépendamment. Au total, 72 trajectoires évolutives ont donc été réalisées, chacune représentée par une souche ancêtre et une souche échantillonnée dans la culture finale et phénotypée au bout des six mois.

Les 72 souches évoluées ont été phénotypées pour 9 traits d'histoire de vie et quatre traits métaboliques et leur phénotype a été comparé à celui de leur ancêtre (Spor et al. 2014). Dans chaque environnement d'évolution, la population converge vers un phénotype multivarié différent. Les conditions où le glucose est le plus limitant (1% glucose et 96h de batch) sélectionnent pour une petite taille de cellule, une plus grande taille de population comme attendue pour la stratégie petite cellule trouvée pour les souches isolées d'environnement plus pauvre en sucre comme la forêt. Les taux de croissances liés à la fermentation et à la respiration ont augmenté, comme on pouvait s'y attendre. Cependant leur augmentation n'est pas différente d'un milieu de culture à l'autre. Ce sont donc d'autres traits qui sont sélectionnés spécifiquement dans chaque milieu. Le trait qui discrimine le mieux les souches est la taille, qui a servi à définir les stratégies. Les concentrations d'éthanol résiduel, qui sont liées à l'environnement et aux capacités de production/consommation d'éthanol des souches, permettent aussi de discriminer les souches évoluées selon les milieux. Il semblerait donc que l'environnement sélectionne différentes stratégies.

Pour mieux comprendre l'évolution des traits d'histoire de vie et le maintien et/ou l'émergence des corrélations au cours de l'évolution dans un environnement changeant, nous avons modélisé des dynamiques de population de levure dans le dispositif de cette expérience.

VII. Questions abordées au cours de la thèse

Le but de ma thèse était de modéliser l'évolution dans un environnement batch de manière utile et facilement interprétable afin :

- d'étudier l'impact d'un environnement changeant (ici modéliser par la succession de phase de fermentation/respiration de longueur variable) sur l'évolution des traits et de leur relations (chapitre 2, article 1 publié dans *Proceedings of the Royal Society B*)
- d'étudier l'impact des boucles de rétroaction éco-evolutives sur cette dynamique évolutive (chapitre 2, article 1 publié dans *Proceedings of the Royal Society B*)
- de comprendre comment les relations préexistantes entre traits (relations entre traits ancestrales) peuvent contraindre les trajectoires évolutives de chaque trait et leur corrélation (chapitre 3, article 2 en préparation)
- de comprendre dans quelles conditions des relations entre traits peuvent émerger au cours d'une trajectoire évolutive (chapitre 3, article 2 en préparation)

Pour cela, j'ai proposé une décomposition simple de la fitness, qui permet de prendre en compte plusieurs traits pour déterminer quels sont les traits sous sélection et de prendre en compte l'environnement pour comprendre son impact sur l'évolution des traits.

Suite à cette introduction, je rappellerai dans le chapitre 1 le protocole de l'évolution expérimentale effectué dans le cadre de la thèse d'Aymé Spor et je présenterai les modèles mathématiques que j'ai développé au cours de ma thèse : i) un modèle d'équations différentielles ordinaires qui décrit la dynamique de la population sur un batch puis, ii) un modèle permettant de simuler la dynamique éco-évolutive au cours de repiquages successifs. Je présenterai ensuite les données utilisées pour calibrer le modèle, les méthodes utilisées pour inférer les valeurs des paramètres, ainsi que l'ajustement obtenu entre les données et le modèle.

Dans le chapitre 2, nous analyserons une décomposition de la fitness que l'on peut faire à partir du modèle. Cette analyse nous permettra d'étudier trois éléments de l'adaptation : les traits sélectionnés, l'environnement et son influence sur les traits sélectionnés et enfin les

boucles de rétroaction éco-évolutives. Ce chapitre est un article accepté par *Proceedings of the Royal Society B* (Collot et al. 2018).

Dans le chapitre 3, je montrerai en utilisant le cadre de modélisation de la dynamique adaptative comment des relations ancestrales (préexistantes) entre des traits peuvent faire émerger des relations entre d'autres traits. Pour cela j'étudierai les stratégies optimales dans le système d'équations différentielles selon le type de contrainte ancestral et l'environnement.

Dans le chapitre 4, dans la perspective de confronter nos prédictions théoriques et des observations, j'analyserai les données obtenues lors de l'évolution expérimentale, ainsi que des données de compétition obtenues par l'équipe ACEP durant ma thèse. Je testerai quelles sont les différences phénotypiques entre les souches en fonction de leur milieu d'évolution. Les données de compétition permettront de mettre en évidence des différences entre les souches ancêtres de l'évolution expérimentale en analysant la fitness à différents moments de la dynamique.

Enfin, je finirai par une discussion.

Bibliographie :

- Andersen, Ken H. et al. 2014. "Modelling Emergent Trophic Strategies in Plankton." *Journal of Plankton Research* 37(5): 862–68.
- Anderson, Roy, and Robert May. 1979. "Population Biology of Infectious Diseases: Part I." *Nature* 280.
- Antoce, Arina Oana, Ioan Constantin Nămoșanu, and Florentina Matei Radoi. 2011. "Comparative Study Regarding the Ethanol Resistance of Some Yeast Strains Isolated from Romanian Vineyards." *Romanian Biotechnological Letters* 16(2): 5981–88.
- Auer, Sonya K., Jeffrey D. Arendt, Radhika Chandramouli, and David N. Reznick. 2010. "Juvenile Compensatory Growth Has Negative Consequences for Reproduction in Trinidadian Guppies (*Poecilia Reticulata*)." *Ecology Letters* 13(8): 998–1007.
- Barrick, Jeffrey, and Richard Lenski. 2013. "Genome Dynamics during Experimental Evolution." *Nat. Rev. Genet.* 14(12): 1301–15.

- Bataillon, Thomas, and Susan F. Bailey. 2014. "Effects of New Mutations on Fitness: Insights from Models and Data." *Annals of the New York Academy of Sciences* 1320(1): 76–92.
- Beardmore, Robert E, Ivana Gudelj, David a Lipson, and Laurence D Hurst. 2011. "Metabolic Trade-Offs and the Maintenance of the Fittest and the Flattest." *Nature* 472: 342–46.
- Brodie, Edmund D. 1992. "Correlational Selection for Color Pattern and Antipredator Behavior in the Garter Snake *Thamnophis Ordinoideis*." *Evolution* 46(5): 1284–98.
- Chevin, Luis-Miguel. 2011. "On Measuring Selection in Experimental Evolution." *Biology letters* 7: 210–13.
- Collot, Dorian et al. 2018. "Feedback between Environment and Traits under Selection in a Seasonal Environment : Consequences for Experimental Evolution." *Proceedings of the royal society B*.
- Darwin, Charles. 1859. *On the Origine of Species*.
- Débarre, F., S. L. Nuismer, and M. Doebeli. 2014. "Multidimensional (Co)Evolutionary Stability." *The American Naturalist* 184(2): 158–71.
<http://www.journals.uchicago.edu/doi/10.1086/677137>.
- de Deken, R. H. 1966. "The Crabtree Effect: A Regulatory System in Yeast." *Journal of General Microbiology* 44(2): 149–56.
<http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-44-2-149>.
- Dhar, R. et al. 2011. "Adaptation of *Saccharomyces Cerevisiae* to Saline Stress through Laboratory Evolution." *Journal of Evolutionary Biology* 24(5): 1135–53.
- Doebeli, M. 2002. "A Model for the Evolutionary Dynamics of Cross-Feeding Polymorphisms in Microorganisms." *Population Ecology* 44(2): 59–70.
- Elbing, Karin et al. 2004. "Role of Hexose Transport in Control of Glycolytic Flux in *Saccharomyces Cerevisiae*." *Applied and Environmental Microbiology* 70(9): 5323–30.
- Eyre-Walker, Adam, and Peter D. Keightley. 2007. "The Distribution of Fitness Effects of New Mutations." *Nature Reviews Genetics* 8(8): 610–18.
- Faust, Karoline, and Jeroen Raes. 2012. "Microbial Interactions: From Networks to Models." *Nature Reviews Microbiology* 10(8): 538–50.
- Fisher, R. 1958. 22 Dover Publication *The Genetical Theory of Natural Selection*.
- Frenkel, Evgeni M et al. 2015. "Crowded Growth Leads to the Spontaneous Evolution of Semistable Coexistence in Laboratory Yeast Populations." *Proceedings of the National*

- Academy of Sciences of the United States of America* 112(36): 11306–11.
- Frohman, Charles, Danielle Widmer, and Mira D E Orduña. 2014. “Amélioration de La Performance Des Levures Œnologiques Avec La Méthode Fed-Batch.” *Revue suisse Viticulture, Arboriculture, Horticulture* 46(3): 182–86.
- Geritz, S A H, Éva Kisdi, G. Meszén, and J.A.J. Metz. 1998. “Evolutionarily Singular Strategies and the Adaptive Growth and Branching of the Evolutionary Tree.” *Evolutionary Ecology* 12: 35–57.
- Gorter, Florian A., Mark M. G. Aarts, Bas J. Zwaan, and J. Arjan G. M. de Visser. 2016. “Dynamics of Adaptation in Experimental Yeast Populations Exposed to Gradual and Abrupt Change in Heavy Metal Concentration.” *The American Naturalist* 187(1): 110–19.
- Grant, P.R. Peter R P.R, and B.R. .R Rosemary Grant. 2006. “Evolution of Character Displacement in Darwin ’ s Finches.” *Science* 313(5784): 224–26.
<http://www.ncbi.nlm.nih.gov/pubmed/16840700>.
- Gresham, David, and Jungeui Hong. 2015. “The Functional Basis of Adaptive Evolution in Chemostats.” *FEMS Microbiology Reviews* 39(1): 2–16.
- Halligan, Daniel L., and Peter D. Keightley. 2009. “Spontaneous Mutation Accumulation Studies in Evolutionary Genetics.” *Annual Review of Ecology, Evolution, and Systematics* 40(1): 151–72.
<http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.39.110707.173437>.
- Henry, Jean-Pierre, and Pierre-Henri Gouyon. 1999. *Précis de Génétique Des Populations*. Dunod.
- Herrel, A. et al. 2008. “Rapid Large-Scale Evolutionary Divergence in Morphology and Performance Associated with Exploitation of a Different Dietary Resource.” *Proceedings of the National Academy of Sciences* 105(12): 4792–95.
<http://www.pnas.org/cgi/doi/10.1073/pnas.0711998105>.
- Immler, S., A. Gonzalez-Voyer, and T. R. Birkhead. 2012. “Distinct Evolutionary Patterns of Morphometric Sperm Traits in Passerine Birds.” *Proceedings of the Royal Society B: Biological Sciences* 279(1745): 4174–82.
<http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.1398>.
- Jasmin, Jean-Nicolas, Marcus Dillon, and Clifford Zeyl. 2012. “The Yield of Experimental Yeast Populations Declines during Selection.” *Proceedings. Biological sciences / The Royal*

- Society* 279: 4382–88.
- Kawecki, Tadeusz J et al. 2012. "Experimental Evolution." *Trends in Ecology & Evolution* 27(10): 547–60.
- Kibota, Travis T., and Michael Lynch. 1996. "Estimate of the Genomic Mutation Rate Deleterious to Overall Fitness in *E. Coli*." *Nature* 381(6584): 694–96.
- Kisdi, Éva. 2015. "Construction of Multiple Trade-Offs to Obtain Arbitrary Singularities of Adaptive Dynamics." *Journal of Mathematical Biology* 70(5): 1093–1117.
- Kremer, Colin T., and Christopher A. Klausmeier. 2013. "Coexistence in a Variable Environment: Eco-Evolutionary Perspectives." *Journal of Theoretical Biology* 339: 14–25.
- Kurtzman, Cletus P, Jack W Fell, and Teun Boekhout, eds. 2011. *The Yeasts (Fifth Edition)* The Yeasts. Fifth Edit. London: Elsevier.
<https://www.sciencedirect.com/science/article/pii/B9780444521491001804>.
- Lenski, Richard E, Michael R Rose, Suzanne C Simpson, and Scott C Tadler. 1991. "Long-Term Experimental Evolution in *Escherichia Coli* . I . Adaptation and Divergence During." *The American Naturalist* 138(6): 1315–41.
- Levins, Richard. 1962. "Theory of Fitness in a Heterogeneous Environment. I. The Fitness Set and Adaptive Function." *The American Naturalist* 96(891): 361.
- Liti, Gianni et al. 2009. "Population Genomics of Domestic and Wild Yeasts." *Nature* 458(7236): 337–41.
- Macarthur, Robert H, and Edward O Wilson. 1967. *The Theory of Island Biogeography*. REV- Revi. Princeton University Press. <http://www.jstor.org/stable/j.ctt19cc1t2>.
- Maharjan, Ram et al. 2013. "The Form of a Trade-off Determines the Response to Competition." *Ecology Letters* 16(10): 1267–76.
- Metcalf, C. Jessica E., and Samuel Pavard. 2007. "Why Evolutionary Biologists Should Be Demographers." *Trends in Ecology and Evolution* 22(4): 205–12.
- Monod, J. 1949. "The Growth of Bacterial Cultures." *Annual Review of Microbiology* 3: 371–94.
- Mukai, T. 1964. "Genetic Structure of Natural Populations of *Drosophila Melanogaster*.v1. Spontaneous Mutation Rate of Polygene Controlling Viability." *Genetics* 50(1): 1–.
- Novak, Maja et al. 2006. "Experimental Tests for an Evolutionary Trade-Off between Growth Rate and Yield in *E. Coli*." *The American Naturalist* 168(2): 242–51.

- Paquin, Charlotte E, and Julian Adams. 1983. "Relative Fitness Can Decrease in Evolving Asexual Populations of *S. Cerevisiae*." *Nature* 306: 368–71.
- Pfeiffer, T, S Schuster, and S Bonhoeffer. 2001. "Cooperation and Competition in the Evolution of ATP-Producing Pathways." *Science (New York, N.Y.)* 292(5516): 504–7. <http://www.ncbi.nlm.nih.gov/pubmed/11283355> (February 19, 2014).
- Pluain, Jessica et al. 2014. "Epistasis and Allele Specificity in the Emergence of a Stable Polymorphism in *Escherichia Coli*." *Science* 343(6177): 1366–69.
- Post, D. M., and E. P. Palkovacs. 2009. "Eco-Evolutionary Feedbacks in Community and Ecosystem Ecology: Interactions between the Ecological Theatre and the Evolutionary Play." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1523): 1629–40.
- Rainey, Paul B., and Michael Travisano. 1998. "Adaptive Radiation in a Heterogeneous Environment." *Nature* 394(6688): 69–72.
- Reznick, David, Michael J Bryant, and Farrah Bashey. 2002. "R- and K- Selection Revisited: The Role of Population Regulation in Life-History Evolution." *Ecology* 83(6): 1509–20. [http://www.esajournals.org/doi/abs/10.1890/0012-9658\(2002\)083\[1509:RAKSRT\]2.0.CO;2](http://www.esajournals.org/doi/abs/10.1890/0012-9658(2002)083[1509:RAKSRT]2.0.CO;2).
- Robert, Lydia et al. 2018. "Mutation Dynamics and Fitness Effects Followed in Single Cells." *Science* 359(6381): 1283–86.
- Rocabert, Charles et al. 2017. 13 *PLoS Computational Biology Beware Batch Culture: Seasonality and Niche Construction Predicted to Favor Bacterial Adaptive Diversification*.
- Roff, Derek A., and D. J. Fairbairn. 2007. "The Evolution of Trade-Offs: Where Are We?" *Journal of Evolutionary Biology* 20(2): 433–47.
- Roop, Jeremy I., Kyu Chul Chang, and Rachel B. Brem. 2016. "Polygenic Evolution of a Sugar Specialization Trade-off in Yeast." *Nature* 530(7590): 336–39. <http://www.nature.com/doi/abs/10.1038/nature16938>.
- Spencer, Christine C., Gerda Saxer, Michael Travisano, and Michael Doebeli. 2007. "Seasonal Resource Oscillations Maintain Diversity in Bacterial Microcosms." *Evolutionary Ecology Research* 9(5): 775–87.
- Spor, Aymé et al. 2008. "'Ant' and 'Grasshopper' Life-History Strategies in *Saccharomyces Cerevisiae*." *PloS one* 3(2).

- . 2009. “Niche-Driven Evolution of Metabolic and Life-History Strategies in Natural and Domesticated Populations of *Saccharomyces Cerevisiae*.” *BMC evolutionary biology* 9: 296.
- . 2014. “Phenotypic and Genotypic Convergences Are Influenced by Historical Contingency and Environment in Yeast.” *Evolution* 68(3): 772–90.
- Stearns, S C, and R F Hoekstra. 2000. *Evolution, an Introduction*. 36, , : Oxford University Press.
- Tenaillon, Olivier. 2014. “The Utility of Fisher ’ s Geometric Model in Evolutionary Genetics.” *Annual Review of Ecology, Evolution, and Systematics*: 179–201.
- Thomas, F, T Lefèvre, and M Raymond, eds. 2016. *Biologie Évolutive*. De Boeck.
<http://www.documentation.ird.fr/hor/fdi:010050890>.
- Warringer, Jonas et al. 2011. “Trait Variation in Yeast Is Defined by Population History.” *Plos Genetics* 7(6).
- Wiser, Michael J, and Richard E Lenski. 2015. “A Comparison of Methods to Measure Fitness in *Escherichia Coli*.” *PLoS ONE* 10(5): 1–11.
- Wloch, Dominika M, Krzysztof Szafraniec, Rhona H Borts, and Ryszard Korona. 2001. “Direct Estimate of the Mutation Rate and the Distribution of Fitness Effects in the Yeast *Saccharomyces Cerevisiae*.” *Genetics* 159: 441–52.
- Wright, Jeremiah et al. 2011. “Batch and Continuous Culture-Based Selection Strategies for Acetic Acid Tolerance in Xylose-Fermenting *Saccharomyces Cerevisiae*.” *FEMS Yeast Research* 11(3): 299–306.
- Yi, Xiao, and Antony M Dean. 2013. “Bounded Population Sizes, Fluctuating Selection and the Tempo and Mode of Coexistence.” *Proc. Natl. Acad. Sci. USA* 110(42): 16945–50.
- Zakhartsev, Maksim, Xuelian Yang, Matthias Reuss, and Hans Otto Pörtner. 2015. “Metabolic Efficiency in Yeast *Saccharomyces Cerevisiae* in Relation to Temperature Dependent Growth and Biomass Yield.” *Journal of Thermal Biology* 52: 117–29.
<http://dx.doi.org/10.1016/j.jtherbio.2015.05.008>.
- Zakrzewska, A. et al. 2011. “Genome-Wide Analysis of Yeast Stress Survival and Tolerance Acquisition to Analyze the Central Trade-off between Growth Rate and Cellular Robustness.” *Molecular Biology of the Cell* 22(22): 4435–46.
<http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E10-08-0721>.

CHAPITRE 1

Modèles et inférence des paramètres

Chapitre 1 : Modèles et inférence des paramètres

L'objectif de ce chapitre est de présenter le modèle que nous avons construit pour représenter la dynamique de *Saccharomyces cerevisiae* lors d'une succession de cultures en batch. Je présenterais également la méthode utilisée pour ajuster les données. Je commenterais ensuite les distributions de traits obtenus par cet ajustement. Enfin je présenterais rapidement la version simplifiée du modèle que nous avons utilisée pour obtenir certains résultats théoriques.

Le modèle est constitué de deux composantes. La première représente la dynamique de population de plusieurs souches sur un batch, et correspond à un système d'équations différentielles ordinaires. Pour paramétrer le modèle, nous avons ajusté une version monosouche de ce modèle à 72 monocultures différentes correspondant à 72 souches différentes. La deuxième composante correspond à des simulations d'évolution dans des batchs successifs où peuvent apparaître de nouveaux mutants qui seront ou ne seront pas sélectionnés.

I. Modèle de la dynamique de population dans une culture en batch (modèle 1)

Le modèle décrit ci-dessous a été adapté de différents modèles proposés précédemment dans la littérature (Lencastre Fernandes et al. 2012; MacLean and Gudelj 2006). Il modélise la dynamique temporelle de la concentration de deux métabolites, le glucose $G(t)$ (en g/mL, t indique le temps en heures) et l'éthanol $E(t)$ (en g/mL), ainsi que la dynamique de la concentration de cellules de chaque souche i , $N_i(t)$ (en cellules/mL).

A $t=0$, on introduit dans le milieu de culture contenant une concentration de glucose G_0 et E_0 d'éthanol, une certaine quantité de cellule de chaque souche $N_i(0)$. Le nombre total de cellules introduites est N_0 .

A. Consommation des ressources et transition diauxique

On suppose que la ressource limitante pour la croissance de la population est le carbone. On rappelle que même en présence d'oxygène, s'il y a suffisamment de glucose dans le milieu, les levures fermentent, c'est l'effet Crabtree (de Deken 1966). Lorsque le glucose vient à manquer, les cellules adoptent un métabolisme respiratoire : c'est la transition diauxique.

La population de levure va d'abord consommer le glucose par fermentation à un taux maximal J_f (en g/cellule/h). Le taux de consommation dépend de la concentration en glucose suivant le modèle de Monod (Monod 1949): $J_f \frac{G}{G+K_f}$. K_f (en g/mL) représente la concentration de glucose pour laquelle le taux de consommation vaut la moitié de J_f . Dans la suite, K_f est appelé « affinité de la réaction de fermentation ». Cette réaction de fermentation va produire de l'éthanol avec un rendement p (en g d'éthanol/g de glucose), du CO₂ (non inclus dans le modèle) et permettre à la population de croître avec un rendement y_f (en cellules/g de glucose consommé). On peut donc modéliser la vitesse de croissance de la population de la souche i selon le modèle de Monod (Monod 1949) par : $r_f \frac{G}{G+K_f}$. Ici, le taux de croissance maximal de la population est noté $r_f = y_f \cdot J_f$ (en h⁻¹).

En respiration, on peut partir du même modèle : l'éthanol est consommé à une vitesse maximale J_r (en g/cellules/h), et cette réaction permet à la population de croître à un taux de croissance maximale r_r (en h⁻¹), et on note K_r (en g/mL) l'affinité de la respiration et y_r le rendement en cellules par gramme d'éthanol consommé. Pour représenter l'inhibition de la respiration par le glucose, nous avons fait le choix de modélisation proposé par (Lencastre Fernandes et al. 2012; Sonnleitnert and Käppli 1986) où la vitesse de consommation de l'éthanol devient $r_r \frac{E}{E+K_r} \frac{K_c}{G+K_c}$. L'inhibition est liée à un seul paramètre K_c , qui est la concentration de glucose pour laquelle le taux de croissance en respiration vaut la moitié du taux de croissance maximal, appelé « constante d'inhibition » dans la suite. Une autre façon de modéliser l'inhibition est de modifier l'affinité. Le taux de croissance devient alors $r_r \frac{E}{E+K_r} \frac{K_c}{G+K_c}$. Dans nos données, nous n'avons trouvé aucune information qui aurait pu nous orienter vers l'une ou l'autre forme. La première étant plus simple à écrire, c'est celle-ci qui sera utilisée.

B. Modélisation de la toxicité de l'éthanol

Lorsque la concentration en éthanol devient trop importante, celui-ci a un effet toxique sur les cellules. Dans (Lima-Costa et al. 2012) , il est montré que l'éthanol réduit d'abord le taux de croissance de la population, puis pour de très fortes concentrations, la vitesse de consommation du glucose. L'éthanol n'induirait cependant pas de mortalité des cellules. Dans (van Uden 1985), huit formules sont proposées pour prendre en compte l'effet de l'éthanol. J'ai choisi la formule proposée par (Aiba, Shoda, and Nagatani 1969) : où E_m est la tolérance à l'éthanol (en g/mL). Cette formule a peu de paramètre afin de simplifier l'estimation de ceux-ci et est définie pour toutes les valeurs de concentration sans prendre de valeur négative.

La toxicité de l'éthanol est contestable biologiquement, bien que très souvent évoquée dans la littérature. En effet, l'éthanol ne serait pas le principal responsable de l'arrêt de la croissance. Une étude a montré que les cellules sont plus résistantes à de l'éthanol ajouté par l'expérimentateur qu'à l'éthanol qu'elles produisent (Hoppe and Hansford 1982). Une explication possible est que des coproduits de la fermentation, autres que l'éthanol, soient responsables de l'arrêt de la croissance, et notamment l'acétate (Frohman, Widmer, and Orduña 2014). De plus, la toxicité de l'éthanol dépend aussi des conditions du milieu comme le pH ou la température (Verduyn 1991). Cependant, l'éthanol est le seul métabolite secondaire mesuré dans nos expériences. Ainsi, nous avons choisi d'exprimer la toxicité en fonction de la concentration d'éthanol mais ces résultats doivent être interprétés avec précaution.

C. Modélisation de la mortalité

D'après les données observées au cours des monocultures de 72 souches différentes, la mortalité des cellules au cours du batch est importante. En effet, lorsque la concentration initiale de glucose est de 15%, environ 50% des cellules sont mortes après 48h, et 99% après 96h de culture. Pour des questions de simplicité, dans le modèle, j'ai considéré un taux de mortalité, m , constant au cours d'un batch. Ce choix a notamment été guidé par le fait que nous n'avons pas réussi à identifier les raisons de cette mortalité. En effet, nous n'avons trouvé dans la littérature aucun cas où une mortalité aussi forte est observée. On suppose

qu'elle n'est pas directement reliée à l'éthanol. Dans (Jones 1989), une dose d'éthanol inférieur à 100g/L n'est pas létal, et dans (Frohman, Widmer, and Orduña 2014; Lima-Costa et al. 2012) aucune mortalité n'est observée alors que les concentrations d'éthanol sont similaires aux conditions expérimentales de phénotypage des souches que j'ai étudiées (Spor et al. 2014).

D. Equations du modèle

La dynamique de la population de n souches de levure est décrite par un système de $2+n$ équations différentielles ordinaires : une pour la concentration en glucose $G(t)$, une pour la concentration en éthanol $E(t)$ et une pour chacune des concentration de souche de cellules $N_i(t)$ ($i=1,\dots,n$). Les souches n'interagissent que via l'environnement décrit par la concentration en éthanol et en glucose. Un batch contient initialement une concentration de glucose G_0 , une concentration de cellules N_0 , et une concentration nulle d'éthanol ($E_0=0$). La durée d'un batch est notée T (en heures).

$$\begin{aligned}\frac{dG}{dt} &= - \sum_i J_{f,i} \frac{G}{K_{f,i} + G} N_i \\ \frac{dN_i}{dt} &= r_{f,i} \frac{G}{K_{f,i} + G} \cdot \exp\left(\frac{-E}{E_{m,i}}\right) \cdot N_i + r_{r,i} \cdot \exp\left(\frac{-E}{E_{m,i}}\right) \cdot \frac{E}{K_{r,i} + E} \cdot \left(\frac{K_{C,i}}{K_{C,i} + G}\right) N_i - m_i \cdot N_i \\ \frac{dE}{dt} &= \sum_i J_{f,i} \frac{G}{K_{f,i} + G} p_i N_i - \sum_i J_{r,i} \frac{E}{K_{r,i} + E} \left(\frac{K_{C,i}}{K_{C,i} + G}\right) N_i \\ G(0) &= 0.15 \text{ g/mL}, \quad E(0) = 0 \text{ g/mL}, \quad \sum_i N_i(0) = 10^6 \text{ cells/mL}\end{aligned}$$

II. Calibration du modèle 1

A. Données de phénotypage

Le modèle a été ajusté sur les données obtenues lors de l'évolution expérimentale menée au laboratoire durant le doctorat d'Aymé Spor (Spor 2009; Spor et al. 2014). Lors de cette expérience, 6 souches de levure provenant de différents environnements ont évolué dans 4 environnements de cultures batch successives, différant par la concentration initiale de glucose et la fréquence des repiquages. Cinq des six souches ancêtres sont diploïdes. Seule la souche S288C, qui est une souche modèle de laboratoire, est haploïde. Pour chaque souche

et chaque environnement, trois évolutions indépendantes ont été menées. A la fin de cette expérience, chacune des 6 souches ancêtres et des 72 souches évoluées ont été cultivées en monoculture dans un batch. Chaque souche a été cultivée sur un milieu YNB 3% avec des acides aminés, contenant initialement 15% ou 1% de glucose, pendant 96 heures à 30°C sous agitation. La concentration initiale de cellules est d'environ 10^6 cell/mL. La concentration initiale d'éthanol est nulle. Les concentrations en éthanol, glucose et cellules ont été mesurées toutes les trois heures pendant les premières 24 heures, puis toutes les 24 heures. Au cours de ma thèse, j'ai utilisé les données de phénotypage dans le milieu à 15% de glucose. Pour ces mesures, les souches ont été réparties aléatoirement en cinq blocs d'expérience.

La concentration en cellules de la culture a été mesurée par un compteur Coulter à différents temps. Cette méthode mesure le nombre d'objets sphériques dans le milieu. Il est donc possible de compter des cellules mortes n'ayant pas éclatées. Afin d'estimer également la concentration de cellules vivantes, des comptages sur boîte ont été réalisés : après avoir estimé la concentration en cellules totales au Coulter, un volume de la culture diluée a été étalé sur une boîte de Pétri, et, après culture le nombre de colonies a été compté. On suppose que la dilution est suffisante pour que chaque colonie soit issue d'une seule cellule de la culture. En divisant le nombre de colonies formées par le nombre attendu de colonies (selon le comptage fait au Coulter), on peut estimer la proportion de cellules vivantes. Les étalements sur boîtes ont été effectués toutes les 24 heures.

B. Données de compétition

Ces données ont été obtenues par Carine Remoué de l'équipe Atelier Cartographie, Expression et Polymorphisme du Moulon. Afin de pouvoir distinguer les deux souches mises en compétition, les cinq souches ancêtres diploïdes ont été transformées avec un plasmide codant une protéine fluorescente rouge (comme indiqué dans (Raffoux et al. 2018)). Cette fluorescence peut être détectée grâce à un cytomètre de flux. Sur les cinq souches, seulement quatre ont gardé le plasmide. Chaque souche transformée a été mise en compétition avec les quatre non-transformées (y compris elle-même). Les 16 compétitions possibles entre souches colorées ou non ont été réalisées trois fois chacune. Les compétitions ont lieu dans un milieu 15% pendant 96 heures. Les deux souches sont introduites avec une concentration de 5.10^5

cellules/mL chacune. Le nombre de cellules a été mesuré au Coulter, et le pourcentage de cellules vivantes transformées, vivantes non transformées et mortes (colorées au Sytox) ont été comptées au cytomètre de flux.

Sur ces données, nous testerons l'effet potentiel du plasmide, ainsi que la dynamique de la fitness, mais nous ne ferons pas d'inférence. En effet, les concentrations de ressources n'ont pas été mesurées ce qui ne permet pas un ajustement correct du modèle aux données.

NB : On peut noter qu'une cellule est « vivante » si elle forme une colonie lorsque qu'elle est placée sur un nouveau milieu. Il est donc possible que des cellules en quiescence soient comptées comme mortes, si cette quiescence est suffisamment forte. Cependant, deux choses sont à noter :

- Si une cellule est quiescente, mais qu'elle ne redémarre pas son cycle sur un nouveau milieu, elle n'aura pas d'effet sur la dynamique de population. Elle peut donc être comptée comme morte.
- La mortalité mesurée par cytométrie en flux donne des résultats similaires. Cette mesure se base sur la polarité de la membrane cellulaire : si la membrane est dépolarisée, la cellule est considérée comme morte.

C. Valeur des paramètres dans la littérature

A partir des données de la littérature, j'ai estimé l'ordre de grandeur des paramètres du modèle (voir Tableau 1-1). Les paramètres de respiration ainsi que la mortalité sont rarement mesurés et le sont dans des conditions très éloignées des nôtres. Les paramètres de respiration ont donc été choisis du même ordre de grandeur que leur homologue en fermentation. Le taux de mortalité a été choisi du même ordre de grandeur que le taux de croissance. Ces valeurs, nous ont permis de proposer des distributions *a priori* des paramètres nécessaires aux algorithmes d'inférence présentés ci-dessous. Nous avons pris l'ordre de grandeur de ces valeurs, et les distributions choisies sont des lois uniformes sur ces intervalles. Les valeurs des paramètres dépendent beaucoup des modèles choisis par les auteurs. Les intervalles des distributions ont donc été choisis assez larges, afin de ne pas donner trop de poids à la distribution *a priori* lors de l'inférence.

Nom du trait	valeur	modèle	souche	référence
Vitesse de consommation du glucose J_f	$2,4.10^{-9}$ g.cell ⁻¹ .h ⁻¹	Modèle de (Sonnleitnert and Käppeli 1986)	CEN.PK 113-5D	(Lencastre Fernandes et al. 2012)
Vitesse de consommation du glucose J_f	$4,70.10^{-10}$ g.cell ⁻¹ .h ⁻¹	Modèle de Monod	Non indiqué	(Loveday and Winger 2007)
Vitesse de consommation du glucose J_f	$1,99.10^{-9}$ g.cell ⁻¹ .h ⁻¹	Modèle de Monod	<i>S. diastaticus</i>	(Wang and Sheu 2000)
Vitesse de consommation du glucose J_f	$1,04.10^{-10}$ g.cell ⁻¹ .h ⁻¹ $2,56.10^{-10}$ g.cell ⁻¹ .h ⁻¹	Estimation d'un taux de croissance avec une inhibition de l'éthanol	ATCC 4126	(Hoppe and Hansford 1982)
Vitesse de consommation du glucose J_f	$1,6.10^{-10}$ g.cell ⁻¹ .h ⁻¹ $2,52.10^{-10}$ g.cell ⁻¹ .h ⁻¹	Modèle de Monod	<i>K. marxianus</i>	(dos Santos et al. 2013)
Vitesse de consommation du glucose J_f	$1,12.10^{-19}$ g.cell-1.h-1	Mesure expérimentale	KOY.PK2-1C83	(Elbing et al. 2004)
Taux de croissance en fermentation r_f	Entre 0,1 et 0,4 h ⁻¹	Mesure expérimentale	Ensemble des <i>Saccharomycotina</i>	(Hagman et al. 2013)

Taux de croissance en fermentation r_f	Entre 0,1 et 0,4 h ⁻¹	Mesure expérimentale	20 souches de <i>Saccharomyces cerevisiae</i>	(Spor et al. 2009)
Taux de croissance en fermentation r_f	Entre 0,18 et 0,5 h ⁻¹	Mesure expérimentale	8 souches de <i>Saccharomyces cerevisiae</i>	(Antoce, Nămoșanu, and Radoi 2011)
Tolérance à l'éthanol E_m	Entre 5.10 ⁻³ et 55.10 ⁻³ g.mL ⁻¹	Fonction d'inhibition $K/(K+E)$	review	(Hoppe and Hansford 1982)
Tolérance à l'éthanol E_m	Entre 0,1 et 0,4 g.mL ⁻¹	Mesure expérimentale (diminution linéaire du taux de croissance)	8 souches de <i>Saccharomyces cerevisiae</i>	(Antoce, Nămoșanu, and Radoi 2011)
Constante d'inhibition K_c	1.10 ⁻⁴ g.mL ⁻¹	Modèle de (Sonnleitner and Käppli 1986)	CEN.PK 113-5D	(Lencastre Fernandes et al. 2012)
Rendement glucose/éthanol p	Entre 0,1 et 0,5	Mesure expérimentale	Ensemble des <i>Saccharomycotina</i>	(Hagman et al. 2013)
Affinité de la fermentation K_f	5.10 ⁻⁴ g.mL ⁻¹	Modèle de (Sonnleitner and Käppli 1986)	CEN.PK 113-5D	(Lencastre Fernandes et al. 2012)
Affinité de la fermentation K_f	6.10 ⁻⁴ g.mL ⁻¹	Modèle de Monod	Non indiqué	(MacLean and Gudelj 2006)

Table 1-1 : valeurs des paramètres trouvés dans la littérature.

D. Inférence par méthode ABC-SMC

1. Algorithme

En utilisant les données expérimentales, nous avons estimé 7 paramètres (J_f , r_f , m , E_m , p , J_r , r_r) du modèle ainsi que les concentrations initiales de cellules N_0 et de glucose G_0 pour 78 souches (les 72 évoluées et les 6 ancêtres). En effet, les mesures n'ont pas été effectuées à $t=0$. Or, la concentration en glucose est difficile à contrôler précisément expérimentalement car le glucose a tendance à cristalliser au bord du milieu de culture, diminuant possiblement la concentration de glucose dans le milieu. De même, la concentration en cellules était difficile à contrôler précisément à partir d'une mesure au Coulter avant la mise en culture, notamment en raison de la présence de cellules mortes. On va donc devoir estimer les concentrations initiales de glucose et de cellules.

Les trois paramètres restants (K_f , K_r and K_c) ont été fixés à 5.10^{-4} g/mL car il s'est révélé malheureusement impossible d'estimer correctement ces paramètres à partir des données disponibles. Nous avons vérifié que la valeur choisie pour ces paramètres n'affectait pas les estimations des autres paramètres.

Pour estimer les paramètres propres à chaque souche, j'ai ajusté la version monosouche du modèle aux données observées correspondantes. Pour cela, j'ai utilisé une version adaptée de l'algorithme ABC-SMC (Toni et al. 2009). Dans un algorithme ABC classique (Tavaré et al. 1997), les paramètres, tirés dans une loi *a priori*, sont utilisés pour simuler le modèle. Les prédictions sont ensuite comparées aux données. On accepte un jeu de paramètres si la distance entre les données et les prédictions est inférieure à un seuil ϵ fixé. Lorsque N_{obj} valeurs de paramètres ont été acceptées, l'algorithme s'arrête et les valeurs de paramètres définissent la loi *a posteriori*. Afin de limiter le nombre de calculs, on peut utiliser des distributions intermédiaires (méthode ABC-SMC, (Toni et al. 2009)) : l'algorithme commence avec un seuil élevé, et calcul une distribution *a posteriori*. Cette nouvelle distribution est utilisée comme distribution *a priori* pour une nouvelle itération j de l'algorithme avec un seuil

plus faible. Dans les algorithmes ABC, le seuil est fixé à l'avance. Dans la version adaptée de l'algorithme, le nouveau seuil est défini à partir des quantiles de la distribution précédente. L'algorithme s'arrête lorsque la variation du seuil entre deux itérations est inférieure à environ 5% (les distances sont ici de l'ordre de 1).

Les variables observées aux temps $t = 3, 6, 9, 12, 16, 24, 48, 72$ et 96 heures sont : $G(t)$, $Z_i(t)$, $E(t)$ où Z_i (en cellules/mL) est le nombre total de cellules produites par la souche i , observée au Coulter. De plus, aux temps $t=24, 48, 72$ et 96 heures, on observe $N(t)=Z(t).s(t)$ où $s(t)$ est estimée à partir des étalements sur boîte). La distance $d(X_{obs}, X_{pred})$ entre le modèle et les observations est la distance euclidienne entre les variables observées aux différents temps (indiquées par « obs ») et les prédictions du modèle (indiquées par « pred »), normalisée par les valeurs observées et le nombre d'observations (n_N, n_Z, n_E and n_G) :

$$d(X_{obs}, X_{pred}) = \frac{1}{n_N} \sum_{i=1}^{n_N} \left(\frac{N_{obs,i} - N_{pred,i}}{N_{obs,i}} \right)^2 + \frac{1}{n_Z} \sum_{i=1}^{n_Z} \left(\frac{Z_{obs,i} - Z_{pred,i}}{Z_{obs,i}} \right)^2 \\ + \frac{1}{n_E} \sum_{i=1}^{n_E} \left(\frac{E_{obs,i} - E_{pred,i}}{E_{obs,i}} \right)^2 + \frac{1}{n_G} \sum_{i=1}^{n_G} \left(\frac{G_{obs,i} - G_{pred,i}}{G_{obs,i}} \right)^2$$

L'algorithme est le suivant :

1. $j=1, \varepsilon_1=4$.
2. a. Echantillonnage d'un vecteur de paramètres θ dans la distribution *a priori* $\pi(\theta)$ décrite dans la table 2.
b. Si la distance $d(X_{obs}, X_{pred}(\theta)) < \varepsilon_1$, Le vecteur est ajouté à l'ensemble P_1 .
c. Si la taille de la population est inférieure à 10 000, retour à l'étape 2.a.
d. Calcul de ε_2 qui est le 36^{ème} percentile de $\{d(X_{obs}, X_{pred}(\theta)), \theta \in P_1\}$. Chaque vecteur a le même poids $\omega_1(\theta)=1/10\,000$. $j=2$.
3. a. Echantillonnage d'un vecteur de paramètres θ^* dans P_{j-1} pondérée par ω_{j-1} .
b. Perturbation de θ^* par un noyau K pour obtenir θ^{**} . Comme le choix du noyau a peu d'effet sur la vitesse de convergence (Toni et al. 2009), nous avons utilisé des distributions uniformes $\theta^{**} \sim U(0,9; 1,1) \theta^*$.
c. Si $d(X_{obs}, X_{pred}(\theta^{**})) < \varepsilon_j$ and $\pi(\theta^{**}) \neq 0$, le vecteur est ajouté à l'ensemble P_j .
d. Si la taille de la population est inférieure à 10 000, retour à l'étape 3.a.
e. Calcul de ε_{j+1} qui est le 36^{ème} percentile de $\{d(X_{obs}, X_{pred}(\theta^{**})), \theta^{**} \in P_j\}$.

f. Calcul des poids $\omega(\theta^{**}) = \frac{\pi(\theta^{**})}{\sum \omega(\theta_{j-1}^k) K(\theta^{**} | \theta_{j-1}^k)}$ avec θ_{j-1}^k les vecteurs de P_{j-1} .

Normalisation des poids.

g. Si $\varepsilon_{j+1} - \varepsilon_j > 0,05$, $j=j+1$ et retour à l'étape 3.a.

L'ensemble final définit la distribution *a posteriori* $\pi(\theta^{**} | d(X_{\text{obs}}, X_{\text{pred}}(\theta^{**})) < \varepsilon)$.

Pour choisir le pourcentage de réduction de l'intervalle à chaque itération, nous avons cherché celui qui minimiserait le nombre de simulations à faire. En cherchant à minimiser le nombre de calculs à faire, on trouve que la proportion de simulations à garder pour définir la nouvelle distance à chaque boucle est d'environ e^{-1} soit 37%.

Démonstration rapide de $p_{\text{boucle}} = e^{-1}$:

On veut garder N_{obj} simulations à chaque boucle et réduire la distribution a priori d'une proportion p_{obj} à la fin de toutes les itérations. A chaque boucle, on réduit la distribution d'une proportion p_{boucle} qui correspond à la proportion de simulations que l'on va garder (i.e. celle que l'on utilise pour définir la distance de l'itération suivante). On cherche à minimiser le nombre total de simulation faites $N_{\text{réelle}}$.

Par boucle, on a $N_{\text{boucle}} = N_{\text{obj}} / p_{\text{boucle}}$.

On souhaite que la taille X de l'intervalle soit réduite de p_{obj} , en le réduisant k fois d'un facteur p_{boucle} . On souhaite donc trouver k tel que $X \cdot p_{\text{obj}} = X(p_{\text{boucle}})^k$.

On va devoir faire $k = \ln(p_{\text{obj}}) / \ln(p_{\text{boucle}})$ boucles au total.

Donc le nombre de simulation est $N_{\text{réelle}} = N_{\text{obj}} / p_{\text{boucle}} \times \ln(p_{\text{obj}}) / \ln(p_{\text{boucle}})$. On trouve que $N_{\text{réelle}}$ est minimum pour $p_{\text{boucle}} = e^{-1} \approx 0.37$, indépendamment du nombre de simulations que l'on souhaite garder, ou de la réduction de distribution que l'on souhaite.

2. Qualité de l'ajustement

Afin de vérifier la qualité de l'ajustement, les résidus ont été calculés pour chaque paramètre, chaque souche et à chaque point d'observation (Figure 2-1). Les résidus ont été calculés pour chacun des 10 000 jeux de paramètres comme la différence entre la prédiction et la valeur observée standardisée par la valeur observée.

$$\text{res}_x = \frac{1}{n_x} \sum_{i=1}^x \text{abs} \left(\frac{X_{\text{obs},i} - X_{\text{pred},i}}{X_{\text{obs},i}} \right)$$

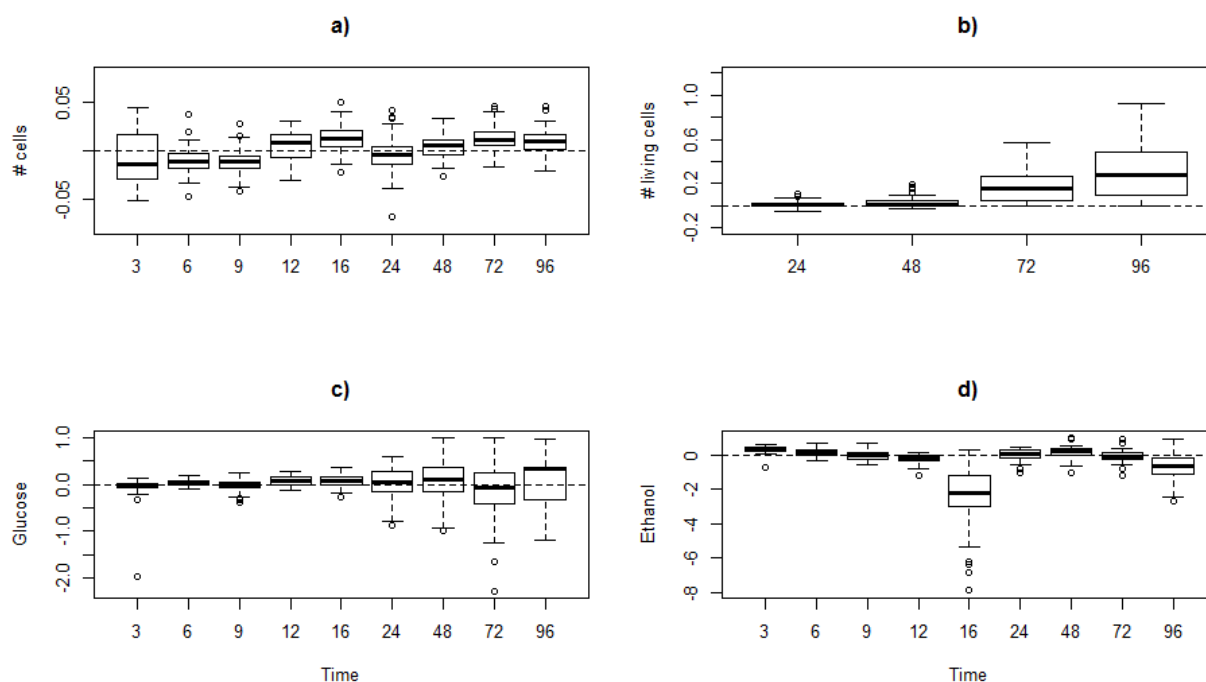


Figure 2-1 : Distribution des résidus en fonction du temps. Les résidus sont calculés comme la différence entre les prédictions du modèle et les données observées pour chaque souche et a) pour la concentration totale de cellules b) pour la concentration de cellules vivantes, c) pour la concentration en glucose, d) pour la concentration en éthanol.

A partir de cette figure, on peut repérer les points qui sont moins bien estimés. Si les résidus sont plus élevés à un temps qu'aux autres, cela nous indique qu'il y a un problème d'estimation. De même si les résidus ne sont pas centrés en 0, cela indique un biais systématique dans nos estimations. Il y a principalement trois types de points problématiques (Figure 2-2) :

- la concentration d'éthanol à 16h est toujours surestimée par les simulations. Cependant, dans les données observées, ce point est toujours très inférieur aux points précédents et suivants. Il ne correspond donc pas à l'allure générale de la courbe. Ce biais systématique pourrait donc être lié à une erreur expérimentale. En effet, le kit utilisé pour mesurer la concentration d'éthanol ne fonctionne que sur une certaine gamme de concentrations. Il est donc nécessaire de diluer l'échantillon pour rester dans cette gamme. Si cette dilution n'était pas assez importante à ce point, il est possible que l'on sorte de la gamme et que la mesure soit faussée.

- Les concentrations de glucose et d'éthanol ont également des résidus assez forts en fin de dynamique. L'éthanol est souvent surestimé alors que le glucose est sous-estimé. Ces biais pourraient être liés au fait que le modèle ne peut s'ajuster correctement aux deux concentrations en même temps. En effet dans le modèle, il ne peut y avoir consommation d'éthanol que si la concentration en glucose est presque nulle, ce qui n'est pas le cas pour les données observées : dans certaines dynamiques la concentration en glucose est encore relativement élevée lorsque les cellules passent en respiration. Dans ce cas le modèle va, soit prédire une concentration de glucose nulle et estimer correctement l'éthanol, soit prédire correctement le glucose, mais pas la consommation d'éthanol. Ceci pourrait être corrigé en choisissant une fonction plus compliquée pour la transition diauxique. Cependant, nous n'avons pas d'indice concernant l'éventuelle forme de celle-ci. Il est également possible que la consommation d'éthanol ne soit pas due aux cellules, mais à l'évaporation de celui-ci.
- Le nombre de cellules vivantes est sous-estimé à la fin de la dynamique. Cela peut s'expliquer par deux phénomènes. D'abord la mortalité n'est pas forcément constante. S'il y a un point où la mortalité est très forte, le paramètre m sera surestimé. La deuxième explication est liée à la transition diauxique : si elle ne se fait pas, le taux de croissance de la population reste à 0, et la taille de la population diminue.

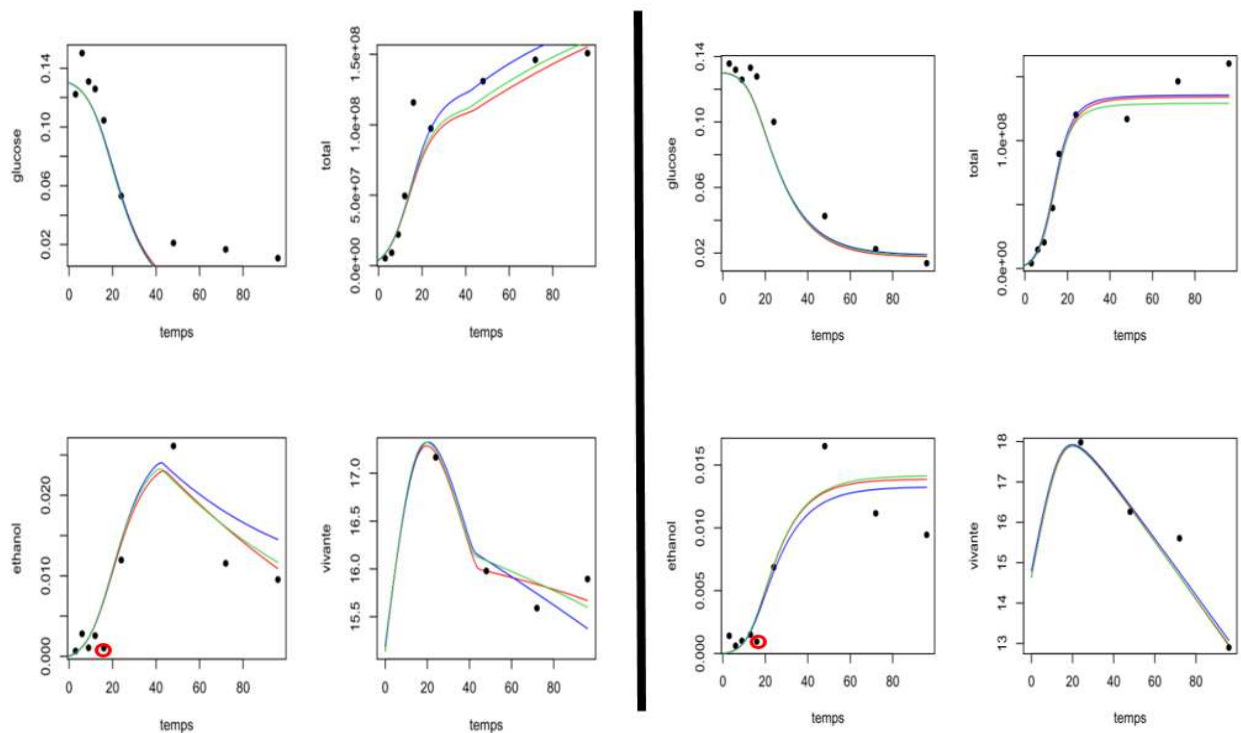


Figure 2-2 : Ajustement du modèle aux données pour deux souches évoluées. Les points représentent les données et les courbes représentent les simulations du modèle pour différentes valeurs de paramètres (en bleu, avec le jeu de paramètre donnant la distance la plus faible ; en vert, avec les paramètres moyens et en rouge, avec les paramètres médians). A gauche on ajuste l'éthanol mais pas le glucose, à droite le glucose mais pas l'éthanol la diminution d'éthanol. Le cercle rouge indique le point d'éthanol toujours plus bas que la prédiction.

Pour chaque souche, les paramètres sont estimés avec une certaine distribution. Les coefficients de variation des différents paramètres sont représentés sur la figure 2-3. Une estimation est satisfaisante si le coefficient de variation de la distribution du paramètre est plus faible dans la distribution *a posteriori* que dans la distribution *a priori*.

Le coefficient de variation des paramètres de respiration (J_r et r_r) est assez fort. Ils semblent donc assez difficiles à estimer. En effet, la consommation est rarement au maximum : si la concentration en éthanol est forte, la consommation est inhibée par l'effet toxique celui-ci, si la concentration en éthanol est faible, la consommation est diminuée à cause de l'affinité de

la réaction. De plus, l'inhibition liée au glucose altère également la vitesse. Il est donc assez difficile d'estimer la vitesse maximum puisqu'elle n'est jamais atteinte.

Le coefficient de variation du taux de croissance en fermentation diminue assez peu, car il était déjà faible dans nos distributions *a priori*. En effet, ce paramètre est souvent estimé dans la littérature. De plus ce paramètre est important pour la dynamique et va donc être fortement contraint, ce qui peut biaiser les estimations des autres paramètres. Les autres paramètres ont tous des coefficients de variations qui ont diminué.

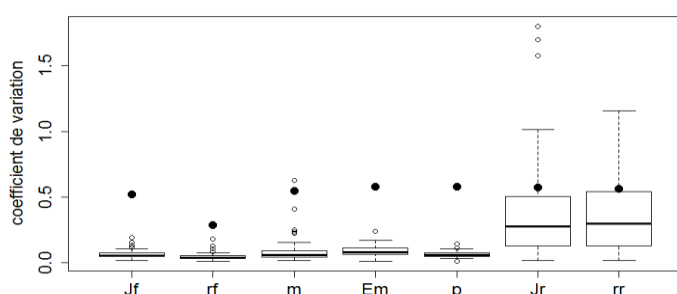


Figure 2-3: Distributions des coefficients de variation des différents traits. *Les coefficients de variation des distributions de chaque paramètre ont été calculés pour chacune des 72 souches évoluées. Les points noirs indiquent le coefficient de variation dans la distribution à priori.*

III. Simulation de l'évolution expérimentale (Modèle 2)

Le modèle 1 permettait de modéliser la compétition entre différentes souches de *S. cerevisiae* au cours d'un batch. Lors de l'évolution expérimentale menée au laboratoire lors de la thèse d'Aymé Spor (Spor 2009), les cultures en batch étaient repiquées dans une nouvelle culture batch à intervalle de temps régulier (48 heures ou 96 heures selon l'environnement). L'objectif du modèle 2 est de simuler cette expérience en utilisant le modèle 1 décrit ci-dessus de façon itérative. Pour cela, à la fin de chaque culture batch (après T heures, la durée entre deux transferts), les conditions du batch sont ramenées aux conditions initiales et pour simuler le repiquage de la population, un échantillon de $N_0=10^6$ cellules est tiré aléatoirement dans la population finale du batch précédent, constituée de différentes souches de levure. Le tirage est fait suivant une loi multinomiale de paramètres $(N_0, f_1, \dots, f_J)$ avec J le nombre de souches. Ainsi, la probabilité de tirer une cellule de la souche i est donnée par sa fréquence à la fin du

batch. Dans la population tirée aléatoirement, on choisit aléatoirement une cellule et on modifie son phénotype (un ou plusieurs des paramètres du modèle). On obtient alors une nouvelle souche mutante. On simule ensuite avec le modèle 1 une culture batch de cette nouvelle population dans un milieu contenant initialement G_0 g/mL de glucose et 0 g/mL d'éthanol.

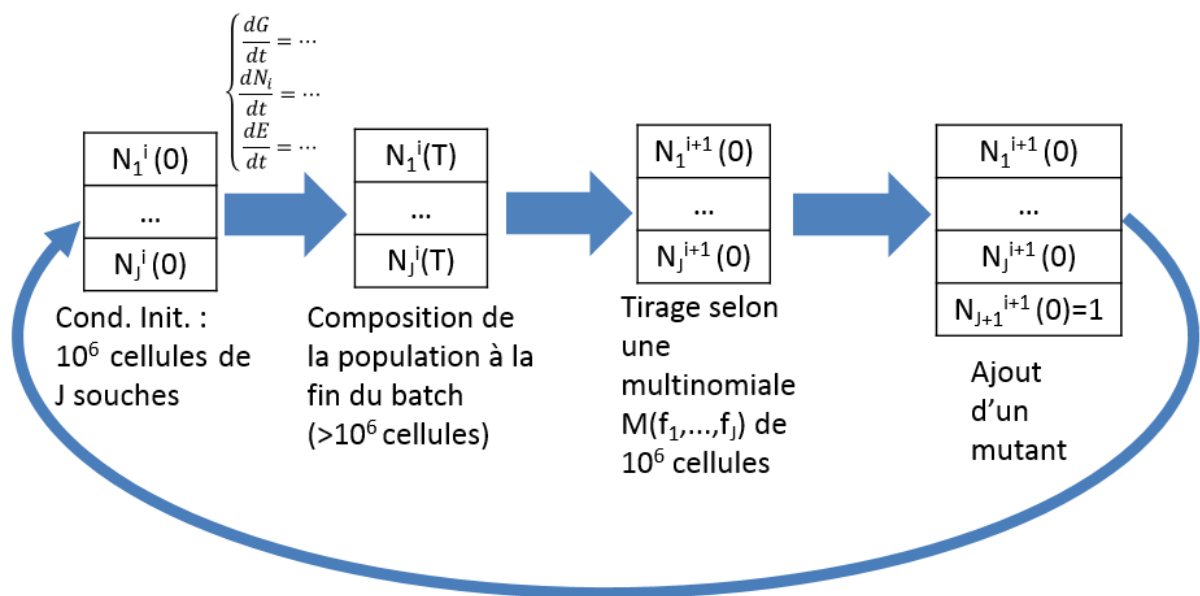


Figure 2-4 : Schéma de l'algorithme de modèle 2.

Deux hypothèses peuvent être discutées :

- Le tirage se fait dans une loi multinomiale. Ce tirage est justifié si la taille de la population finale est très grande par rapport à 10⁶, alors qu'un tirage uniforme sans remise des individus dans la population aurait été possible. Cette hypothèse a pour but de simplifier le calcul de la probabilité d'extinction $P_{\text{ext},i}=(1-f_{\text{fin},i})10^6$. Elle évite également un problème si la taille de la population est inférieure à 10⁶ cellules/mL au moment du repiquage. Cela peut se produire si la mortalité est assez forte, notamment en échantillonnant à 96 heures. Dans ce cas, on ne pourrait pas échantillonner N_0 cellules dans la population finale et l'algorithme bloquerait.
- Le mutant apparaît avec une concentration de 1 cellule/mL. Selon le volume, le mutant peut donc apparaître avec une taille de population forte. De plus, on ne fait apparaître qu'un mutant à la fin d'un batch. Ceci est lié au fait que les naissances ne sont pas prises en compte de manière explicite. On suppose donc qu'à chaque batch un seul

mutant atteint le seuil de 1 cellule/mL. Comme le mutant atteint 1 cellule/mL indépendamment de sa fitness, cela revient à supposer que la dérive est beaucoup plus importante que la sélection pour des tailles de population très faibles.

Ce modèle a été utilisé dans l'article présenté au chapitre 3.

Bibliographie

Aiba, S., M. Shoda, and M. Nagatani. 1969. "Kinetics of Product Inhibition in Alcohol Fermentation-Temperature Effect in the 'Sake' Brewing." *Biotechnology and bioengineering* 11: 1285–87.

Antoce, Arina Oana, Ioan Constantin Nămoșanu, and Florentina Matei Radoi. 2011. "Comparative Study Regarding the Ethanol Resistance of Some Yeast Strains Isolated from Romanian Vineyards." *Romanian Biotechnological Letters* 16(2): 5981–88.

de Deken, R. H. 1966. "The Crabtree Effect: A Regulatory System in Yeast." *Journal of General Microbiology* 44(2): 149–56.
<http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-44-2-149>.

Elbing, Karin et al. 2004. "Role of Hexose Transport in Control of Glycolytic Flux in *Saccharomyces Cerevisiae*." *Applied and Environmental Microbiology* 70(9): 5323–30.

Frohman, Charles, Danielle Widmer, and Mira D E Orduña. 2014. "Amélioration de La Performance Des Levures Œnologiques Avec La Méthode Fed-Batch." *Revue suisse Viticulture, Arboriculture, Horticulture* 46(3): 182–86.

Hagman, Arne, Torbjörn Säll, Concetta Compagno, and Jure Piskur. 2013. "Yeast 'Make-Accumulate-Consume' Life Strategy Evolved as a Multi-Step Process That Predates the Whole Genome Duplication." *PloS one* 8(7).

Hoppe, GK, and GS Hansford. 1982. "Ethanol Inhibition of Continuous Anaerobic Yeast Growth." *Biotechnology Letters* 4(1): 39–44.

Jones, RP. 1989. "Biological Principles for the Effects of Ethanol." *Enzyme and Microbial Technology* 11(June 1988). <http://www.sciencedirect.com/science/article/pii/0141022989900732> (February 12, 2014).

Lencastre Fernandes, Rita et al. 2012. "Cell Mass and Cell Cycle Dynamics of an Asynchronous Budding Yeast Population: Experimental Observations, Flow Cytometry Data Analysis, and Multi-Scale Modeling." *Biotechnology and bioengineering* 110(3): 812–26.

Lima-Costa, Maria Emília et al. 2012. "Kinetics of Sugars Consumption and Ethanol Inhibition in Carob Pulp Fermentation by *Saccharomyces Cerevisiae* in Batch and Fedbatch Cultures." *Journal of Industrial Microbiology and Biotechnology* 39(5): 789–97.

Loveday, S M, and R J Winger. 2007. "Mathematical Model of Sugar Uptake in Fermenting Yeasted Dough." *Journal of agricultural and food chemistry* 55(15): 6325–29.
<http://www.ncbi.nlm.nih.gov/pubmed/17595109>.

MacLean, R Craig, and Ivana Gudelj. 2006. "Resource Competition and Social Conflict in Experimental Populations of Yeast." *Nature* 441: 498–501.

Monod, J. 1949. "The Growth of Bacterial Cultures." *Annual Review of Microbiology* 3: 371–94.

- Raffoux, Xavier et al. 2018. "High-Throughput Measurement of Recombination Rates and Genetic Interference in *Saccharomyces Cerevisiae*." *Yeast*. <http://doi.wiley.com/10.1002/yea.3315> (May 5, 2018).
- dos Santos, Valdilene Canazart, Caio Roberto Soares Bragança, Frederico José Vieira Passos, and Flávia Maria Lopes Passos. 2013. "Kinetics of Growth and Ethanol Formation from a Mix of Glucose/Xylose Substrate by *Kluyveromyces Marxianus* UFV-3." *Antonie van Leeuwenhoek* 103(1): 153–61.
- Sonnleitnert, B, and O. Käppeli. 1986. "Growth of *Saccharomyces Cerevisiae* Is Controlled by Its Limited Respiratory Capacity: Formulation and Verification of a Hypothesis*." *Biotechnology and Bioengineering* 28: 927–37. <https://onlinelibrary-wiley-com.insb.bib.cnrs.fr/doi/pdf/10.1002/bit.260280620>.
- Spor, Aymé. 2009. "Adaptive Strategies to Contrasted Environments in *Saccharomyces Cerevisiae*." Université Paris-Sud XI.
- . 2009. "Niche-Driven Evolution of Metabolic and Life-History Strategies in Natural and Domesticated Populations of *Saccharomyces Cerevisiae*." *BMC evolutionary biology* 9: 296.
- . 2014. "Phenotypic and Genotypic Convergences Are Influenced by Historical Contingency and Environment in Yeast." *Evolution* 68(3): 772–90.
- Tavaré, S, D J Balding, R C Griffiths, and P Donnelly. 1997. "Inferring Coalescence Times from DNA Sequence Data." *Genetics* 145(2): 505–18. <http://www.ncbi.nlm.nih.gov/pubmed/9071603> (May 5, 2018).
- Toni, Tina et al. 2009. "Approximate Bayesian Computation Scheme for Parameter Inference and Model Selection in Dynamical Systems." *Journal of the Royal society Interface* 6: 187–202.
- van Uden, N. 1985. "Chapter 2 - Ethanol Toxicity and Ethanol Tolerance in Yeasts." In *Annual Reports on Fermentation Processes. Volume 8*, ed. GEORGE T. TSAO. Elsevier, 11–58.
- Verduyn, C. 1991. "Physiology of Yeasts in Relation to Biomass Yields." *Antonie van Leeuwenhoek* 60(3–4): 325–53.
- Wang, Feng-Sheng, and Jyh-Woei Sheu. 2000. "Multiobjective Parameter Estimation Problems of Fermentation Processes Using a High Ethanol Tolerance Yeast." *Chemical Engineering Science* 55(18): 3685–95. <http://linkinghub.elsevier.com/retrieve/pii/S0009250900000385>.

CHAPITRE 2

Boucles éco-évolutives entre l'environnement et les traits sous sélection dans un environnement saisonnier : conséquences pour l'évolution expérimentale

Chapitre 2 : Boucles éco-évolutives entre l'environnement et les traits sous sélection dans un environnement saisonnier : conséquences pour l'évolution expérimentale

Les cultures en batch de microorganismes sont fréquemment utilisées en évolution expérimentale pour étudier la dynamique de l'adaptation. Généralement, ce type d'environnement est supposé conduire à une augmentation du taux de croissance sous l'effet de la sélection naturelle. Cependant, d'autres traits d'histoire de vie pourraient être sélectionnés. En effet, des batchs successifs forment un environnement saisonnier, et des caractères différents peuvent être ciblés par la sélection au cours des différentes saisons. De plus, dans un environnement clos comme celui de cultures en batchs, les organismes peuvent avoir un impact sur la composition de leur environnement, conduisant à des boucles de rétroaction éco-évolutives qui doivent être prises en compte pour comprendre l'adaptation.

Dans ce chapitre, je présente un modèle mathématique décrivant une expérience d'évolution chez la levure *Saccharomyces cerevisiae* (Spor et al. 2014), permettant de comprendre quels traits ont été ciblés par la sélection, et de décrire le rôle de l'environnement saisonnier dans des cultures en batch. Je montre que dans cet environnement saisonnier, deux types de traits sont soumis à la sélection : des traits d'histoire de vie en lien avec la croissance et la mortalité à chaque saison, mais aussi des traits de transition, décrivant la capacité des individus à réagir à des changements environnementaux. L'impact de l'environnement peut être résumé par la longueur des saisons, qui détermine les poids respectifs de chaque type de trait dans les composantes de la valeur sélective. Plus la saison est longue, plus la sélection sur les traits caractéristiques de la saison est importante. Par ailleurs, la longueur des saisons est déterminée par les phénotypes des individus, ce qui conduit à l'émergence de boucles de rétroaction éco-évolutives. Mes résultats montrent donc comment l'évolution en batchs successifs affecte à la fois la longueur des saisons et la force de la sélection sur les différents traits.

Research



Cite this article: Collot D, Nidelet T, Ramsayer J, Martin OC, Méléard S, Dillmann C, Sicard D, Legrand J. 2018 Feedback between environment and traits under selection in a seasonal environment: consequences for experimental evolution. *Proc. R. Soc. B* **284**: 20180284.
<http://dx.doi.org/10.1098/rsob.2018.0284>

Received: 5 February 2018

Accepted: 22 March 2018

Subject Category:

Evolution

Subject Areas:

evolution

Keywords:

seasonal environment, mathematical modelling, fitness components, life-history traits evolution

Author for correspondence:

Judith Legrand

e-mail: judith.legrand@inra.fr

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4052783>.

Feedback between environment and traits under selection in a seasonal environment: consequences for experimental evolution

Dorian Collot¹, Thibault Nidelet², Johan Ramsayer^{1,2}, Olivier C. Martin¹, Sylvie Méléard³, Christine Dillmann¹, Delphine Sicard² and Judith Legrand¹

¹GQE-Le Moulon, INRA, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, 91190 Gif-sur-Yvette, France

²SPO, Inra, Univ. Montpellier, Montpellier, SupAgro, Montpellier, France

³CMAF, Ecole Polytechnique, CNRS, Palaiseau, France

DC, 0000-0002-6072-8429; JL, 0000-0001-5157-9268

Batch cultures are frequently used in experimental evolution to study the dynamics of adaptation. Although they are generally considered to simply drive a growth rate increase, other fitness components can also be selected for. Indeed, recurrent batches form a seasonal environment where different phases repeat periodically and different traits can be under selection in the different seasons. Moreover, the system being closed, organisms may have a strong impact on the environment. Thus, the study of adaptation should take into account the environment and eco-evolutionary feedbacks. Using data from an experimental evolution on yeast *Saccharomyces cerevisiae*, we developed a mathematical model to understand which traits are under selection, and what is the impact of the environment for selection in a batch culture. We showed that two kinds of traits are under selection in seasonal environments: life-history traits, related to growth and mortality, but also transition traits, related to the ability to react to environmental changes. The impact of environmental conditions can be summarized by the length of the different seasons which weight selection on each trait: the longer a season is, the higher the selection on associated traits. Since phenotypes drive season length, eco-evolutionary feedbacks emerge. Our results show how evolution in successive batches can affect season lengths and strength of selection on different traits.

1. Introduction

Experimental evolution is a choice tool to study adaptation in controlled environmental conditions [1]. For microorganisms, the two main devices are chemostats and batch cultures [2]. In a chemostat, the population is kept under a constant growth rate while resources are added continuously. By contrast, for batch cultures, the resources are consumed and are not renewed. In the present paper, we focus on the consequences of environmental changes, as mimicked by successive batch cultures, on selective pressures and evolution of traits.

Batch culture often involves 24 h batches, with low resource concentrations, notably for yeast [3–11]. The cell population exhibits roughly an exponential growth followed by a stationary phase. However, population dynamics can be more complex. The depletion and production of resources and/or toxins can lead to a succession of different phases. When these phases are repeated periodically in successive batches, they can be seen as seasons. The make–accumulate–consume strategy of several yeast species [12] typically leads to a seasonal environment. In the presence of oxygen, several yeast species are able to ferment glucose if its concentration is sufficient. When glucose

concentration reaches too low a level, cells enter a transition phase called the 'diauxic shift' during which they switch to respiration metabolism. In the respiration season, yeasts consume the ethanol they have previously produced and the population grows at a lower rate than during fermentation. Considering the ubiquity of metabolic plasticity in microorganisms [13,14], such examples of complex metabolic dynamics leading to seasonal structure might not be an exception but rather the norm, and should be accounted for in the interpretation of experimental evolution.

The consequences of seasonality have been studied experimentally in batch cultures [15] and theoretically using mathematical or computer models [16–18]. Different models (considering different levels of organization, from phenotypic to metabolic levels) have been developed and predict that coexistence may be observed when a trade-off exists between two traits selected in two different seasons [16,17,19]. This coexistence has been observed experimentally, notably in the long-term experimental evolution (LTEE) of *Escherichia coli* [20], where it has been shown that a few mutations distinguish the two coexisting strains, and affect resource consumption and the efficiency of the diauxic shift [17,21–23].

The analysis of fitness in seasonal environments still remains limited by our ability to take into account the interaction between genotype and environment. Usually, fitness is characterized and measured using the ratio of growth rates of two strains over the period of the whole culture; the strain having the higher net growth rate being the fittest [24–26]. However, this does not provide much insight into why one strain performs better than another in a given seasonal environment. One way to predict the population dynamics of competing strains is to decompose fitness into components [27,28]. Different decompositions have been proposed for seasonal environments [27–29]. Some only include growth rates [28] and others account for frequency dependence or trait dependence [29]. However, fitness also depends on the environment and its descriptors should be included in fitness [30].

In this paper, using a mathematical modelling approach, we propose a decomposition of the fitness function accounting for the strain-induced environmental changes. We show that this decomposition retrieves results on coexistence in seasonal environments [15–18]. It also provides a simple tool to understand which traits are under selection in a given environment and how the selection of these traits acts on the environment and feeds back on the weight of the different fitness components. This provides new insights into ecoevolutionary feedbacks and the traits potentially involved in historical contingency and intransitive fitness.

2. Models and methods

(a) The model

We modelled the population dynamics of n competing strains in one batch culture where cells produce biomass successively by fermentation and respiration. In this model, cells only interact indirectly via competition for resources and production of ethanol, which is both a resource and a toxic compound. It is described by a set of differential equations

adapted from MacLean & Gudelj [31].

$$\left. \begin{aligned} \frac{dG}{dt} &= - \sum_i J_{f,i} \frac{G}{K_{f,i} + G} N_i \\ \frac{dN_i}{dt} &= r_{f,i} \frac{G}{K_{f,i} + G} \cdot \exp\left(\frac{-E}{E_{m,i}}\right) \cdot N_i \\ &\quad + r_{r,i} \cdot \exp\left(\frac{-E}{E_{m,i}}\right) \cdot \frac{E}{K_{r,i} + E} \cdot \left(\frac{K_{c,i}}{K_{c,i} + G}\right) N_i - m_i \cdot N_i \\ \frac{dE}{dt} &= \sum_i J_{f,i} \frac{G}{K_{f,i} + G} p_i N_i - \sum_i J_{r,i} \frac{E}{K_{r,i} + E} \left(\frac{K_{c,i}}{K_{c,i} + G}\right) N_i \\ G(0) &= 0.15 \text{ g ml}^{-1}, E(0) = 0 \text{ g ml}^{-1}, \sum_i N_i(0) = 10^6 \text{ cells ml}^{-1} \end{aligned} \right\} \quad (2.1)$$

$G(t)$, simplified to G , is the glucose concentration at time t , $E(t)$, simplified to E , is the ethanol concentration at time t , and for all i in 1 to n , $N_i(t)$, simplified to N_i , is the cell concentration of strain i at time t . When the glucose concentration is high, the cells consume the glucose by fermentation at a rate $J_{f,i} \cdot G / (K_{f,i} + G)$, where $J_{f,i}$ ($\text{g h}^{-1} \text{ cell}^{-1}$) is the maximum glucose consumption rate of strain i and $K_{f,i}$ is the glucose concentration at which the glucose consumption of strain i is half its maximal glucose consumption, called saturation constant for glucose. For fermentation, the ethanol/glucose yield is denoted p_i (in grams of ethanol produced per gram of glucose consumed) and the 'cell' yield is denoted $y_{f,i}$ (in a number of cells produced per gram of glucose). The maximum population growth rate in fermentation is denoted $r_{f,i} = J_{f,i} y_{f,i}$ (h^{-1}). When the glucose resource is depleted, yeast switch to respiration and consume the ethanol at a rate $J_{r,i} \cdot E / (K_{r,i} + E)$, where $J_{r,i}$ ($\text{g h}^{-1} \text{ cell}^{-1}$) is the maximum ethanol consumption rate and $K_{r,i}$ is the ethanol concentration at which the ethanol consumption is the half of the maximal ethanol consumption, called the saturation constant for ethanol. To model the Crabtree effect, we multiply the growth rate by the term $K_{c,i} / (K_{c,i} + G)$ which simulates the inhibition of respiration by the glucose and where $K_{c,i}$ (g ml^{-1}) is an inhibition constant. For respiration, the 'cell' yield is denoted $y_{r,i}$. The maximum population growth rate in respiration is denoted $r_{r,i} = J_{r,i} y_{r,i}$ (h^{-1}). Owing to the toxicity of ethanol, the yield $y_{f,i}$ is decreased by a factor $\text{tox}(E) = \exp(-E/E_{m,i})$ where $E_{m,i}$ denotes the tolerance to ethanol (g ml^{-1}) [32]. The cells' mortality rate, denoted m_i (h^{-1}) is considered constant over the batch for the sake of simplicity. Indeed, the number of observations would not allow the inference of mortality variations along the batch.

For n strains in competition, our model is therefore composed of $(2 + n)$ differential equations with a total set of $10n$ parameters ($J_{f,i}, r_{f,i}, m_i, E_{m,i}, p_i, J_{r,i}, r_{r,i}, K_{c,i}, K_{f,i}, K_{r,i}$) where $G(t)$, $E(t)$ and $N_i(t)$ are, respectively, the glucose (g ml^{-1}), ethanol (g ml^{-1}) and cell concentration of strain i (cell ml^{-1}) at time t .

The system was implemented in C and solved by the Runge–Kutta method.

(b) Model calibration

(i) Experimental data

We fitted the 'one-strain version' of the model to data from 72 different strains obtained from a 6-month experimental evolution described in detail in [11] and in electronic supplementary material S1. To fit the one-strain model, for each strain, we used observed data of glucose concentration $G(t)$, ethanol concentration $E(t)$, cell density $Z(t)$ and the

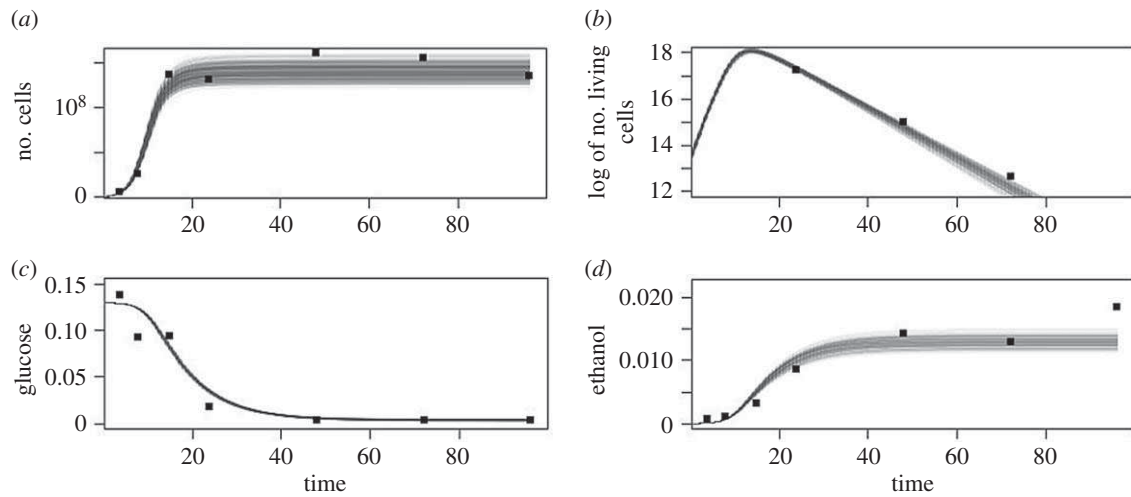


Figure 1. Example of predictions of the model for one strain evolved in a 1%–96 h environment (a) for the total density of cells, (b) for the logarithm of the density of living cells, (c) for the glucose concentration, (d) for the ethanol concentration. Grey curves are the 10 000 best simulations obtained from the ABC algorithm. Black points are experimental data.

survival rate $s(t)$ monitored over a 96 h batch monoculture in a 15% glucose medium [11].

(ii) Parameters inference

For each strain, we decided to estimate seven parameters (J_f , r_f , m , E_m , p , J_r , r_r) of the model. The remaining parameters (K_f , K_r and K_c) were set at 5×10^{-4} g ml $^{-1}$. We checked that parameter estimates were not sensitive to (K_f , K_r and K_c) values. We also allowed the model calibration to involve the initial glucose concentration G_0 and the initial population concentration N_0 . These initial conditions were estimated as they may vary from one batch to the other due to experimental variations.

To estimate the parameters $\theta = (J_f, r_f, m, E_m, p, J_r, r_r, G_0, N_0)$, we used the ABC-SMC algorithm [33]. The distance $d(X_{\text{obs}}, X_{\text{pred}})$ between data and model predictions was calculated as the Euclidian distance between the corresponding values of $G(t)$, $Z(t)$, $E(t)$ and $N(t) = Z(t) \cdot s(t)$ at times 3, 6, 9, 12, 16, 24, 48, 72 and 96 h, normalized by the observed values and the number of observations (n_N , n_Z , n_E and n_G). Observed ethanol concentration at 16 h was always lower than at 12 and 24 h, presumably because an additional ethanol dilution had to be carried out to stay in the condition of proper linear quantification by the kit. That extra dilution may have impacted all dosages at this point. Therefore, we decided to discard these observations in the estimation procedure. In order to test the accuracy of the model, the distributions of standardized residuals were examined for each parameter. More details and the algorithm are provided in electronic supplementary material S1 and S2, as well as the detailed analysis of these results.

(c) *In silico* experiments of competitions

One thousand strains were generated *in silico* by sampling 1000 sets of parameters in the parameter prior distributions (calculated from the estimated parameters based on experimental data) using uniform latin-hypercube sampling [12] (electronic supplementary material, table S2). For each couple of strains, we simulated pairwise competitions over one batch between a resident strain (initial concentration = 10^6 cells ml $^{-1}$) and a mutant strain (initial concentration = 10 cells ml $^{-1}$), leading to a total of 999 000 competitions.

For each competition, the effective length of seasons was calculated for each strain using Riemann formula with time

step of 0.02 h and 4800 steps. The sensitivity of fitness and effective season lengths to the trait values was analysed using partial rank correlation coefficients (PRCC) using the package *epiR* [34].

We also quantified the frequency-dependence of fitness by simulating *in silico* competitions over one batch as described above with different initial frequencies of the mutant and resident strains. For each couple of strains, we simulated pairwise competitions with different initial frequencies of the mutant strain (10^{-6} , 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.999999 cell ml $^{-1}$). For each couple, we determined whether coexistence was possible by finding initial frequencies where the fitness reached 0.

Finally to address the long-term dynamics of competition between two strains, we simulated the competition over several successive batches. To go from one batch to the next, we reset the glucose concentration to 0.15 g ml $^{-1}$, the ethanol concentration to 0, the population concentration to 10^6 cells ml $^{-1}$ while keeping frequency of each strain at its final frequency in the preceding batch.

3. Results

To explore traits under selection, we built a system of differential equations that simulates a population of *S. cerevisiae* and its effects on resource dynamics in respiro-fermentative batches. We considered three seasons according to resource availability: first a fermentation phase where glucose is consumed and ethanol is produced; then a respiration phase where ethanol is consumed; and finally a mortality phase where resources are exhausted and cells are quiescent or dead. The environment is described by glucose and ethanol concentrations. Ethanol is considered as both a resource and a poison inhibiting growth. The model has been built for multiple strains, interacting only through the environment. Each strain is characterized by its population dynamics, which depends only on resource concentration. To calibrate and validate the model, we fitted the ‘one strain version of the model’ to experimental data regarding 72 strains grown separately (figure 1; electronic supplementary material S1). These fits provided 72 sets of parameter values that were used to define their range in simulations.

The model was then used to analyse and simulate the dynamics of competing strains to study fitness components. We analysed the traits under selection during each season and how season lengths affect selection. We then highlighted the feedbacks between selection and the environment and explored their potential consequences on life-history traits adaptation.

(a) Fitness components

To assess which traits are selected during a batch culture, we used a decomposition of the fitness function. Using the Malthusian fitness function [14], the fitness in a batch of length T , $W_{1/2}(T)$, can be defined as

$$\begin{aligned} W_{1/2}(T) &= \ln\left(\frac{N_1(T)}{N_1(0)}\right) - \ln\left(\frac{N_2(T)}{N_2(0)}\right) \\ &= \ln\left(\frac{N_1}{N_2}\right)(T) - \ln\left(\frac{N_1}{N_2}\right)(0) \\ &= \int_0^T \frac{d(\ln(N_1/N_2))}{dt} dt. \end{aligned} \quad (3.1)$$

If the two strains have the same traits, $\ln(N_1/N_2)(T) = \ln(N_1/N_2)(0)$ leading to a null fitness. If $W_{1/2}(T) > 0$, strain 1 is fitter than strain 2.

Combining equations (2.1) and equation (3.1), one obtain:

$$\begin{aligned} W_{1/2}(T) &= \int_0^T \frac{d(\ln(N_1/N_2))}{dt} dt \\ &= \int_0^T r_{f,1} \cdot \exp\left(\frac{-E}{E_{m,1}}\right) \cdot \frac{G}{K_{f,1} + G} dt \\ &\quad - \int_0^T r_{f,2} \cdot \exp\left(\frac{-E}{E_{m,2}}\right) \cdot \frac{G}{K_{f,2} + G} dt \\ &\quad + \int_0^T r_{f,1} \cdot \exp\left(\frac{-E}{E_{m,1}}\right) \cdot \frac{E}{K_{r,1} + E} \cdot \frac{K_{c,1}}{K_{c,1} + G} dt \\ &\quad - \int_0^T r_{f,2} \cdot \exp\left(\frac{-E}{E_{m,2}}\right) \cdot \frac{E}{K_{r,2} + E} \cdot \frac{K_{c,2}}{K_{c,2} + G} dt - (m_1 - m_2)T \end{aligned} \quad (3.2)$$

To simplify notation in this expression, $E(t)$ and $G(t)$ were, respectively, denoted E and G . For each strain ($i = 1, 2$), let us define

$$T_i^{\text{ferm}} = \int_0^T \exp\left(\frac{-E}{E_{m,i}}\right) \cdot \frac{G}{K_{f,i} + G} dt \quad (3.3)$$

and

$$T_i^{\text{resp}} = \int_0^T \exp\left(\frac{-E}{E_{m,i}}\right) \cdot \frac{E}{K_{r,i} + E} \cdot \frac{K_{c,i}}{K_{c,i} + G} dt. \quad (3.4)$$

We called these quantities ‘the effective season length’ of fermentation (T_i^{ferm}) and respiration (T_i^{resp}). Considering the simplified model where $E_m \rightarrow \infty$, $K_r \rightarrow 0$, $K_f \rightarrow 0$ and $K_c \rightarrow 0$, T_i^{ferm} is the time to glucose exhaustion and T_i^{resp} the time to ethanol exhaustion (see electronic supplementary material S3). In this simple case where concentrations of resources vanish in a finite time, the ‘effective season length’ corresponds to the period of time where the resource is effectively used. In the general case, they depend on resources, the strains’ parameters and how these strains react to abiotic conditions.

With (3.3) and (3.4), fitness can be written as

$$\begin{aligned} W_{1/2}(T) &= (r_{f,1} \cdot T_1^{\text{ferm}} - r_{f,2} \cdot T_2^{\text{ferm}}) \\ &\quad + (r_{r,1} \cdot T_1^{\text{resp}} - r_{r,2} \cdot T_2^{\text{resp}}) - (m_1 - m_2)T. \end{aligned} \quad (3.5)$$

Alternatively, (3.5) can be written in terms of differences between strains as

$$\begin{aligned} W_{1/2}(T) &= \Delta r_f \cdot T_1^{\text{ferm}} + \Delta r_r \cdot T_1^{\text{resp}} - \Delta m \cdot T + \Delta T^{\text{ferm}} \cdot r_{f,2} \\ &\quad + \Delta T^{\text{resp}} \cdot r_{r,2}. \end{aligned} \quad (3.6)$$

For each parameter θ (in $r_{\text{ferm}}, r_{\text{resp}}, m, T_{\text{ferm}}, T_{\text{resp}}$), $\Delta\theta = \theta_1 - \theta_2$. Therefore, the fitness is composed of a sum of terms, each one describing a different component of differences between two strains weighted by the length of a season.

The first three terms correspond to the direct advantages given by life-history traits: growth rates during fermentation (Δr_f) or respiration (Δr_r) and mortality (Δm), weighted by their respective season length ($T_1^{\text{ferm}}, T_1^{\text{resp}}, T$) for the reference strain 1. The last two terms depend on strains differences for their effective season length ΔT^{ferm} , and ΔT^{resp} , weighted by the corresponding seasonal growth rate of the competing strain 2. For each season, a positive ΔT term means that strain 1 is able to take advantage of the resources efficiently over a longer period than strain 2. Hence these terms, called transition terms in this article, measure how strains respond differently to environmental changes.

To interpret this equation, consider two strains differing by one single model parameter. Two cases can be considered. Case a: this parameter is a life-history trait (r_f, r_r or m): effective season lengths will be equal (equations (3.3) and (3.4)) so $\Delta T^{\text{ferm}} = 0$ and $\Delta T^{\text{resp}} = 0$ and $W_{1/2}(T) = \Delta r_x \cdot T_1^x$ ($x = \text{ferm}$ or resp) or

$$W_{1/2}(T) = -\Delta m \cdot T \quad (3.7)$$

depending on the considered parameter. Case b: this parameter is not a life-history trait but affects the effective season lengths T^{ferm} and T^{resp} of each strain. Then

$$W_{1/2}(T) = r_{f,2} \cdot (\Delta T^{\text{ferm}}) \quad (3.8)$$

or

$$W_{1/2}(T) = r_{r,2} \cdot (\Delta T^{\text{resp}}). \quad (3.9)$$

Hence, any parameter θ for which a strain difference leads to a difference in effective season length is under selection.

The effective season length of one strain can depend on the characteristics of its competitor. For example, having strain 2 consume resources faster will lead to a shorter season for strain 1. As the effective season length weights the advantage conferred by a life-history trait (case a), any competitor trait that changes season length will change selection strength on life-history traits. Similarly, for traits for which strain differences impact ΔT (case b), selection will be higher if the growth rate of the corresponding season is high.

In summary, fitness is composed of five terms: three can be considered as life-history terms and two are transition terms, related to the ability to keep the same metabolic behaviour (fermentation or respiration) despite the environment changes. Life-history terms depend on differences in growth or mortality rates in each season weighted by its length. The transition terms are differences in response to environment changes weighted by growth rates. Hence, a strain has two ways to be fitter than its competitor: having beneficial life-history traits (i.e. having faster growth or lower mortality), or taking better advantage of the available resources.

Table 1. Link between difference of mutant and resident season length and traits. Partial rank correlations between the difference of length of fermentation and respiration season and traits of the two strains computed with 100 000 strain couples. The traits indexed R are the traits of the resident, the sign Δ indicates the difference between the traits of the mutant and the traits of the resident. Italicized correlations are correlations larger than 0.2 in absolute value.

	$r_{f,R}$	$r_{r,R}$	m_R	$E_{m,R}$	$y_{f,R}$	$y_{r,R}$	p_R	$K_{f,R}$	$K_{r,R}$	$K_{c,R}$
ΔT^{ferm}	0.123	−0.004	−0.004	0.455	0.062	0.035	−0.327	−0.038	0.036	−0.023
ΔT^{resp}	−0.001	0.115	0.002	0.274	0.150	−0.053	−0.374	−0.030	0.094	−0.051
	Δr_f	Δr_r	Δm	ΔE_m	Δy_f	Δy_r	Δp	ΔK_f	ΔK_r	ΔK_c
ΔT^{ferm}	0.014	−0.180	0.014	0.940	0.148	0.028	0.009	−0.029	0.006	0.001
ΔT^{resp}	0.006	−0.010	0.019	0.837	0.009	0.021	0.001	−0.004	−0.230	0.033

Table 2. Link between season length and traits. Partial rank correlations between the length of the fermentation and respiration season and traits of the two strains computed with 100 000 strain couples. The traits indexed by R are the traits of the resident, traits indexed by M are the one of the mutant. Italicized correlations are correlations larger than 0.2 in absolute value.

	$r_{f,R}$	$r_{r,R}$	m_R	$E_{m,R}$	$y_{f,R}$	$y_{r,R}$	p_R	$K_{f,R}$	$K_{r,R}$	$K_{c,R}$
T_M^{ferm}	−0.948	0.031	0.548	−0.272	0.788	0.057	−0.522	0.129	0.097	−0.044
T_M^{resp}	−0.174	−0.578	0.475	−0.442	−0.764	0.628	0.829	−0.013	0.116	−0.085
	$r_{f,M}$	$r_{r,M}$	m_M	$E_{m,M}$	$y_{f,M}$	$y_{r,M}$	p_M	$K_{f,M}$	$K_{r,M}$	$K_{c,M}$
T_M^{ferm}	0.004	0.007	0.002	0.751	0.020	−0.007	0.014	−0.019	0.007	0.007
T_M^{resp}	−0.005	−0.006	0.009	0.589	0.011	0.001	0.029	−0.007	−0.093	0.016

(b) Impact of traits on effective season lengths

According to equation (3.6), traits subject to selection are either life-history traits (r_f , r_r or m) or ones changing ΔT^{ferm} or ΔT^{resp} . To explore how fitness may be affected by variations of traits, we therefore explored how these variations impact the terms T_1^{ferm} , T_1^{resp} , ΔT^{ferm} and ΔT^{resp} .

We tackled this question using a set of *in silico* competitions of 1000 strains, computing the effective season lengths (T_R^{ferm} , T_R^{resp} , T_M^{ferm} , T_M^{resp}). Each strain was characterized by a set of traits sampled in a prior distribution based on experimental data and the literature. In each simulated competition, strain 2 is the resident strain (R, high initial frequency) and strain 1 is a mutant strain (M) whose very low frequency is such that it does not impact the resource dynamics. Hence if a resident trait is correlated with the season length of the mutant, this trait impacts season length via resource dynamics. On the contrary if a mutant trait correlates with its own season length, this means that the mutant impacts season length via the strain's response to the environment.

The partial rank correlation coefficient (PRCC) between season length and the individual traits of the mutant and the resident were estimated. We first explored traits that are subject to selection and then traits that affect selection strength.

(i) Traits subject to selection

A trait was considered as a life-history trait when the difference for this trait between the mutant and the resident led to variations in one of the two transition terms ΔT^{ferm} or ΔT^{resp} . PRCC between the effective season length differences and the trait differences are shown in table 1. Only two metabolic traits were subject to selection: ethanol tolerance E_m and respiration saturation constant K_r . The difference in ethanol tolerance between both strains (ΔE_m) tended to increase both ΔT^{ferm} and ΔT^{resp} , indicating that a higher tolerance to

ethanol will be selected for. In the same way, ΔT^{resp} slightly decreased when the difference between saturation constants increased (ΔK_r), indicating that a lower saturation constant will be selected for. Indeed, the lower the saturation constant, the longer a strain is effective in using that resource.

To summarize, higher growth rates, lower mortality rates, higher tolerance to ethanol and, to a lesser extent, lower respiration saturation constants will be under selection.

(ii) Traits that change selection strength

We first investigated the effect of mutant and resident phenotype on the mutant season lengths T_M^{resp} and T_M^{ferm} . PRCC between the mutant effective season lengths (T_M^{resp} and T_M^{ferm}) and the traits are shown in table 2. A positive correlation is expected to enhance the efficiency of selection on growth rate, while a negative correlation is expected to decrease it.

Tolerance to ethanol $E_{m,M}$ is the only trait of the mutant which affects the effective season lengths of the mutant (PRCC = 0.75 for T_M^{ferm} and 0.58 for T_M^{resp}). Otherwise the mutant effective season lengths are affected by traits of the resident strains, via the resource dynamics which can be modified by different mechanisms. Increasing the cell concentration reduces resources more rapidly, reducing the corresponding season length. This is in line with the negative correlation of the season lengths T_M^{ferm} and T_M^{resp} with the growth rates of the resident ($r_{f,R}$ and $r_{r,R}$) and their positive correlation with the mortality rate of the resident m_R . In addition, the respiration season length T_M^{resp} is negatively impacted by the tolerance to ethanol of the resident $E_{m,R}$, and the fermentation yield of the resident $y_{f,R}$, as they both increase the cell concentration in this season. Similarly, traits that impact the efficiency of resource consumption (at constant growth rate) will spare the resources and consequently increase the corresponding season length. This is in line with the positive correlation of the fermentation

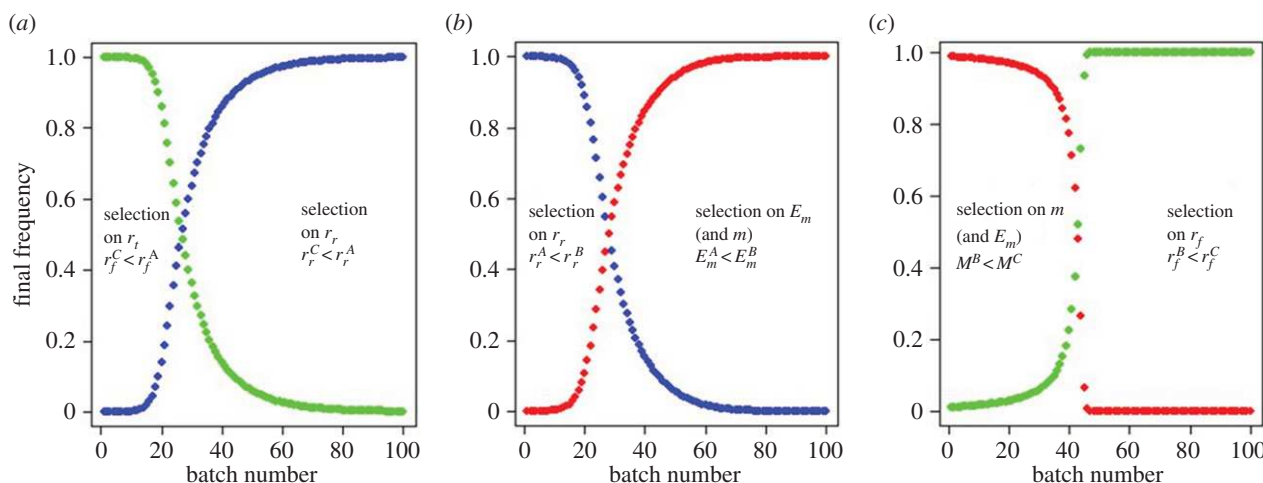


Figure 2. Competition between three strains and non-transitivity. On each panel, the frequency of the strain (blue for strain A, red for strain B and green for strain C) at the end of each batch is represented. The strains are propagated through 100 batches of 90 h. The text indicates the traits which influence fitness the most and are responsible for the invasion in the first batches (on the left) and in the later batches (on the right). Tolerance to ethanol and mortality are selected together, but we only indicate the traits responsible for the positive fitness. Parameter values are given in electronic supplementary material S3.

(respectively respiration) season length T_M^{ferm} (resp. T_M^{resp}) with the fermentation of the resident (resp. respiration) yield $y_{f,R}$. Finally, traits that impact resource production such as the ethanol/glucose yield p_R can be positively correlated to the season length where this resource is consumed.

We also investigated the effect of mutant characteristics on the transition terms ΔT^{ferm} and ΔT^{resp} . As shown previously, the only trait under strong selection through the transition terms' fitness component is ethanol tolerance. If (ΔT) is correlated with the trait of the resident, this trait will therefore affect the efficiency of selection on ethanol tolerance. PRCC between the transition terms and the resident traits are shown in table 1. We found that an increase in ethanol tolerance increased both ΔT^{ferm} and ΔT^{resp} , while an increase in the ethanol/glucose yield (p_R) decreased both terms.

Altogether, our results showed that ethanol tolerance and saturation constants are under selection and the strength of selection on these traits is weighted by the ethanol/glucose yield of the resident strains. Besides, the strength of selection on growth rates is modulated by almost all resident characteristics, except for the saturation and inhibition constants, as well as by the mutant ethanol tolerance.

(c) Feedback loops between traits under selection

Just as trait values affect season lengths, season lengths drive the strength of selection on traits (equation (3.6)). As a consequence, a feedback loop between selected traits and the environment arise: selection affects trait values, the associated changes modify season lengths, and then these modified season lengths impact the trait values via selection. Such eco-evolutionary feedbacks have been described in different situations and are known to have consequences on the evolutionary trajectories by promoting diversification such as branching or coexistence [35]. In this section, we explore and illustrate these feedbacks.

(i) Description of feedback loops

Four traits are involved in feedback loops between selection and environment: the growth rates (r_f and r_r), the mortality rate (m) and the tolerance to ethanol (E_m). As seen in §3a, selective pressures tend to increase the fermentation and respiration

growth rates (equation (3.7)), to decrease the mortality rate (equation (3.7)) and to increase the tolerance to ethanol (table 1). The evolution of these four traits has different consequences on the effective season lengths and thus on the strength of selection. All the results described here are a compilation of the results from the previous section, and are summarized in electronic supplementary information S4. We consider the effect of a change in the resident trait on the mutant's fitness, and therefore which mutants are selected.

An increase of the growth rate (resp. r_f and r_r) shortens the season (resp. T_f and T_r) (table 2) and thus decreases the weight of the difference of the growth rate in the fitness function (equation (3.6)). Growth rates also weight ΔT^{ferm} and ΔT^{resp} (table 1), and consequently increase selection on ethanol resistance.

A decrease of the mortality rate shortens both the fermentation and the respiration seasons (table 2). Therefore, selection on mortality rates tends to decrease selection strength on growth rates, and indirectly increases selection on mortality rate.

Finally, selection tends to increase the tolerance to ethanol, which in turn decreases the length of the fermentation and respiration seasons (table 2), but increases ΔT^{ferm} and ΔT^{resp} (table 1). Consequently, selection on tolerance to ethanol tends to decrease selection on growth rates and increase selection on itself.

The fermentation yield, respiration yield and ethanol/glucose yield are not selected in our model, and consequently they are not involved in any feedback loop. Indeed, if only the yield varies between two strains, every ΔT term and $\Delta \theta$ term will be null, and so will be fitness (equation (3.6)). However, yield can affect season lengths and therefore enhance or inhibit selection on one trait. For instance, a higher yield (resp. y_f and y_r) enhances selection for higher growth rate (resp. r_f and r_r) (table 2).

(ii) Consequences of eco-evolutionary feedbacks in batch cultures

As shown above, in batch cultures, because of the dynamics of nutrients and toxins, the strength of selection is not constant during the evolution and can sometimes depend strongly on the resident strain. One potential consequence is non-transitivity of fitness. For instance, C can invade B but not A, B can invade A but not C, A can invade C but not B (figure 2;

electronic supplementary material S5). Another consequence is the possible coexistence between strains via frequency dependence. When the mutant is at very low frequency, its impact on the dynamics of the medium is negligible. However, as its frequency increases, its impact on the resources and toxin dynamics can become essential. As a result, the effective season lengths will be modified gradually as well as the mutant's fitness. Its fitness can decrease and even reach zero. Note that as $W_{1/2} = -W_{2/1}$, coexistence implies that the two strains have a negative fitness when they are resident [36], and this is possible if at least two life-history traits are different between the strains. Indeed, if only one trait differs between two strains, the fitness is reduced to a single term (equations (3.7), (3.8) or (3.9)) and cannot vanish.

To explore the relationship and the occurrence of coexistence between two strains, we have generated 100 strains and simulated 10 000 pairwise competitions (see Models and methods). In 44% of these competitions, fitness was negative frequency dependent, indicating a negative feedback. In 8% of the simulations fitness sign changed indicating possible coexistence. In half of these 8%, there was a stable coexistence equilibrium. Lastly 93% of the simulated strains coexisted with at least one other strain.

4. Discussion

There have been several studies in the literature regarding how to define and measure fitness [24–26,30,37–39]. For a seasonal environment, these definitions are often based on the average growth rate in competition but they do not incorporate explicitly the determinants of this growth rate nor the environment. Here, we developed a mechanistic mathematical model which explicitly includes life history and resource use, as well as toxicity traits and their impact on both the population dynamics and the environment. For that we followed previous authors who modelled environmental variations such as the duration of daylight [27] or the amount of different resources [20,31]. Because our purpose was to provide a definition of fitness that includes the effect of the environment on the strain dynamics, we defined the 'effective season length', which can be interpreted as the time needed for a specific strain to reach the same fitness if growth rates were constant and maximal. In a general case, effective season lengths do not only rely on the level of resources, they also depend on the strains interaction with environmental changes. This definition implies that season lengths are not exactly the same for every strain, even if they are put into the same initial environment. The fitness function can be interpreted as a sum of the advantages in each season, weighted by the season's length. This decomposition allows classifying traits and studying eco-evolutionary feedbacks.

Our work shows that traits can be grouped according to their response to selection and their impact on the environment. The first group is composed of life-history traits (growth and mortality rates). These traits are well known as fitness components [37] and are responsible for a 'selective advantage' in each season. The second group contains transition traits (saturation and inhibition constants related to resource consumption and metabolism switches) which are related to environmental changes: the strains which are less sensitive to changes in levels of resources will have a selective advantage. These traits are less frequently studied, because their effect on the usual function of fitness (i.e. ratio of

growth rates [24,28]) arises solely via their impact on average growth rates. An example of selection on these two kinds of traits can be found in the LTEE on *E. coli* [20]. After 20 000 generations, evolved strains are able to grow faster, but also have a lower saturation constant and consume resources sooner than their ancestor when resources are renewed. Other works showed that the ability to switch faster from one resource to the other has been selected during the LTEE [15,17,19,40,41].

Selection pressure also depends on season length. The impact of season length on selected traits has been studied both theoretically and experimentally in batches by considering cheaters and cooperators [31], in chemostat cultures [27], and in organisms with more complex life cycle and thus having for instance more elaborated consumption strategies [42]. Several experimental evolution studies have also shown that season lengths drive traits selection in batch culture. This has been found for instance in a study by Pekonnen *et al.* [43] where bacteria (*Serratia marcescens* and *Novosphingobium capsulatum*) evolving in long batches of 7 days were mainly selected for a lower mortality rates rather than higher growth rates. On the contrary, in many other works [3–8,10,20,44,45] using one-day-long batches, the growth rate increased while the mortality rate remained unchanged. This is for example the case in the experimental work on which our model was based and in which we compared the evolution of life-history and metabolic traits in batches differing by their batch lengths [11]. In media with long batches, the growth rate seems to increase less than in media with short batches. These examples show how experimental evolution design can affect the outcome of selection.

Besides the effect of season lengths on selection outcomes, the dynamics of strains will affect season lengths. The variation of strain frequencies over successive batches may change season lengths, forming feedback loops that can favour or disfavour different traits. Fitness in batch systems is inevitably frequency dependent. The associated feedback loops are crucial to understand evolution and adaptation [36]. In an experimental evolution with successive batches, the emergence of different strains will typically alter season lengths, which will in turn change the strength of selection on different traits. For example, in case of cross-feeding [17,19,40,41], one strain produces a resource used by another one; the ability of the 'producing strain' to release resources will modulate the length of the season where this resource is consumed, and consequently selection on the consuming strain.

Here, we have highlighted several feedback loops where the strength of selection should decrease over the course of evolution. For example, fermentation growth rate increases due to selection, but the strength of selection on this trait will decrease with batch number. This phenomenon is likely to be quite general: as a trait responds to selection, the corresponding season where the selection is effective typically shortens, and consequently the selective pressure will decrease with evolutionary time.

Regarding experimental evolution, the feedback loops provide an explanation of some observed non-trivial patterns. For instance, there are examples of experimental evolution aiming at comparing evolutionary outcomes in stressed and non-stressed environments. One way to design a stressful environment is to add a poison [3–5], which can increase the mortality rate. However, according to our simulations, fermentation season length is positively correlated with mortality rate. Hence, in the 'stressed' environment, the

fermentation season will be longer and, as a consequence, the stressed and non-stressed environments will also differ in the resource dynamics. Then, a higher selection on growth rate is expected in the 'stressed' environment than in the non-stressed environment. As the experimental measurement of the effective season length may be difficult, this effect will be hard to quantify and account for without the use of a mathematical model incorporating traits and their impact on the environment. However, taking the seasons or a proxy into account should allow us to predict the outcome of adaptation more reliably than looking solely at phenotypes.

Feedback loops in batch culture may lead to difficulties in practical fitness measurement due to frequency dependence. Fitness is often measured by performing competitions between the initial ancestor(s) and the evolved strains. In a single season environment, this measurement provides an accurate measurement of adaptation and changing the competing strain or its frequency does not affect the ranking of fitness values [39]. But in more complex cases, the environment and notably season lengths can evolve with competing strains, and consequently an evolved strain may be less fit than its ancestor due to non-transitivity of fitness. This non-transitivity of fitness, observed in several experimental evolutions [46], is described by the 'rock-paper-scissors' game [47] and has been studied in many scenarios within spatially structured environments or via clonal interference [48,49]. Our work highlights that modelling the population and environment dynamics for decomposing fitness over a season cycle can help understand such non-transitivity.

In the present work, we quantified the genericity of coexistence using *in silico* competitions, sampling the traits values independently. Adding some metabolic constraints between traits could either prevent or enhance the coexistence pattern [27,50] and could help to explain why in some cases no significant change of the fermentation growth rate over the course on the evolution is observed [51]. For instance, coexistence by cross-feeding is possible if a negative relation exists between the ability to produce the resources and the efficiency to use it [17,19,40,41]. In a future work, we are planning to study the interplay between trade-offs and eco-evolutionary feedbacks in seasonal environments.

Exploiting the data obtained in [11], the model captures the fermentation-respiration behaviour of most strains although two biases have been observed. First, for strains which begin respiration while the glucose concentration is still high, the model cannot fit concentrations of both resources. Second, the mortality tends to be overestimated. This trend could be due to the choice of a constant mortality. However, because of the low number of observations for

mortality, it was not possible to consider more complex mortality functions. We cannot dismiss that mortality rate has increased along the batch due to a lack of resources or osmotic stress. To investigate the variation of mortality, one would need more detailed data throughout the batch. However, whatever the form of the mortality rate, it should appear in the decomposition of fitness as terms accounting for the trait differences between the two strains and terms accounting for the difference between efficient season lengths. Then, the fitness decomposition could be used to understand selection pressures on traits related to mortality. Hence, despite its simplicity and limits, this approach helps investigating how eco-evolutionary dynamics can shape evolutionary paths in experimental evolution in batch cultures of organisms exhibiting several metabolic phases.

To conclude, we proposed a decomposition of the fitness function which allows to better grasp selection pressure. In the perspective of setting up an experimental evolution experiment, these results can be useful at two levels. First, results on feedback loops can be used to propose alternative interpretations of evolutionary outcomes. Second, they can guide the scientist in the design of the experimental evolution conditions. The scientist could play on environmental factors to exploit positive feedbacks that will strengthen the selective pressure while avoiding negative feedbacks. In the context of technological applications such as strain selection for industrial fermentation, this could provide a way to better select for a targeted character and avoid collateral damage coming from indirectly selecting for the wrong trait. In practice, further work is needed to explore how to translate such theoretical results into a tool for adapting batch length or other environmental factors to real-time observations of evolutionary dynamics.

Data accessibility. From the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.5vn65dh>.

Authors' contributions. D.C., T.N., D.S., J.L. conceived the study and developed the model; D.C., T.N., S.M., O.M., C.D., J.L. contributed to the mathematical development and statistical analysis; D.C., T.N., J.R., C.D., D.S., J.L. interpreted the results; D.C. performed the analysis, D.C., T.N., D.S. and J.L. wrote the manuscript which was edited by all authors. All authors read and approved the final manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work is supported by the 'IDI 2014' project funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-02.

Acknowledgements. We are grateful to L. M. Chevin, S. de Monte, C. Nizak and D. de Vienne for their discussion and advice on this work. We also want to thank the Chaire 'Modélisation Mathématique et Biodiversité' of Veolia Environnement-Ecole Polytechnique-Museum National d'Histoire Naturelle-Fondation X.

References

- Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. 2012 Experimental evolution. *Trends Ecol. Evol.* **27**, 547–560. (doi:10.1016/j.tree.2012.06.001)
- Dunham MJ. 2010 *Experimental evolution in yeast: a practical guide*, 2nd edn. Cambridge, MA: Elsevier Inc.
- Cowen LE, Kohn LM, Anderson JB. 2001 Divergence in fitness and evolution of drug resistance in experimental populations of *Candida albicans*. *J. Bacteriol.* **183**, 2971–2978.
- Dhar R, Sägeser R, Weikert C, Yuan J, Wagner A. 2011 Adaptation of *Saccharomyces cerevisiae* to saline stress through laboratory evolution. *J. Evol. Biol.* **24**, 1135–1153. (doi:10.1111/j.1420-9101.2011.02249.x)
- Gorter FA, Aarts MMG, Zwaan BJ, de Visser JAGM. 2016 Dynamics of adaptation in experimental yeast populations exposed to gradual and abrupt change in heavy metal concentration. *Am. Nat.* **187**, 110–119. (doi:10.1086/684104)
- Joseph SB, Kirkpatrick M. 2008 Effects of the [PSI⁺] prion on rates of adaptation in yeast. *J. Evol. Biol.* **21**, 773–780. (doi:10.1111/j.1420-9101.2008.01515.x)
- Laan L, Koschwanez JH, Murray AW. 2015 Evolutionary adaptation after crippling cell

- polarization follows reproducible trajectories. *Elife* **4**, 1–18. (doi:10.7554/eLife.09638)
8. Samani P, Bell G. 2010 Adaptation of experimental yeast populations to stressful conditions in relation to population size. *J. Evol. Biol.* **23**, 791–796. (doi:10.1111/j.1420-9101.2010.01945.x)
 9. Jasmin J, Zeyl C. 2012 Life-history evolution and density-dependent growth in experimental population of yeast. *Evolution* **66**, 3789–3802. (doi:10.1111/j.1558-5646.2012.01711.x)
 10. Frenkel EM, McDonald MJ, Van Dyken JD, Kosheleva K, Lang GI, Desai MM. 2015 Crowded growth leads to the spontaneous evolution of semistable coexistence in laboratory yeast populations. *Proc. Natl Acad. Sci. USA* **112**, 11 306–11 311. (doi:10.1073/pnas.1506184112)
 11. Spor A *et al.* 2014 Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. *Evolution* **68**, 772–790. (doi:10.1111/evo.12302)
 12. Hagman A, Säll T, Compagno C, Piskur J. 2013 Yeast ‘make-accumulate-consume’ life strategy evolved as a multi-step process that predates the whole genome duplication. *PLoS ONE* **8**, e68734. (doi:10.1371/journal.pone.0068734)
 13. Comte J, Fauteux L, Del Giorgio PA. 2013 Links between metabolic plasticity and functional redundancy in freshwater bacterioplankton communities. *Front. Microbiol.* **4**, 1–11. (doi:10.3389/fmicb.2013.00112)
 14. Cumming BM, Steyn AJC. 2015 Metabolic plasticity of central carbon metabolism protects mycobacteria. *Proc. Natl Acad. Sci. USA* **112**, 13 135–13 146. (doi:10.1073/pnas.1518171112)
 15. Spencer CC, Saxer G, Travisano M, Doebeli M. 2007 Seasonal resource oscillations maintain diversity in bacterial microcosms. *Evol. Ecol. Res.* **9**, 775–787.
 16. Herron MD, Doebeli M. 2011 Adaptive diversification of a plastic trait in a predictably fluctuating environment. *J. Theor. Biol.* **285**, 58–68. (doi:10.1016/j.jtbi.2011.06.007)
 17. Großkopf T *et al.* 2016 Metabolic modelling in a dynamic evolutionary framework predicts adaptive diversification of bacteria in a long-term evolution experiment. *BMC Evol. Biol.* **16**, 163. (doi:10.1186/s12862-016-0733-x)
 18. Rocabert C, Knibbe C, Consuegra J, Schneider D, Beslon G. 2017 Beware batch culture: seasonality and niche construction predicted to favor bacterial adaptive diversification. *PLoS Comput. Biol.* **13**, e1005459. (doi:10.1371/journal.pcbi.1005459)
 19. Doebeli M. 2002 A model for the evolutionary dynamics of cross-feeding polymorphisms in microorganisms. *Popul. Ecol.* **44**, 59–70. (doi:10.1007/s101440200008)
 20. Vasi F, Travisano M, Lenski RE. 1994 Long-term experimental evolution in *Escherichia coli*. Changes in life-history traits during adaptation to a seasonal environment. *American Naturalist* **144**, 432–456.
 21. Le Gac M, Pluain J, Hindre T, Lenski RE, Schneider D. 2012 Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **109**, 9487–9492. (doi:10.1073/pnas.1207091109)
 22. Pluain J *et al.* 2014 Epistasis and allele specificity in the emergence of a stable polymorphism in *Escherichia coli*. *Science* **343**, 1366–1369. (doi:10.1126/science.1248688)
 23. Spencer CC, Bertrand M, Travisano M, Doebeli M. 2007 Adaptive diversification in genes that regulate resource use in *Escherichia coli*. *PLoS Genet.* **3**, 0083–0088. (doi:10.1371/journal.pgen.0030015)
 24. Chevin L-M. 2011 On measuring selection in experimental evolution. *Biol. Lett.* **7**, 210–213. (doi:10.1098/rsbl.2010.0580)
 25. Roff DA. 2008 Defining fitness in evolutionary models. *J. Genet.* **87**, 339–348. (doi:10.1007/s12041-008-0056-9)
 26. Orr HA. 2010 Fitness and its role in evolutionary genetics. *Nat. Rev. Genet.* **10**, 531–539. (doi:10.1038/nrg2603)
 27. Kremer CT, Klausmeier CA. 2013 Coexistence in a variable environment: eco-evolutionary perspectives. *J. Theor. Biol.* **339**, 14–25. (doi:10.1016/j.jtbi.2013.05.005)
 28. Yi X, Dean AM. 2013 Bounded population sizes, fluctuating selection and the tempo and mode of coexistence. *Proc. Natl Acad. Sci. USA* **110**, 16 945–16 950. (doi:10.1073/pnas.1309830110)
 29. Ribick N, Lenski RE. 2015 Modeling and quantifying frequency-dependent fitness in microbial populations with cross-feeding interactions. *Evolution* **69**, 1313–1320. (doi:10.1111/evo.12645)
 30. Metz JAJ, Nisbet RM, Geritz SA. 1992 How should we define ‘fitness’ for general ecological scenarios? *Tree* **7**, 198–202.
 31. MacLean RC, Gudelj I. 2006 Resource competition and social conflict in experimental populations of yeast. *Nature* **441**, 498–501. (doi:10.1038/nature04624)
 32. Aiba S, Shoda M, Nagatani M. 1969 Kinetics of product inhibition in alcohol fermentation-temperature effect in the ‘Sake’ brewing. *Biotechnol. Bioeng.* **11**, 1285–1287. (doi:10.1002/bit.260110621)
 33. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. 2009 Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface.* **6**, 187–202. (doi:10.1098/rsif.2008.0172)
 34. Blower SM, Dowlatabadi H. 1994 Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *Int. Stat. Rev.* **62**, 229–243.
 35. Post DM, Palkovacs EP. 2009 Eco-evolutionary feedbacks in community and ecosystem ecology: interactions between the ecological theatre and the evolutionary play. *Phil. Trans. R. Soc. B* **364**, 1629–1640. (doi:10.1098/rstb.2009.0012)
 36. Heino M, Metz JAJ, Kaitala V. 1998 The enigma of frequency-dependent selection. *Tree* **13**, 367–370.
 37. Brommer JE. 2000 The evolution of fitness in life-history theory. *Biol. Rev.* **75**, 377–404. (doi:10.1017/S000632310000551X)
 38. Wu B, Gokhale CS, van Veelen M, Wang L, Traulsen A. 2013 Interpretations arising from Wrightian and Malthusian fitness under strong frequency dependent selection. *Ecol. Evol.* **3**, 1276–1280. (doi:10.1002/ece3.500)
 39. Wiser MJ, Lenski RE. 2015 A comparison of methods to measure fitness in *Escherichia coli*. *PLoS ONE* **10**, 1–11. (doi:10.1371/journal.pone.0126210)
 40. Gerlee P, Lundh T. 2010 Productivity and diversity in a cross-feeding population of artificial organisms. *Evolution* **64**, 2716–2730. (doi:10.1111/j.1558-5646.2010.01020.x)
 41. Pfeiffer T, Bonhoeffer S. 2004 Evolution of cross-feeding in microbial populations. *Am. Nat.* **163**, E126–E135. (doi:10.1086/383593)
 42. Akhmetzhanov AR, Groganard F, Maillet L. 2011 Optimal life-history strategies in seasonal consumer-resource dynamics. *Evolution* **65**, 3113–3125. (doi:10.1111/j.1558-5646.2011.01381.x)
 43. Pekkonen M, Ketola T, Laakso JT. 2013 Resource availability and competition shape the evolution of survival and growth ability in a bacterial community. *PLoS ONE* **8**, 1–12. (doi:10.1371/journal.pone.0076471)
 44. Zeyl C. 2005 The number of mutations selected during adaptation in a laboratory population of *Saccharomyces cerevisiae*. *Genetics* **169**, 1825–1831. (doi:10.1534/genetics.104.027102)
 45. Jasmin J-N, Dillon M, Zeyl C. 2012 The yield of experimental yeast populations declines during selection. *Proc. R. Soc. B* **279**, 4382–4388. (doi:10.1098/rspb.2012.1659)
 46. Paquin CE, Adams J. 1983 Relative fitness can decrease in evolving asexual populations of *S. cerevisiae*. *Nature* **306**, 368–371. (doi:10.1038/306368a0)
 47. Maynard Smith J. 1999 *Evolutionary genetics*. Oxford, UK: Oxford University Press.
 48. Billiard S, Smadi C. 2016 Beyond clonal interference: Scrutinizing the complexity of the dynamics of three competing clones. See <http://arxiv.org/abs/1607.04656>.
 49. Frean M, Abraham ER. 2001 Rock-scissors-paper and the survival of the weakest. *Proc. R. Soc. Lond. B* **268**, 1323–1327. (doi:10.1098/rspb.2001.1670)
 50. Friesen M, Saxer G, Travisano M. 2004 Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in *Escherichia coli*. *Evolution* **58**, 245–260. (doi:10.1111/j.0014-3820.2004.tb01642.x)
 51. Gerstein AC, Otto SP. 2011 Cryptic fitness advantage: diploids invade haploid populations despite lacking any apparent advantage as measured by standard fitness assays. *PLoS ONE* **6**, e26599. (doi:10.1371/journal.pone.0026599)

CHAPITRE 3

Emergence de corrélation entre traits d'histoire de vie et métabolisme au cours de l'adaptation à un environnement saisonnier

Chapitre 3 : Emergence de corrélation entre traits d'histoire de vie et métabolisme au cours de l'adaptation à un environnement saisonnier

I. Article

Emergence of correlations between life-history traits and metabolism during adaptation to a fluctuating environment

A. Introduction

Organisms are characterized by a number of quantitative traits that can be correlated, i.e., that don't vary independently. Relations between two traits can be shaped by different mechanisms such as thermodynamic constraints in metabolic networks leading to resources allocation issues (1), genome architecture including linkage disequilibrium, epistatic interactions or pleiotropy (2). They are the result of the interplay between intracellular physical constraints and past evolutionary history. Here, we considered life-history traits under selection and also other traits such as metabolic traits that may affect or be affected by evolutionary trajectories. For this reason, we used the term "relation" rather than "trade-off" when two traits under selection are negatively related (3).

Despite the different mechanistic origins of these relations (3,4), they have a common consequence: they restrict the phenotypic space and, in this way, they may affect future evolutionary dynamics. The effect of these "ancestral" relations on adaptation has been studied theoretically as well as experimentally (5–8). For example, it is well known that the shape of a trade-off between two life-history traits can impact the outcome of life-history trait evolution (9,10). For instance, in a constant environment with two resources, a generalist strategy will be selected if the trade-off is concave, while a specialist strategy will be selected

if the trade-off is convex (9). Similar results have been found in mathematical modeling studies regarding adaptation in chemostat environments (11–13), in seasonal environment (7) and when the phenotype is based on more than two traits (14). These effects have also been observed experimentally, for instance in the case of the “transmission-virulence” trade-off in viruses or the “mortality-growth rate” trade-off in bacteria (10).

In this work, our aim was to examine whether relations between traits, either life-history or metabolic traits, can emerge from eco-evolutionary dynamics in a seasonal environment and how “ancestral” relations impact this potential emergence. Here, the term “ancestral” is used to refer to any preexisting relation between two traits in contrast with a relation that may emerge from evolutionary or eco-evolutionary dynamics during the course of evolution from the ancestral state.

When measuring relations between traits experimentally, it may prove difficult to disentangle the respective effects of cellular constraints and eco-evolutionary dynamics on the observed relations between traits. In the literature, different experimental settings have been used to study relations between traits. Depending on the experimental design, the observed relation between two traits may be shaped by mechanisms occurring at different scales. For example, at the molecular (DNA) scale, using collections of mutants allowed researchers to understand how mutations can affect trait variations and also trade-offs between traits (15–17). By using instead strains in natural collections (18,19) or obtained through experimental evolution (20–22), only the fittest phenotypes are observed (23–25) and in these case, the relations between traits may be shaped by both ancestral relations and evolutionary dynamics.

Alternatively, to understand how the interplay between “ancestral” constraints and eco-evolutionary dynamics shape relations between traits, one can use mathematical modeling. Here, using such an approach, we focused on the specific context of experimental evolution in successive batch cultures, where eco-evolutionary feedbacks can occur (26). In batch culture, a small fraction of the population is transferred to a new medium at regular time intervals. Contrary to chemostat designs where population growth is kept constant by continuously adding resources, in batches, populations can experience different metabolic phases due to variations of the biotic and abiotic environment. In a previous paper (26), using a mathematical model of yeast population dynamics in batch culture, we have proposed a

decomposition of fitness that allowed us to study eco-evolutionary feedbacks and their potential impacts on trait evolution. In this previous work, we have highlighted that growth rates in fermentation and respiration as well as mortality rate and ethanol resistance are selected while metabolic traits such as the glucose consumption rate are not. In the present work, our aim was to investigate whether a relation between one of these selected traits and one metabolic trait or between two of these selected traits could emerge along the course of evolution in successive batch cultures using an adaptive dynamics modeling framework. We investigated whether the emergence of trade-offs during the course of evolution depends on the “ancestral” relation between traits and in which environmental conditions these emergent relations can be observed.

B. Models and Methods

We used both simulations and mathematical analysis to study the emergence of relations between traits. Our work is based on a model of the population dynamics of n competing yeast strains analyzed in a previous article (26) and described below in a). Here, we used the fitness decomposition previously proposed (see b) to determine the strategies toward which evolution will drive and how that may shape relations between traits (see c). However, in some cases, we were not able to find analytical solutions and addressed the question via simulations of evolutionary paths of strains along successive batch cultures, as described in d.

a. Population dynamics over one batch

We modeled the population dynamics of n competing strains in one batch culture where cells produce biomass via two different metabolic ways the fermentation path and the respiration path. In this model, cells only interact indirectly via competition for resources and production of ethanol which is both a resource and a toxic compound. It is described by a set of differential equations adapted from MacLean and al. (27).

$$\begin{aligned}
\frac{dG}{dt} &= - \sum_i J_{f,i} \frac{G}{K_{f,i}+G} N_i \\
\frac{dN_i}{dt} &= r_{f,i} \frac{G}{K_{f,i}+G} \cdot \exp\left(\frac{-E}{E_{m,i}}\right) \cdot N_i + r_{r,i} \cdot \exp\left(\frac{-E}{E_{m,i}}\right) \cdot \frac{E}{K_{r,i}+E} \cdot \left(\frac{K_{c,i}}{K_{c,i}+G}\right) N_i - m_i \cdot N_i \\
\frac{dE}{dt} &= \sum_i J_{f,i} \frac{G}{K_{f,i}+G} p_i N_i - \sum_i J_{r,i} \frac{E}{K_{r,i}+E} \left(\frac{K_{c,i}}{K_{c,i}+G}\right) N_i
\end{aligned} \tag{1}$$

$$G(0) = 0.15 \text{ g/mL}, E(0) = 0 \text{ g/mL}, \sum_i N_i(0) = 10^6 \text{ cells/mL}$$

$G(t)$, simplified by G , is the glucose concentration at time t , $E(t)$, simplified by E , is the ethanol concentration at time t , and for all i in $(1, \dots, n)$, $N_i(t)$, simplified by N_i , is the cells concentration of strain i at time t . When the glucose concentration is high, the cells consume the glucose by fermentation at a rate $J_{f,i} \cdot G / (K_{f,i} + G)$, where $J_{f,i}$ (g/h) is the maximum glucose consumption rate of strain i and $K_{f,i}$ is the glucose concentration at which the glucose consumption strain i is the half of its maximal glucose consumption, called saturation constant for glucose. During fermentation, the ethanol/glucose yield is denoted p_i (in grams of ethanol produced per gram of glucose consumed) and the yield of cells production is denoted $y_{f,i}$ (in number of cells produced per gram of glucose). When the glucose resource is almost depleted, yeast switch to respiration metabolism and consume the ethanol at a rate $J_{r,i} \cdot E / (K_{r,i} + E)$, where $J_{r,i}$ (g/h) is the maximum ethanol consumption rate and $K_{r,i}$ is the ethanol concentration at which the ethanol consumption is the half of the maximal ethanol consumption, called the saturation constant for ethanol. To model the Crabtree effect, we multiplied the growth rate by the term $K_{c,i} / (K_{c,i} + G)$ which simulates the inhibition of respiration by the glucose and where $K_{c,i}$ (g/mL) is the glucose concentration for which the rate of the ethanol consumption reaches half of its maximum, called the inhibition constant. During respiration the yield of cell production is denoted $y_{r,i}$. Due to the toxicity of ethanol, the yield $y_{f,i}$ is decreased by a factor $\text{tox}(E) = \exp\left(-\frac{E}{E_{m,i}}\right)$ here $E_{m,i}$ denotes the toxic effect of the ethanol (28). The cells' mortality rate is denoted m_i (h^{-1}) and considered constant over the batch. Finally for a strain i , the maximum population growth rate is denoted $r_{f,i} = J_{f,i} \cdot y_{f,i}$ (h^{-1}) in fermentation and $r_{r,i} = J_{r,i} \cdot y_{r,i}$ (h^{-1}) in respiration.

For n strains in competition, the model is therefore composed of $(2+n)$ differential equations with a total set of $10n$ parameters ($J_{f,i}$, $r_{f,i}$, m_i , $E_{m,i}$, p_i , $J_{r,i}$, $r_{r,i}$, $K_{c,i}$, $K_{f,i}$, $K_{r,i}$) where $\mathbf{G(t)}$, $\mathbf{E(t)}$ and $\mathbf{N_i(t)}$ are respectively the glucose (g/mL), ethanol (g/mL) and cells concentration of strain i (cell/mL) at time t .

b. Fitness function

In (26), we proposed a decomposition of the fitness function accounting for the different seasons within one batch culture containing two strains.

Considering an environment with S seasons, we can define $T_{i,j}$ efficient length of season j strain i (see (26)). These season lengths depend on the resources availability during season j but also on strain i . Fitness of strain 1 in competition with strain 2 in a seasonal environment can be written:

$$W_{1/2}(T) = \sum_{j \in \text{season}} (T_{1,j}) \cdot \Delta r_j + \sum_{j \in \text{season}} (r_{2,j}) \cdot \Delta T_j \quad (2)$$

where $r_{2,i}$ is the growth rate of strain 2 during season j , Δr_j is the difference of growth rates in season j and ΔT_j is the difference of efficient season lengths : $T_{1,j} - T_{2,j}$.

As $T_{i,j}$ depends on the strains, the term ΔT_j appears in the fitness function. This term quantifies the ability of a strain to continue to grow on a given resource longer than the other one. This term is therefore influenced by the ability to use resources efficiently at low concentration and/or by the tolerance to toxin.

Two kinds of traits are selected: the growth rates r_j , which appear directly in the fitness function and “transition” traits, which affect the sign of ΔT_j .

In the case of our model, $W_{1/2}(T) = \Delta r_f \cdot T_1^{ferm} + \Delta r_r \cdot T_1^{resp} - \Delta m \cdot T + \Delta T^{ferm} \cdot r_{f,2} + \Delta T^{resp} \cdot r_{r,2}$ (3)

With:

$$T_i^{ferm} = \int_0^T \exp\left(\frac{-E}{E_{m,i}}\right) \cdot \frac{G}{K_{f,i} + G} dt \quad (4)$$

$$T_i^{resp} = \int_0^T \exp\left(\frac{-E}{E_{m,i}}\right) \cdot \frac{E}{K_{r,i} + E} \cdot \frac{K_{c,i}}{K_{c,i} + G} dt \quad (5)$$

In this model, r_f , r_r , m are growth or mortality rates, E_m , K_r , K_f and K_c , are transition traits and y_f , y_r and p are the three metabolic traits which affect season length but are not selected. More details are provided in (26).

c. Existence of evolutionary stable strategies

Our aim was to study whether relations between two independent traits could emerge during the course of evolution. In the case of directional selection for at least one of the traits, that trait will be maximized (or minimized) and no relation can emerge between both traits unless an additional constraint modifies the type of selection. Hence, to study the possible emergence of relations between two traits, one needs to study the existence of a reachable evolutionary stable strategy. Using the fitness decomposition presented above, we studied the existence of an evolutionary stable strategy using the adaptive dynamics framework (29). For this, we first studied the existence of a singular strategy, i.e., a strategy which minimizes or maximizes the fitness function. Then we determined whether this strategy can be reached via an evolutionary path (convergence stability) and whether it can be invaded by a mutant having a different strategy (evolutionary stability). The values of the traits defining this strategy are called singular values of the trait.

We considered batch cultures with one resident and one mutant strain. The frequency of the mutant is close to 0 so that the dynamics of the resources depends only on the resident strain. We considered small mutation effects, so the mutant is close to the resident strain, and only local properties of fitness matter.

i. Case where the phenotype is defined by a single trait

The phenotype is defined by a single trait λ_1 if only λ_1 varies among the population, or if λ_1 and λ_2 vary but a relation ties these two traits which we write in the form $\lambda_2 = f(\lambda_1)$.

A singular strategy is defined by $g = \frac{\partial W_{\text{mut/res}}(T)}{\partial \lambda_{1,\text{mut}}}(\lambda_{1,\text{mut}}) = 0$ when $\lambda_{1,\text{mut}} = \lambda_{1,\text{res}}$.

The evolutionary and convergence stability are determined by the second derivative of fitness according to the mutant traits (denoted E) or according to the resident trait (denoted M) (see (29)).

$$E = \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{1,\text{mut}}^2} \text{ and } M = \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{1,\text{res}} \partial \lambda_{1,\text{mut}}}$$

Eight types of singularity have been described, depending on the signs of E and M. In this work, we consider together the three types of CSS and the three types of “repellor”, we therefore only have four types of singular strategies as described in Table 1.

	E+M<0 : Convergence Stable Strategy	E+M>0 : Convergence Unstable Strategy
E<0 : Evolutionary Stable Strategy	Continuously Stable Strategy (CSS) or optimal strategy	Garden of Eden
E>0: Evolutionary Unstable Strategy	Branching points	"Evolutionary Repellors"

Table 1: Type of singular strategies. *A repellor strategy is a strategy which can be invaded, and cannot be reached by any adaptive trajectories. It corresponds to a minimum of fitness function. A “Garden of Eden” strategy is a strategy that cannot be invaded, but cannot be reached by any adaptive trajectories. So an optimal strategy exists but cannot be reached by adaptation. A branching point is a strategy toward which the adaptive trajectories converges, but it can be invaded. It is invaded by two coexisting strains with divergent phenotypes. Hence, a branching points leads to two different phenotypes which coexist, and can be a step towards sympatric speciation (30).*

The optimal strategy case is the only one in which the trait has one single value at the equilibrium. In the others, the trait is directionally selected. As a consequence, the optimal strategy is the only case in which a correlation can emerge.

ii. Case where the phenotype is defined by two traits

If two traits λ_1 and λ_2 vary independently among the population, a singular strategy is defined by:

$$\begin{cases} \frac{\partial W_{mut/res}(T)}{\partial \lambda_{1,mut}}(\lambda_{1,mut}) = 0 \text{ when } \lambda_{1,mut} = \lambda_{1,res} \\ \frac{\partial W_{mut/res}(T)}{\partial \lambda_{2,mut}}(\lambda_{2,mut}) = 0 \text{ when } \lambda_{2,mut} = \lambda_{2,res} \end{cases}$$

The stability is now determined by the eigenvalues of the Hessian matrix (E) and the Jacobian matrix of the fitness gradient (M).

$$M = \begin{pmatrix} \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{1,res} \partial \lambda_{1,mut}} & \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{2,res} \partial \lambda_{1,mut}} \\ \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{1,res} \partial \lambda_{2,mut}} & \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{2,res} \partial \lambda_{2,mut}} \end{pmatrix} \text{ and } E = \begin{pmatrix} \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{1,mut}^2} & \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{2,mut} \partial \lambda_{1,mut}} \\ \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{1,mut} \partial \lambda_{2,mut}} & \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{2,mut}^2} \end{pmatrix}$$

The two traits define an optimal strategy if the real part of the eigenvalues of E and M+E are negative, otherwise the singular strategy is unstable (14).

C. Modeling evolutionary paths

The model described in B.a predicts the population dynamic of n strains over a single batch. In order to simulate an adaptive trajectory of several successive batch cultures, we added 1) the simulation of successive serial transfers and 2) the simulation of mutations leading to the possible emergence of new phenotypes. To simulate several successive batch cultures, we used the previous model to simulate the dynamics over each batch culture with an initial glucose concentration G_0 , no ethanol and a population of $N_0 = 10^6$ cells sampled in the final population (at $t=T$) of the previous batch. The sampling of the final population of one batch provided the initial frequency of each strain at the beginning of the next batch.

To simulate mutations, at the end of each batch, we randomly selected one cell and modified one or several of its parameters. For each parameter $\theta_{resident}$, the new value θ_{mutant} was chosen as: $\theta_{mutant} = \theta_{resident} + \epsilon$, where ϵ follows a normal centered law of variance σ^2 . For each simulated adaptive trajectory, the initial population of the first batch was a monomorphic population of phenotype Φ_{ref} (see Table 1 for parameter values). A schematic representation of the model is presented in figure S1.

This model allows the coexistence of multiple phenotypes and is easier to simulate than an individual based model. The analysis and parameterization of the one batch ODE model has been done in (26).

D. Results

In this section, we analyze under which conditions a correlation between two independent traits can emerge during adaptation in a seasonal environment using our mathematical model of yeast traits evolution in successive batch cultures. In particular, we study whether the emergence of a correlation during the course of evolution depends on the type of traits studied (metabolic traits, life-history traits), the existence of ancestral trade-off and eco-evolutionary feedbacks. According to a previous study of the one batch-model (26), a strain has three metabolic traits (glucose and ethanol consumption rate (J_f and J_r) and ethanol production p) and four life-history traits (fermentation growth rate r_f , respiration growth rate r_r , mortality rate m and ethanol tolerance E_m). Other traits have little effect in the range of parameters we used.

For the sake of simplicity, we focused on the case where only two independent traits can vary by mutation although other traits, linked to one of these two independent traits can also vary due to ancestral relations such as pleiotropic mutations, resource allocation trade-offs or physical constraints. Hence, the phenotype of each strain is described by the value of two independent traits, each being either a metabolic (non-selected) or a life-history (selected) trait.

All combinations of the two traits are assumed to be “physiologically” possible. However, these phenotypes may not be equally fit and, therefore, only some of them will be observed after adaptation, potentially leading to the emergence of correlations between those traits in evolved populations. To determine whether a correlation can emerge between two specific traits, we examined the existence of singular strategies and their types for three different cases: case a) two metabolic traits may vary independently, case b) one metabolic trait and one life-history trait may vary independently; case c) two life-history traits may vary independently.

In addition, in section d, we examined the impact of the variations of the environment on the emergence of relations between two traits.

a. Case (a): Two independent metabolic traits

In this case, we assumed that only two metabolic traits among (J_f , J_r or p) vary independently while other traits are fixed. As metabolic traits are not selected, the relative fitness of each strain compared to any other is null. As a consequence, evolution is only driven by drift. In this case, no correlation emerges between the two independently metabolic traits. All the strategies are singular (derivative of fitness is null), but none are optimal.

b. Case (b): One metabolic trait and one life-history trait vary independently

First, we assumed that one metabolic trait (either J_f , J_r or p) and one life-history trait (either r_f , r_r , m or E_m) vary by mutation independently with no ancestral trade-offs between both traits. As shown in Table 2, the derivative of fitness with respect to each of the four life-history trait is never null, indicating that each life-history trait is under directional selection. Hence, if there is no trade-off between life-history traits preventing directional selection, the considered trait is selected in one direction unless it reaches a physiological limit (3). For growth rates and ethanol tolerance, the derivative is positive, meaning they will increase whereas a decrease of the mortality rate is selected. Consequently, if only one life-history trait and one metabolic trait vary by mutation independently, no correlation can emerge between them.

Table 2: Derivative of fitness when only one life-history varies

trait	derivative of fitness, g	Traits which influence g
r_f	$\frac{\partial W_{1/2}(T)}{\partial r_f}(r_f) = T_{\text{ferm}}$	J_f, p
r_r	$\frac{\partial W_{1/2}(T)}{\partial r_r}(r_r) = T_{\text{resp}}$	J_f, J_r, p
M	$\frac{\partial W_{1/2}(T)}{\partial m}(m) = -T$	$\emptyset$
E_m	$\frac{\partial W_{1/2}(T)}{\partial E_{m,1}}(E_m) = r_{\text{ferm}} \frac{\partial T_{\text{ferm}}}{\partial E_m} + r_{\text{resp}} \frac{\partial T_{\text{resp}}}{\partial E_m}$ (the two terms are positive)	J_f, J_r, p

The maximization of each trait if they are independent is expected. It corresponds to the “Darwinian Daemon”. However, it could be prevented by the presence of ancestral trade-off (31). Hence, we re-examined the question assuming that another life-history trait varies due to a trade-off between two life-history traits. In this case, three traits can vary in the population: one metabolic trait and two life-history traits. However, as the two life-history traits are linked by a trade-off, the whole phenotype can still be described by one life-history trait and one independent metabolic trait.

Let's denote f the function describing the relationship between the two life-history traits λ_1 and λ_2 . The derivative of fitness with respect to each trait can be calculated depending on the trade-off (Table 3). For each trait λ_1 , the solution of the equation $\frac{\partial W_{mut/res}(T)}{\partial \lambda_{1,mut}}(\lambda_{1,mut}) = 0$ provides the singular trait value, if it exists. This equation depends on season length and their derivative. In most cases, this equation can't be solved analytically as it depends on season lengths and their derivative with respect to the trait for which we have no analytical solution (see Table 3 and the example below). However, if a singular strategy exists, it will verify this relation. Note also that, season lengths depend on the phenotypes of competing strains and consequently, singular values of the life-history traits can depend on the metabolic trait.

Table 3: singular condition for each couple of trait.

trait 1	trait 2	derivative of fitness	singularity condition
r_f	r_r	$\frac{\partial W_{1/2}(T)}{\partial r_f}(r_f, f(r_f)) = T^{ferm} + f'(r_f)T^{resp}$	$-f'(r_f) = \frac{T^{ferm}}{T^{resp}}$
r_f	m	$\frac{\partial W_{1/2}(T)}{\partial r_f}(r_f, f(r_f)) = T^{ferm} - f'(r_f)T$	$f'(r_f) = \frac{T^{ferm}}{T}$
r_f	E_m	$\begin{aligned} & \frac{\partial W_{1/2}(T)}{\partial r_f}(r_f, f(r_f)) \\ &= T^{ferm} + f'(r_f) \left(r_f \frac{\partial T^{ferm}}{\partial E_m} + r_r \frac{\partial T^{resp}}{\partial E_m} \right) \end{aligned}$	$\begin{aligned} & -f'(r_f) = \\ & \frac{T^{ferm}}{\left(r_f \frac{\partial T^{ferm}}{\partial E_m} + r_r \frac{\partial T^{resp}}{\partial E_m} \right)} \end{aligned}$
r_r	m	$\frac{\partial W_{1/2}(T)}{\partial r_r}(r_r, f(r_r)) = T^{resp} - f'(r_r)T$	$f'(r_r) = \frac{T^{resp}}{T}$
r_r	E_m	$\begin{aligned} & \frac{\partial W_{1/2}(T)}{\partial r_r}(r_r, f(r_r)) \\ &= T^{resp} + f'(r_r) \left(r_f \frac{\partial T^{ferm}}{\partial E_m} + r_r \frac{\partial T^{resp}}{\partial E_m} \right) \end{aligned}$	$\begin{aligned} & -f'(r_r) = \\ & \frac{T^{resp}}{\left(r_f \frac{\partial T^{ferm}}{\partial E_m} + r_r \frac{\partial T^{resp}}{\partial E_m} \right)} \end{aligned}$
m	E_m	$\begin{aligned} & \frac{\partial W_{1/2}(T)}{\partial m}(m, f(m)) \\ &= f'(m) \left(r_f \frac{\partial T^{ferm}}{\partial E_m} + r_r \frac{\partial T^{resp}}{\partial E_m} \right) - T \end{aligned}$	$\begin{aligned} & f'(m) = \\ & \frac{T}{\left(r_f \frac{\partial T^{ferm}}{\partial E_m} + r_r \frac{\partial T^{resp}}{\partial E_m} \right)} \end{aligned}$

We investigated the existence of a singular value for a life-history trait and the relation between this singular value and a metabolic trait. As mentioned in the methods section, the stability of the equilibrium, the possibility of coexistence, and the possibility of branching are given by the sign of $E = \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{mut}^2}$ and $M = \frac{\partial^2 W_{1/2}(T)}{\partial \lambda_{res} \partial \lambda_{mut}}$ (29) which can be derived for each trade-off function f (Table 4).

Table 4: Second derivative of fitness.

trait 1	trait 2	E	M
r_f	r_r	$\frac{\partial^2 W_{1/2}(T)}{\partial r_f^2}(r_f) = f''(r_f)T^{resp}$	$\frac{\partial W_{1/2}(T)}{\partial r_{f,res} \partial r_f}(r_f) = \frac{\partial T^{ferm}}{\partial r_{f,res}} + f'(r_f) \frac{\partial T_{resp} T^{resp}}{\partial r_{f,res}}$
r_f	m	$\frac{\partial^2 W_{1/2}(T)}{\partial r_f^2}(r_f) = -f''(r_f)T$	$\frac{\partial W_{1/2}(T)}{\partial r_{f,res} \partial r_f}(r_f) = \frac{\partial T^{ferm}}{\partial r_{f,res}}$
r_r	m	$\frac{\partial^2 W_{1/2}(T)}{\partial r_r^2}(r_r) = -f''(r_r)T$	$\frac{\partial W_{1/2}(T)}{\partial r_{r,res} \partial r_r}(r_r) = \frac{\partial T^{resp}}{\partial r_{r,res}}$
r_f	E_m	$\frac{\partial^2 W_{1/2}(T)}{\partial r_f^2}(r_f)$ $= r_f \frac{\partial^2 T^{ferm}}{\partial r_f^2} + r_r \frac{\partial^2 T^{resp}}{\partial r_f^2} + 2 \cdot \frac{\partial T^{ferm}}{\partial r_f}$	$\frac{\partial^2 W_{1/2}(T)}{\partial r_{f,res} \partial r_f}(r_f)$ $= \frac{\partial T^{ferm}}{\partial r_{f,res}} + r_f \frac{\partial^2 T^{ferm}}{\partial r_{f,res} \partial r_f} + r_r \frac{\partial^2 T^{resp}}{\partial r_{f,res} \partial r_f}$
r_r	E_m	$\frac{\partial^2 W_{1/2}(T)}{\partial r_r^2}(r_r)$ $= r_f \frac{\partial^2 T^{ferm}}{\partial r_r^2} + r_r \frac{\partial^2 T^{resp}}{\partial r_r^2} + 2 \cdot \frac{\partial T^{resp}}{\partial r_r}$	$\frac{\partial^2 W_{1/2}(T)}{\partial r_{r,res} \partial r_r}(r_r)$ $= \frac{\partial T_{resp}}{\partial r_{r,res}} + r_f \frac{\partial^2 T^{ferm}}{\partial r_{r,res} \partial r_r} + r_r \frac{\partial^2 T^{resp}}{\partial r_{r,res} \partial r_r}$
m	E_m	$\frac{\partial^2 W_{1/2}(T)}{\partial m^2}(m)$ $= r_f \frac{\partial^2 T^{ferm}}{\partial m^2}$ $+ r_r \frac{\partial^2 T^{resp}}{\partial m^2}$	$\frac{\partial^2 W_{1/2}(T)}{\partial m_{res} \partial m}(m)$ $= r_f \frac{\partial^2 T^{ferm}}{\partial m_{res} \partial m} + r_r \frac{\partial^2 T^{resp}}{\partial m_{res} \partial m}$

We considered separately the case of trade-offs involving growth or mortality rates (r_f , r_r , or m) and those involving transition terms (E_m).

For trade-offs involving only growth rates, it can be noticed that:

- M could be always negative (see proof in SI 2), at least under biologically reasonable condition, meaning that coexistence is possible near the equilibrium.
- The sign of E is determined by the second derivative of f (Table 4). The singular strategy is a branching point if E is positive, i.e, if f is convex.
- As M is always negative, the sign of $M+E$ (and consequently the stability of the singular value (32)) is driven by the value of E , i.e., by the second derivative of f :
 - If f is concave, then E is negative and $M+E$ is negative. So according to table 1, the equilibrium is convergent and evolutionary stable, so this strategy is an optimal strategy
 - If f is convex, then E is positive. So according to the table 1, the singular strategy can be a “repellor” or a branching point, depending on the sign of $M+E$. If $M < -E < 0$, the singular strategy is a branching point. As E is proportional to f' , this condition is true if f' is close to 0.

As M is always negative, the metabolic trait can affect the value of M but not its sign. Similarly, the sign of E is determined by the intracellular constraint, so a metabolic trait can affect the value of E . Hence, it can affect the value of the optimal strategy, but cannot destabilize it. A relation can emerge only if an optimal strategy exists. So a relation can emerge if the trade-off is concave.

If the trade-off involves mortality rate, the behaviour is similar. But as mortality is a « neagitive growth rate », the sign of E changes: if f is concave, E is positive ; and if f is convex, E is negative.

For trade-offs involving at least one transition trait (in our model E_m), the sign of E and M depends on the environment but cannot be determined analytically. As a consequence, we cannot determine a condition for the existence of optimal strategy and therefore emergence of a relation.

As the metabolic trait is not selected and is not related to a selected trait, its evolution is only driven by drift, meaning that it can take any possible value in a biologically realistic range in the adapted population. The metabolic trait affects the season lengths, and as a consequence, affects the optimal values of the selected traits. Therefore, different optimal values of the life-history trait will be observed in each population, depending on the values of the metabolic trait. This leads to the emergence of a correlation between the metabolic and the selected

trait in adapted populations, through eco-evolutionary feedback loops. The shape of the relation can be determined using a graphical method as shown in the example below.

The derivative of the fitness with respect to the life-history-trait depends on the season lengths (Table 3). These lengths and their derivatives can be computed numerically for each phenotype. The singular strategies can then be plotted on a graph representing the life-history trait as a function of the metabolic trait. To determine the direction of the trait evolution, we examined the sign of the fitness derivative presented in Table 3. To illustrate this, we focused on the case where the glucose consumption rate J_f (a metabolic trait) and the fermentation growth rate r_f (a life-history-trait) can vary independently due to mutations and where the mortality rate m is related to the fermentation growth rate r_f by a convex positive trade-off arbitrarily chosen as: $m = f(r_f) = 2.32r_f^5 + 0.02$. The whole phenotype is therefore described by J_f and r_f . For each phenotype (J_f, r_f) , the fermentation season length is computed numerically. The fitness function can be written as: $W_{1/2}(T) = \Delta r_f \cdot T_1^{ferm} - \Delta f(r_f) \cdot T$ (7) and one can derive from this an equation for the singular strategy: $T_1^{ferm} - f'(r_{f,opt}) \cdot T = 0 \Leftrightarrow f'(r_f) = \frac{T_1^{ferm}}{T}$ (8). In this case, a singular strategy exists as f is an increasing function. As f is convex and growth rate and mortality rates are involved, the singular strategy is an optimal strategy. Figure 1 shows the singular strategies depending on the value of the glucose consumption rate for this example and several possible evolutionary paths. In this example, the existence of an ancestral intracellular trade-off between the growth rate r_{ferm} and the mortality rate m leads to an optimal value of r_{ferm} related to the glucose consumption rate J_f . Hence, for several strains evolved independently in the same environment, a relation could emerge between both traits. The relation is calculated analytically in a simplified case provided in SI.

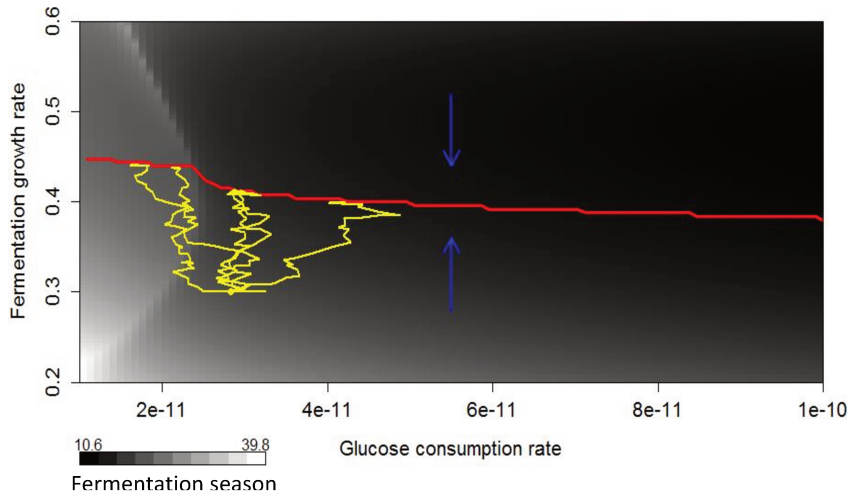


Figure 1 : Determination of the optimal values of the growth rate when i) growth, mortality and glucose consumption rates vary and ii) fermentation growth rate and mortality rate are linked by an ancestral constraint (case 2). Each square indicates the fermentation season length (gray scale) for different combinations of glucose consumption rate (J_f) and fermentation growth rate (r_f). The red line indicates the optimal values of the growth rate (r_f). The direction of the selection gradient is given by the blue arrows. The yellow lines represent five simulations of evolutionary paths starting from one ancestral population with phenotype ($J_f=3.10^{-11}, r_f=0.25$).

c. Case (c): Two life-history traits vary independently

Finally, we examined whether a correlation between two life-history traits varying independently can emerge during adaptation. The two life history traits can be taken among growth rates, mortality rate or ethanol tolerance (r_f , r_r , m or E_m).

Again, to address this question we studied the existence of singular strategies and their impact on relations between varying traits.

As shown in b, for one given life-history trait, the existence of a singular strategy is conditioned to the existence of an intracellular trade-off with another life-history trait. Hence, we considered the case where a third life-history trait (λ_3) is related to the two other independently varying traits (λ_1 and λ_2) at the intracellular level via one constraint such as

$\lambda_3 = g(\lambda_1, \lambda_2)$. Again, the phenotype of one given strain can be described by the values of the two independently varying traits.

As in part b, for each life-history trait, we identified singular values. As shown in Table 3, the singular values of each life-history-trait depend on season lengths, which themselves depends on other life history traits. Contrary to metabolic traits, the other varying life-history-trait cannot take any possible value at the equilibrium. Hence, the singular phenotypes are the one which verify the equations presented in Table 3 for the two considered life-history traits. Furthermore, the number of singular phenotypes is finite.

The stability of the singular strategies depends on the eigenvalues of the Hessian and Jacobian matrix (14,33) which we have solved below for one example. In this example, the respiration growth rate r_{resp} , and the fermentation growth rate r_{ferm} vary independently and, at the intracellular level, the mortality rate m depends on both growth rates via the function $m = f(r_f, r_r) = 0,91 \cdot r_f^3 + 0,59 \cdot r_r^2 - 0,02$. The batches contain 0.15g/mL of glucose initially, and last $T=48h$.

In order to find the singular strategy, the following system has to be solved:

$$\begin{cases} \frac{\partial W_{mut/res}(T)}{\partial r_f}(r_f, m) = 0 \\ \frac{\partial W_{mut/res}(T)}{\partial r_r}(r_r, m) = 0 \end{cases} \text{ which leads to } \begin{cases} r_{f,opt} = \sqrt{\frac{T_f}{2.73T}} \\ r_r = \frac{T_r}{1.18T} \end{cases} \quad (9)$$

The singular strategy can be computed and is equal to ($r_f=0.33$, $r_r=0.20$).

As there are two selected traits, E and M are now matrices. The stability of the equilibrium can be studied by looking at the sign of the eigenvalues of E and M where:

$$E = \begin{pmatrix} -1.80T & 0 \\ 0 & -1.18T \end{pmatrix} = \begin{pmatrix} -86.4 & 0 \\ 0 & -56.64 \end{pmatrix}.$$

$$M = \begin{pmatrix} \frac{\partial T_{ferm}}{\partial r_{ferm,res}} & \frac{\partial T_{ferm}}{\partial r_{resp,res}} \\ \frac{\partial T_{resp}}{\partial r_{ferm,res}} & \frac{\partial T_{resp}}{\partial r_{resp,res}} \end{pmatrix} = \begin{pmatrix} -25.4 & 8.62 \\ -1.34 & -24.7 \end{pmatrix}. \text{ Hence, } M+E = \begin{pmatrix} -111.8 & 8.62 \\ -1.34 & -81.34 \end{pmatrix}.$$

The matrices have been calculated numerically. As E is a diagonal matrix, its eigenvalues are the diagonal terms which are both negative (-86.4 and -56.64). Hence the equilibrium is evolutionary stable and branching is not possible. In addition, as the eigenvalues of M+E are

negative (-111.4 and -81.72), the equilibrium is convergent and evolutionary stable. The singular strategy is therefore an optimal strategy, represented in red in Figure 2. The sign of the fitness derivative of each trait indicates the direction of the selection. As predicted, according to the selection gradient, this equilibrium is stable. We simulated the evolution of 10 different ancestral populations over 300 successive 48 hours-batch cultures, and they all converged toward the predicted equilibrium.

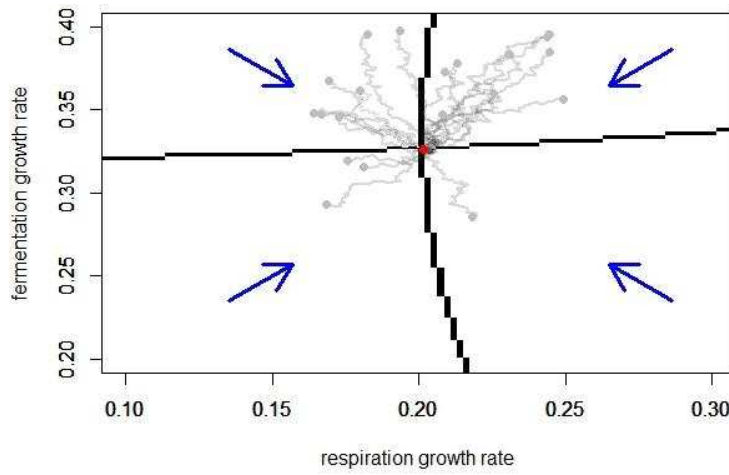


Figure 2: Determination of the optimal phenotype when i) mortality, fermentation and respiration growth rates vary and ii) the three traits are linked by an intracellular constraint (case 3). The vertical black line represents the optimal values of the respiration growth rate (r_r) and the horizontal black line represents the optimal values for the fermentation growth rate (r_f). The red point indicates the optimal phenotype. The direction of the selection gradient is given by the blue arrows. The grey lines represent 20 simulations of evolutionary paths. The grey points represent the different ancestral populations.

d. Effect of the environment

For all the scenarios examined in section a. to c., when an optimal strategy existed, the values of the optimal phenotypes depended on season lengths. The season lengths can be modified either by eco-evolutionary feedbacks or by experimental conditions. Consequently, we can distinguish different mechanisms and conditions that can lead to the emergence of correlation between two traits varying independently.

We have shown in previous parts that a correlation between a selected and a metabolic trait can emerge when ancestral populations evolved in the same environment (same initial concentration of resources at the beginning of each batch, same initial population density at the beginning of each batch, same batch length). In this section, we examined the possible emergence of a correlation between two traits when ancestral populations evolved in different environments.

Season lengths will vary from one experimental condition to another depending on the initial concentration of resources, the population density or the batch length. Hence, even if there is only one optimal phenotype in each environment, this equilibrium depends on experimental conditions. In example c), if populations are cultivated using different batch lengths, equation (9) shows that the optimal phenotype will be different in each condition: the longer the batch is, the smaller are the optimal values of r_f and r_r (Figure 3). As a consequence, by observing the evolved populations, the two life-history traits r_f and r_r are correlated.

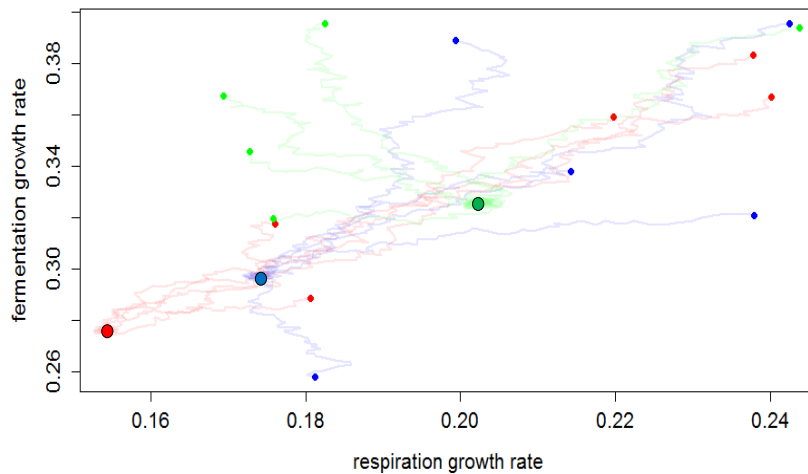


Figure 3 : Adaptive trajectories in environment with different batch lengths, when i) mortality, fermentation and respiration growth rates vary and ii) the three traits are linked by an intracellular constraint (case 3). *The colored lines represent five simulations of evolutionary path in each environment. Colors indicate batch lengths (red=72h, blue=60h and green=48h). The bullet points represent the different ancestral populations. The circled points represent the optimal phenotypes.*

In case b), where there are an infinite number of strategies distributed along a line, variations of the environment will also shift the optimal strategies as shown in Figure 4. Lastly, if several populations were evolved in various environments, a relation would be observed within environment.

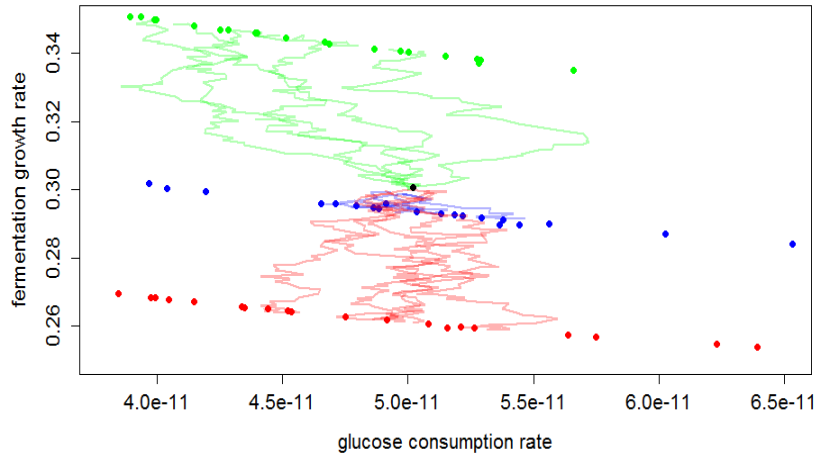


Figure 4 : Adaptive trajectories in environment with different batch lengths, when i) growth, mortality and glucose consumption rates vary and ii) fermentation growth rate and mortality rate are linked by an intracellular constraint (case 2). *The colored lines represent five simulations of evolution in each environment. Colors indicates batch lengths (red=72h, blue=60h and green=48h). The black points represent the ancestral populations. The points represent the phenotypes of 20 populations after 300 batches in each environment.*

E. Discussion

In this paper, we examined when and why a correlation between two traits emerge along the course of evolution in successive batch cultures. A correlation will be observed if the optimal value of a trait depends on the values of another one. Therefore, using a mathematical modeling approach, we first studied the condition in which an optimal value exists, and then what affects this optimal value.

Existence and determination of the optimum. An optimal strategy is a strategy which is evolutionary stable (once it has invaded it cannot be defeated), and convergence stable, i.e. evolutionary trajectories converge toward this strategy. Three conditions are necessary for a trait to have an optimal value. First the trait must be selected. Second the selected trait has

to be linked to another one by a trade-off to observe a singular strategy, otherwise the evolution will lead to Darwinian Daemon (31). This strategy depends on the ancestral trade-off (f') and the environmental constraints, via season lengths (Table 3). Finally, the stability of a singular strategy depends on the shape of the internal constraints, and the way environment is modified by the strains (Table 4). In our model, it is possible to study further the stability when growth rates or mortality rate are considered.

If we only consider growth rates (with fixed transition traits), we have shown that an optimal strategy may exist only if the trade-off is concave. This property has been shown in the case of a constant environment (9), in the case of a chemostat (12) and had been tested by simulations in the case of evolution in successive batches (34). Moreover, if the feedback of the strains on the environment is always deleterious to the strains, or if the environment is fixed ($M \leq 0$), this condition becomes sufficient, and an optimal strategy exists if and only if the trade-off is concave. In our simulation, M is always negative, but have not proved it is always so.

If transition traits are considered, the condition cannot be determined analytically. The concavity of the trade-off curve is no longer necessary; as an example, in (7,34) optimal strategies are observed with a convex trade-off when transition traits are under selection.

- a. Optimal values of the traits are affected by season lengths and therefore traits which affect those lengths.

Season lengths (and consequently optimal values of traits) are affected by three quantities: the values of life-history traits, the values of metabolic traits, and the values of environmental variables (such as glucose concentration or the duration of the batch). If only life-history traits evolve the optimal strategy is unique. In (35), the eighteen populations which evolved in the same environment converged to similar phenotypes. We have shown that, theoretically, that the emergence of a correlation between a metabolic and life-history trait is possible. By studying the adapted phenotype of several lineages in the same environment, we should observe a correlation. But this correlation should not be observed within the population. In the Long Term Experimental Evolution, the yield and the growth rate of twelve adapted

population are measured (20). A positive correlation between the two traits is expected (see SI). However, a negative correlation between the two traits is observed within the population. This negative correlation constrains the possible phenotypes, and could therefore restrict the optimal strategy to a single value. In fact, the emergence of a correlation relies on the assumption that traits are independent. However, this assumption is quite rare in biological systems. For instance, in yeast, the consumption rates are correlated with growth rates (15,36). Environmental variables can affect the values of the equilibrium, and a correlation is therefore expected between the adapted phenotype of different strains which evolved in different environments. A correlation between the growth rate and the yields of 10 strains evolved in different natural environment has been observed (35). As explained above growth rates and yields defined a unique optimal strategy (20), but this strategy depends on the resources abundance of the environment. In high-resources-level environment, growth rate is lower than in low-resources-level environment.

b. Adding constraints

The main limiting assumption in our model is the independence of the traits. The traits are not independent in living organisms and many relation between traits have been documented (36). In *S. cerevisiae*, we found 11 relevant relations in the literature (see Table 1-1 in the Introduction chapter of the thesis). For instance, growth rate is a well-studied trait and is correlated with yield (22,37), ethanol tolerance (16,19), mortality rate (10), and glucose consumption rate (15). As a consequence, emergence of a relation is not expected during an experimental evolution.

If more than two selected traits are linked, the condition verified by the optimal values of the traits will be modified, as well as the stability condition. If metabolic traits are linked to selected traits, the conditions will remain the same, but the stability can be affected.

c. Diversification and coexistence

In this paper we only considered the case of an optimal strategy, in which case, at equilibrium, coexistence is not possible. However, studying coexistence and branching could allow to

better understand diversity. In our model, diversity can be high. For instance when the trade-off is linear (for instance $m=a.r_f+b$) the condition verified by the optimal strategy depends only on the environment ($a=T_f/T$). In this case, the evolution of the system stops as soon as the environment verifies the condition, disregarding the composition of the population. Therefore, several strains can coexist, with very different phenotypes. We did not study the case of unstable strategies either. This occurs notably for a convex trade-off between two growth rates. The sign of $E+M$ is critical in this case and allows switching between branching points and a “repellor”. The condition of a convex trade-off had been tested by (12). This condition is general to every seasonal environment, as soon as fitness can be written as we do.

If metabolism changes, i. e the values of metabolic traits, the environmental conditions will change (i.e., the values of $M+E$) and the system can switch from branching to “repellor”. The coexistence allowed by the branching configuration will stop, and one strain can be eliminated. The switch between the two situations could partially explain the observation made by Rocabert et al. (38). Starting with a situation of two strains coexisting after a branching, one of the two strains is eliminated after a few evolutionary steps in 50% of the batch simulations. As the metabolic traits are modified, the system could switch from branching to “repellor”. This explanation is only partial, as in their system the number of resources increases, going beyond the validity regime of our model. This example shows that coexistence in a seasonal environment is interesting and applying our method to study such cases could lead to interesting results.

Reference:

1. MacLean RC. The tragedy of the commons in microbial populations: insights from theoretical, comparative and experimental studies. *Heredity (Edinb)* [Internet]. 2008;100(3):233–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17957201>
2. Zhang J, Wagner GP. On the definition and measurement of pleiotropy. *Trends Genet* [Internet]. 2013;29(7):383–4. Available from: <http://dx.doi.org/10.1016/j.tig.2013.05.002>

3. Roff DA, Fairbairn DJ. The evolution of trade-offs: Where are we? *J Evol Biol.* 2007;20(2):433–47.
4. Antonovics J, van Tienderen PH. Ontoecogenophyloconstraints? The chaos of constraint terminology. *Trends Ecol Evol.* 1991;6(5):166–8.
5. Mazancourt C De, Dieckmann U. Trade-Off Geometries and Frequency-Dependent Selection. 2004;164(6).
6. Kjørboe T. How zooplankton feed: Mechanisms, traits and trade-offs. *Biol Rev.* 2011;86(2):311–39.
7. Kremer CT, Klausmeier CA. Coexistence in a variable environment: Eco-evolutionary perspectives. *J Theor Biol.* 2013;339:14–25.
8. Edwards KF, Klausmeier CA, Litchman E. A Three-Way Trade-Off Maintains Functional Diversity under Variable Resource Supply. *Am Nat.* 2013;182(6).
9. Levins R. Theory of Fitness in a Heterogeneous Environment. I. The Fitness Set and Adaptive Function. *Am Nat.* 1962;96(891):361.
10. Maharjan R, Nilsson S, Sung J, Haynes K, Beardmore RE, Hurst LD, et al. The form of a trade-off determines the response to competition. *Ecol Lett.* 2013;16(10):1267–76.
11. Gudelj I, Beardmore RE, Arkin SS, MacLean RC. Constraints on microbial metabolism drive evolutionary diversification in homogeneous environments. *J Evol Biol.* 2007;20(5):1882–9.
12. Doebeli M. A model for the evolutionary dynamics of cross-feeding polymorphisms in microorganisms. *Popul Ecol.* 2002;44(2):59–70.
13. Champagnat N, Jabin P-E, Méléard S. Adaptation in a stochastic multi-resources chemostat model. *J Math Pures Appl [Internet].* 2014 Jun 1 [cited 2018 May 5];101(6):755–88. Available from: <https://www.sciencedirect.com/insb.bib.cnrs.fr/science/article/pii/S002178241300127X>
14. Débarre F, Nuismer SL, Doebeli M. Multidimensional (Co)Evolutionary Stability. *Am Nat [Internet].* 2014;184(2):158–71. Available from:

<http://www.journals.uchicago.edu/doi/10.1086/677137>

15. Elbing K, Larsson C, Bill RM, Albers E, Snoep JL, Boles E, et al. Role of Hexose Transport in Control of Glycolytic Flux in *Saccharomyces cerevisiae* Role of Hexose Transport in Control of Glycolytic Flux in *Saccharomyces cerevisiae*. *Appl Environ Microbiol*. 2004;70(9):5323–30.
16. Zakrzewska A, van Eikenhorst G, Burggraaff JEC, Vis DJ, Hoefsloot H, Delneri D, et al. Genome-wide analysis of yeast stress survival and tolerance acquisition to analyze the central trade-off between growth rate and cellular robustness. *Mol Biol Cell* [Internet]. 2011;22(22):4435–46. Available from: <http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E10-08-0721>
17. Roop JI, Chang KC, Brem RB. Polygenic evolution of a sugar specialization trade-off in yeast. *Nature* [Internet]. 2016;530(7590):336–9. Available from: <http://www.nature.com/doi/10.1038/nature16938>
18. Spor A, Nidelet T, Simon J, Bourgaïs A, de Vienne D, Sicard D. Niche-driven evolution of metabolic and life-history strategies in natural and domesticated populations of *Saccharomyces cerevisiae*. *BMC Evol Biol*. 2009;9:296.
19. Antocea AO, Nămolesanu IC, Radoi FM. Comparative study regarding the ethanol resistance of some yeast strains isolated from Romanian vineyards. *Rom Biotechnol Lett*. 2011;16(2):5981–8.
20. Novak M, Pfeiffer T, Lenski RE, Sauer U, Bonhoeffer S. Experimental Tests for an Evolutionary Trade-Off between Growth Rate and Yield in *E. coli*. *Am Nat*. 2006;168(2):242–51.
21. Samani P, Bell G. Experimental evolution of the grain of metabolic specialization in yeast. *Ecol Evol*. 2016;6(12):3912–22.
22. Jasmin J-N, Dillon M, Zeyl C. The yield of experimental yeast populations declines during selection. *Proc Biol Sci*. 2012;279:4382–8.
23. Sinervo B, Svensson E. Correlational selection and the evolution of genomic architecture. *Heredity (Edinb)* [Internet]. 2002;89(5):329–38. Available from:

<http://www.nature.com/doi/10.1038/sj.hdy.6800148>

24. McGlothlin JW, Parker PG, Nolan V, Ketterson ED. Correlational selection leads to genetic integration of body size and an attractive plumage trait in dark-eyed juncos. *Evolution*. 2005;59(3):658–71.
25. Forsman A, Appelqvist S. Visual predators impose correlational selection on prey color pattern and behavior. *Behav Ecol*. 1998;9(4):409–13.
26. Collot D, Nidelet T, Ramsayer J, Martin OC, Dillmann C, Sicard D, et al. Feedback between environment and traits under selection in a seasonal environment : consequences for experimental evolution. *Proc R Soc B*. 2018;
27. MacLean RC, Gudelj I. Resource competition and social conflict in experimental populations of yeast. *Nature*. 2006;441:498–501.
28. Aiba S, Shoda M, Nagatani M. Kinetics of product inhibition in alcohol fermentation- Temperature effect in the “Sake” brewing. *Biotechnol Bioeng*. 1969;11:1285–7.
29. Geritz SAH, Kisdi É, Meszén G, Metz JAJ. Evolutionarily Singular Strategies and the Adaptive Growth and Branching of the Evolutionary Tree. *Evol Ecol*. 1998;12:35–57.
30. Doebeli M, Dieckmann U. Evolutionary Branching and Sympatric Speciation Caused by Different Types of Ecological Interactions. *Am Nat* [Internet]. 2000 Oct [cited 2018 May 5];156(S4):S77–101. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29592583>
31. Stearns SC. Trade-Offs in Life-History Evolution. *Funct Ecol*. 1989;3:259–68.
32. Eshel I. Evolutionary and continuous stability. *J Theor Biol* [Internet]. 1983 Jul 7 [cited 2018 May 5];103(1):99–111. Available from: <https://www.sciencedirect.com/science/article/pii/0022519383902011>
33. Ito H, Sasaki A. Evolutionary branching under multi-dimensional evolutionary constraints. *J Theor Biol* [Internet]. 2016;407:409–28. Available from: <http://dx.doi.org/10.1016/j.jtbi.2016.07.011>
34. Friesen M, Saxer G, Travisano M, M. D. Experimental Evidence for Sympatric

- Ecological Diversification Due To Frequency-Dependent Competition in *Escherichia Coli*. *Evolution* (N Y). 2004;58(2):245–60.
35. Spor A, Kvitek DJ, Nidelet T, Martin J, Legrand J, Dillmann C, et al. Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. *Evolution* (N Y). 2014;68(3):772–90.
36. Ferenci T. Trade-off Mechanisms Shaping the Diversity of Bacteria. *Trends Microbiol* [Internet]. 2016;24(3):209–23. Available from: <http://dx.doi.org/10.1016/j.tim.2015.11.009>
37. Zakhartsev M, Yang X, Reuss M, Pörtner HO. Metabolic efficiency in yeast *Saccharomyces cerevisiae* in relation to temperature dependent growth and biomass yield. *J Therm Biol* [Internet]. 2015;52:117–29. Available from: <http://dx.doi.org/10.1016/j.jtherbio.2015.05.008>
38. Rocabert C, Knibbe C, Consuegra J, Schneider D, Beslon G. Beware batch culture: Seasonality and niche construction predicted to favor bacterial adaptive diversification. Vol. 13, *PLoS Computational Biology*. 2017. 1-32 p.

CHAPITRE 4

Caractérisation de souches dans un
environnement saisonnier par leurs traits
d'histoire de vie et leur fitness

Chapitre 4 : Caractérisation de souches dans un environnement saisonnier par leurs traits d'histoire de vie et leur fitness

L'objectif de ce chapitre est d'analyser les traits d'histoire de vie et la valeur sélective des souches de levure ayant évolué dans quatre milieux de « batchs successifs » différant par la longueur du batch (48h ou 96h) et par la quantité de glucose disponible (1% ou 15%). Deux types de données étaient disponibles pour les souches évoluées (et/ou leurs ancêtres) : des données en monoculture (voir Chapitre 2 et (Spor 2009; Spor et al. 2014)) ainsi que des données de compétitions entre certaines des souches ancêtres utilisées dans (Spor et al. 2014) menées par Carine Remoué (équipe ACEP, GQE-Le Moulon).

Dans un premier temps, je présenterai une analyse statistique des valeurs des traits obtenues en ajustant mon modèle aux données de phénotypage des souches évoluées au laboratoire par A. Spor (Spor et al. 2014) et je mettrai en lumière les traits qui diffèrent selon le milieu d'évolution et le génotype de l'ancêtre. Je comparerai ces résultats à ceux d'Aymé Spor dont les estimations étaient réalisées à partir de modèle plus simples. Dans un second temps, j'étudierai les variations de la valeur sélective au cours d'un batch à partir des données de compétition, obtenues plus récemment par Carine Remoué au laboratoire. Il s'agit de données de compétitions entre souches fluorescentes et leurs isogéniques non fluorescentes ainsi que des compétitions entre les souches ancêtres de l'évolution expérimentale. Ces souches ont été choisies pour être contrastées et représenter le continuum de stratégies de vie caractérisées chez *S. cerevisiae* (de la stratégie de vie « petite cellule » aussi appelée « fourmis » à la stratégie « grande cellule » aussi appelée « cigale » (Spor et al. 2009)). Ces données n'ont jamais encore été analysées. Les métabolites n'y ont pas été quantifiés. L'analyse que j'ai réalisée n'a donc pas pu se faire sur la base de mon modèle dynamique. En revanche, j'ai pu étudier les variations de fitness le long du batch.

I. Caractérisation des souches et de la dynamique adaptative selon leur environnement d'évolution

La caractérisation des souches et l'analyse de leurs stratégies d'histoire de vie en fonction de leur milieu d'évolution ont déjà été menées (Spor et al. 2014). Cependant, grâce à l'ajustement du modèle aux données, d'autres traits ont pu être caractérisés. Par exemple, dans Spor et al (Spor et al. 2014), les traits liés à l'éthanol étaient résumés par la concentration en éthanol lorsque la souche est cultivée en monoculture pendant 48h. Ce paramètre a été identifié comme l'un des paramètres pour lequel il y avait de la convergence selon le milieu d'évolution. Toutefois, il était impossible de savoir si cette différence de concentration était liée à une différence de production de l'éthanol ou une différence de consommation. L'inférence, grâce au modèle, de paramètres comme la production et la tolérance à l'éthanol ou encore la mortalité permet de décrire plus finement les stratégies d'histoire de vie.

Dans cette partie, nous avons étudié l'effet du milieu d'évolution et de la contingence historique (effet du génotype de l'ancêtre) sur la valeur de chacun des traits caractérisés grâce au modèle. Dans un second temps, nous avons cherché à déterminer pour quels traits on observait de la convergence selon le milieu d'évolution.

A. Caractérisation des stratégies d'histoire de vie et convergence phénotypique

1. Matériel biologique et inférence des traits

Les données de phénotypage concernent 72 souches issues de l'évolution expérimentale menée par Aymé Spor et 6 ancêtres. Lors de l'évolution expérimentale, 6 souches ancêtres ont évolué pendant 6 mois dans 4 milieux différents (15%-48h, 1%-48h, 15%-96h, 1%-96h où le premier chiffre indique la concentration de glucose et le second le temps entre chaque repiquage), en trois réplicats indépendants. Le phénotypage a été réalisé sur un milieu batch de 96h, contenant initialement 15% de glucose, et ce en 5 blocs d'expérience où les 72 souches ont été réparties aléatoirement. Le modèle a été ajusté aux données de chaque culture batch comme décrit au chapitre 1.

2. Effet de la contingence, de l'environnement d'évolution sur la valeur des traits

Pour chaque trait X de chaque souche évoluée, nous avons posé le modèle suivant : $X_{i,j,k,l} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{i,j,k,l}$, où α_i est l'effet du bloc d'expérience, β_j est l'effet de l'ancêtre de la souche, et γ_k est l'effet de l'environnement dans lequel la souche a évolué. Les résidus $\varepsilon_{i,j,k,l}$ sont indépendants et identiquement distribués selon une loi normale centrée de variance σ^2_x . Pour chaque trait, chaque effet a été testé par une analyse de variance. Les résultats sont présentés dans le Tableau 5-1.

Trait	Effet blocs			Effet ancêtre			Effet environnement			résidus	
	ddl	SCM	P.val	ddl	SCM	P.val	Ddl	SCM	P.val	Ddl	SCR
Taux de croissance en fermentation r_f	4	0,314	0,00033	5	0,0570	0,493	3	0,0871	0,0895	59	0,754
Taux de mortalité m	4	0,0956	$2,45 \cdot 10^{-5}$	5	0,0419	0,021	3	0,0027	0,82	59	0,171
Tolérance à l'éthanol E_m	4	0,000417	0,00517	5	0,000209	0,17	3	0,000246	0,036	59	0,00152
Rendement éthanol/glucose p	4	0,16	$1,55 \cdot 10^{-8}$	5	0,0061	0,673	3	0,00515	$5,61 \cdot 10^{-5}$	59	0,113
Vitesse de consommation du glucose J_f	4	5,35 $\cdot 10^{-20}$	$7,63 \cdot 10^{-5}$	5	8,38 $\cdot 10^{-21}$	0,477	3	2,50 $\cdot 10^{-21}$	0,71	59	1,08 $\cdot 10^{-19}$
Taux de croissance en respiration r_r	4	0,012	0,021	5	0,0032	0,64	3	0,0086	0,036	59	0,056
Vitesse de consommation du glucose J_r	4	6,87 $\cdot 10^{-21}$	0,188	5	5,269 $\cdot 10^{-21}$	0,44	3	1,48 $\cdot 10^{-21}$	0,0059	59	6,36 $\cdot 10^{-20}$

Table 5-1: Anova sur les différents traits des souches évoluées. Les p-valeurs inférieures à 5% sont en gras.

On constate que l'effet bloc est significatif pour tous les traits, sauf la vitesse de consommation de l'éthanol. L'effet ancêtre est significatif pour le taux de mortalité. Enfin, le milieu d'évolution a un effet significatif sur la tolérance à l'éthanol, la production d'éthanol et la vitesse de consommation de l'éthanol.

Effet ancêtre

On peut noter, que la souche qui est « responsable » de l'effet ancêtre sur la mortalité est la S288C, qui est la seule haploïde et qui a le taux de mortalité le plus élevé parmi les souches ancêtres. Ses descendants ont une mortalité en moyenne plus forte que les autres.

Effet environnement

Trois traits, tous liés à l'éthanol, dépendaient significativement du milieu d'évolution : le rendement éthanol/glucose (**p**), la toxicité à l'éthanol (**E_m**) et la vitesse de consommation de l'éthanol (**J_r**).

3. Convergence phénotypique

Afin de déterminer par quels traits se différenciaient les souches ayant évolué dans des environnements différents, j'ai effectué une analyse factorielle discriminante (LDA) des valeurs de traits estimées par mon modèle (figure 5-1). Cette analyse permet de trouver le système d'axes qui discrimine au mieux les souches en fonction du milieu dans lequel elles ont évolué; chaque axe étant une combinaison linéaire des traits caractérisant les souches.

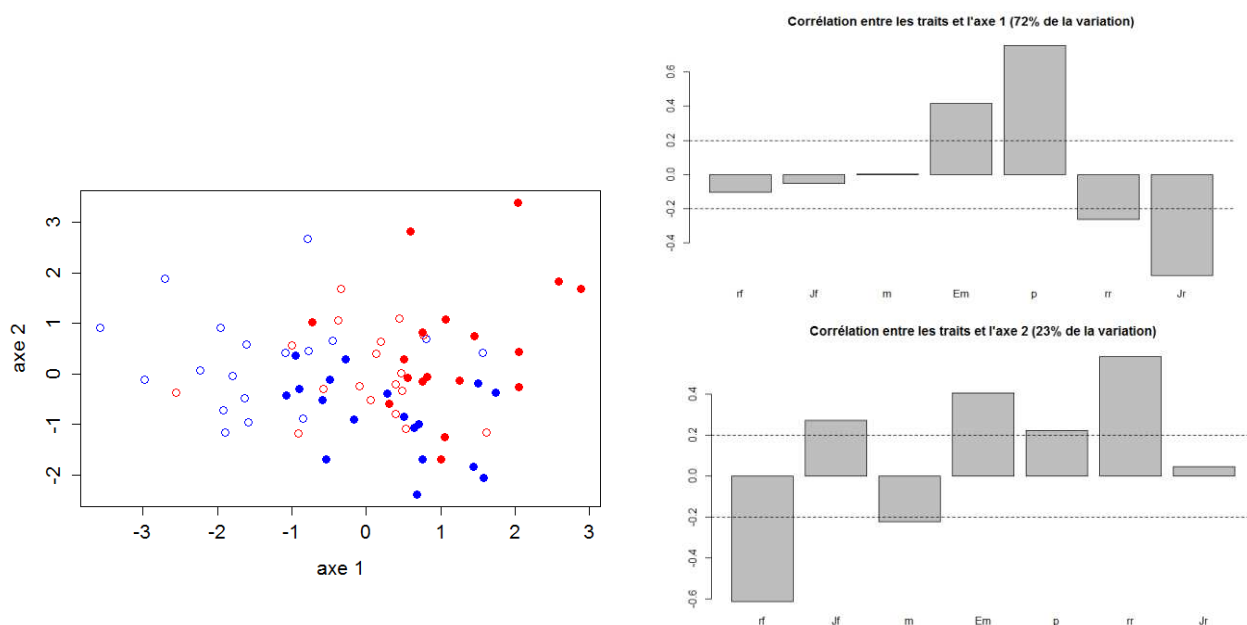


Figure 5-1: Analyse factorielle discriminante pour discriminer les souches évoluées dans différents milieu à partir des traits. Les valeurs des traits sont corrigées de l'effet bloc. La couleur représente l'environnement dans lequel la souche a évolué (rouge =15%, bleu=1%, plein=96h et vide =48h).

Le premier axe discrimine principalement les souches ayant évolué dans le milieu 1%-48h (point bleus vides) des souches ayant évolué dans les trois autres milieux. Cet axe est corrélé à la tolérance à l'éthanol (E_m), au rendement éthanol/glucose (p) et à la vitesse de consommation de l'éthanol (J_r). La tolérance à l'éthanol et la production d'éthanol sont plus faibles pour les souches ayant évolués dans le milieu 1%-48h mais elles consomment l'éthanol plus rapidement.

Le second axe discrimine les souches ayant évolué dans le milieu 15%-96h (points rouges pleins) de celles ayant évolué dans le milieu 1%-96h (points bleus pleins). Cet axe est corrélé à la tolérance à l'éthanol et aux taux de croissance en fermentation et respiration. La tolérance à l'éthanol et le taux de croissance en respiration sont plus élevés pour les souches ayant évolué dans un milieu riche en glucose, mais le taux de croissance en fermentation y est plus faible. Les groupes ne sont cependant pas très bien discriminés et il est donc possible que ces différences ne soient pas significatives.

II. Analyse de la valeur sélective des souches ancêtres et relation avec les traits d'histoire de vie

En mettant en compétition deux souches, il est possible de suivre l'évolution de leur fréquence et d'en déduire l'évolution de leur valeur sélective au cours d'un batch. De plus il est possible, en utilisant également les données de phénotypage (présentées au paragraphe I de ce chapitre), de confronter (au moins qualitativement) la valeur sélective à la décomposition de la fitness proposée au chapitre 2.

Le travail expérimental présenté ci-dessous a été effectué par Carine Remoué (équipe ACEP, GQE-Le Moulon).

A. Matériel biologique

Pour mettre deux souches en compétition et suivre la dynamique de population de chacune d'entre elles au cours d'un batch, il est nécessaire de pouvoir distinguer les deux souches. Pour cela, il est possible de transformer les souches en introduisant dans le génome un plasmide codant une protéine fluorescente détectable au cytomètre de flux.

Au laboratoire, Carine Remoué a mené i) la transformation de plusieurs des souches ancêtres utilisées dans l'évolution expérimentale menée au cours de la thèse d'Aymé Spor, ii) des expériences de compétition de chaque souche transformée avec la souche isogénique non transformée afin de vérifier qu'il n'y a pas d'effet du plasmide sur la fitness des souches, iii) la compétition entre les différentes souches ancêtres, l'une transformée l'autre non.

Transformation des souches

Quatre souches ancêtres (une cinquième n'a pas conservé le plasmide) diploïdes ont été transformées avec un plasmide codant une protéine fluorescente rouge. Ce sont les souches 97 (de laboratoire), 104 (de forêt), 270 (de fruit), et 304 (clinique). La construction contenant le gène de fluorescence a été introduite dans le génome au niveau du gène HO par

recombinaison homologue(avec une méthode comparable à celle utilisée dans (Raffoux et al. 2018)).

Expériences de compétition

Chaque souche transformée, capable de produire une protéine fluorescente rouge, a été mise en compétition 2 à 2 avec les quatre souches non-transformées (y compris elle-même) avec des fréquences initiales de 50%, une concentration totale de 10^6 cellules/mL, dans un milieu à 15% de glucose pendant 96 heures. Cela constitue un total de 16 compétitions. Trois blocs d'expérience incluant chacun les 16 compétitions possibles entre souches fluorescentes ou non ont été réalisés. Des prélèvements de chaque culture ont été faits après 0, 2, 4, 6, 8, 24, 48 et 96 heures de culture. A chaque point de prélèvement, le nombre de cellules a été mesuré au Coulter, et le pourcentage de cellules vivantes fluorescentes (noté « rouge » plus bas), vivantes non fluorescentes (noté « noire » plus bas) et de cellules mortes ont été comptées au cytomètre de flux. Dans la culture, une coloration au Sytox permet de différencier au cytomètre les cellules mortes des vivantes.

On peut noter comme dit précédemment que ces données ne permettent pas d'estimer les paramètres du modèle dynamique que j'ai développé car les concentrations de ressources n'ont pas été mesurées ce qui rend impossible l'ajustement correct du modèle aux données.

Quantifier la valeur sélective d'une souche par rapport à une autre

La fitness est mesurée par : $W_{\text{rouge/noire}}(t) = \ln(N_{\text{rouge}}(t)/N_{\text{noire}}(t)) - \ln(N_{\text{rouge}}(0)/N_{\text{noire}}(0))$ où $N_{\text{rouge}}(t)$ est la taille de la population de la souche transformée et $N_{\text{noire}}(t)$ la taille de la population de la souche non transformée.

Pour chaque temps de prélèvement t , on écrit le modèle suivant : $W_{i,j,k}(t) = \alpha_{i,j} + \epsilon_{i,j,k}$, où $W_{i,j,k}(t)$ est la fitness de la souche i au temps t lorsqu'elle est en compétition avec la souche j dans le réplicat k et $\alpha_{i,j}$ est la fitness moyenne de la souche i par rapport à la souche j . Les résidus $\epsilon_{i,j,k}$ sont indépendants et suivent une loi normale centrée de variance σ^2 . Les trois réplicats des 16 compétitions sont analysés ensemble. Les souches transformées sont donc considérées comme des souches différentes afin de pouvoir vérifier s'il n'y a pas un effet de la transformation. Cette analyse est préliminaire et un effet aléatoire du bloc devrait y être ajouté. La nullité de la fitness a été testée au seuil de 5% par un test de Wald.

Nous présentons ici les analyses pour les points de prélèvement à t=8h (dans la phase de fermentation) et à t=24h (dans la phase de mortalité). Ces deux points ont été choisis car ils représentent chacun une saison. Le point t=8h est le point le plus tard dans la phase de fermentation, on s'attend donc à voir plus facilement les différences entre souches. Aux points t=48h et 96h, certaines souches ont un nombre de cellules vivantes observées nulles. On ne peut donc pas calculer la fitness pour certains couples, ce qui limiterait la puissance du test. La fitness sera donc testée au point t=24h. L'analyse pour les autres points de prélèvement est en cours.

1. Analyse de la valeur sélective à 8 heures

La fitness des différents couples est représentée sur la figure 5-2. L'analyse nous montre que la fitness de chaque souche contre sa version transformée n'est pas significativement différente de 0 après 8 heures de culture, indiquant un coût faible ou négligeable de la transformation. De plus la matrice des compétitions deux à deux semble antisymétrique ce qui confirme un effet négligeable de la construction fluorescente. On constate en regardant cette matrice que la souche 270 est la seule à montrer une valeur sélective significativement différente. Elle perd à la compétition face aux trois autres ancêtres.

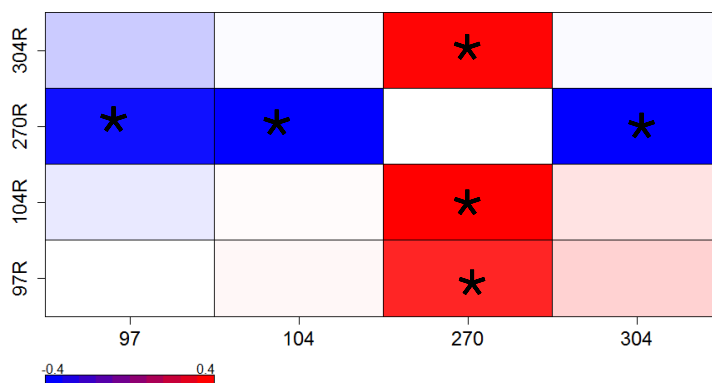


Figure 5-2: Fitness des souches transformées (représentées sur l'axe des ordonnées) dans les différentes compétitions. Les étoiles indiquent les fitness significativement différentes de zéro au seuil de 5%.

D'après les valeurs estimées dans les données de phénotypage, la souche 270 a un taux de croissance net (taux de croissance moins taux de mortalité) estimé de 0.10, la souche 104 de

0.14, la souche 97 de 0.19 et la souche 304 de 0.42. D'après notre décomposition de la fitness, cette dernière est proportionnelle à la différence des taux de croissance nets des deux souches en compétition ce qui est en accord avec nos observations expérimentales d'une valeur sélective négative de la souche 270. Cependant les autres écarts de taux de croissance ne se retrouvent pas dans la fitness. On s'attend en effet à ce que la souche 104 ait une fitness négative lorsqu'elle est en compétition avec la 97, ou que la 304 est une fitness très élevée quel que soit sa compétitrice, puisque son taux de croissance net est très supérieur à ceux des autres souches. Cette différence peut s'expliquer par une erreur dans l'estimation des paramètres, ou une différence expérimentale liée aux conditions expérimentales. En effet, les taux de croissance des souches ont été estimés à partir des données de compétition par Anthony Petit lors de son stage de L2 en 2016, que j'ai co-encadré avec Judith Legrand. Ces taux de croissance sont présentés sur la figure 5-3. Les estimations ont été obtenues avec régression linéaire de la forme $\log(N_{i,k}(t)) = \alpha_i + \beta_i \cdot t + \varepsilon_{i,k}$, avec $N_{i,k}(t)$ le nombre de cellules de la souche i au temps t . Au moment où l'analyse a été seulement deux réplicats étaient disponibles. Cette estimation donne un taux de croissance de la souche 304 plus proche des trois autres. La différence entre les souches 104 et 97 n'est plus significative non plus. Avec cette méthodologie, la souche 270 a également un taux de croissance significativement plus faible que les souches 104 et 97, ce qui correspond aux mesures de fitness. D'autres expériences devront être menées pour comprendre les différences entre l'expérience de compétition et les données de phénotypage présentées dans le chapitre A.

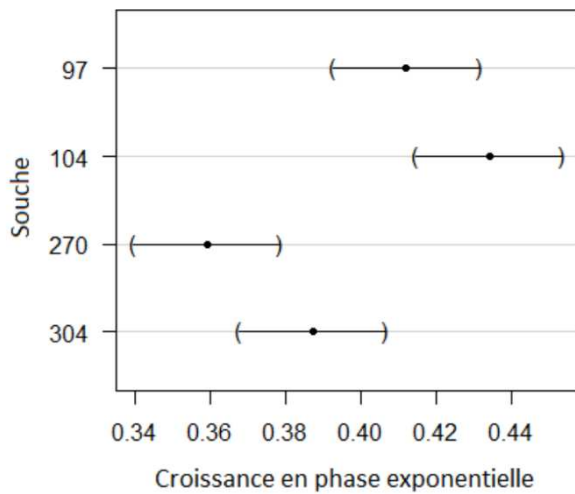


Figure 5-3: Estimation des taux de croissance net en phase exponentielle. Estimation ponctuelle et intervalles de confiance à 95%.

2. Analyse de la valeur sélective après 24h de culture

Les fitness à 24h sont présentées sur la figure 5-4 (panel de gauche). A 24h, comme à 8h, l'effet de la construction avec le gène de fluorescence apparaît négligeable (souches en compétition contre leur version transformée ont une fitness non significativement différente de 0 et la matrice apparaît symétrique).

La valeur sélective est significativement différente de 0 pour les deux compétitions impliquant la souche 104 et la souche 304, La valeur sélective de la souche 270 par rapport aux autres semblent négative mais ce n'est pas significatif. D'après note décomposition, $W(24) = T_f \cdot \Delta r_f + r_f \Delta T_f - 24 \cdot \Delta m$. La différence de fitness serait donc liée soit à une différence de taux de mortalité, soit à une différence de tolérance à l'éthanol, soit à une différence de taux de croissance. Nous avons vu dans le paragraphe précédent l'effet de la différence de taux de croissance sur la fitness. Afin de diminuer l'importance de ce terme dans la fitness et de n'observer que ceux liés à la mortalité et à la toxicité de l'éthanol, nous avons refait la même analyse mais en considérant l'écart de fitness entre 8 et 24h. Les résultats sont présentés sur la figure 5-4 (à droite). Les mêmes compétitions sont significatives. Cependant la fitness de la souche 270 ne semble pas varier entre 8 et 24h. Le fait que la souche 270 ait une fitness

négative (mais non significative) à 24h semble donc lié aux taux de croissance. La fitness de la souche 104 semble diminuer entre 8 et 24h dans toutes les compétitions. Le fait qu'elle ait une fitness négative à 24h semble donc lié à la mortalité et à la tolérance à l'éthanol.

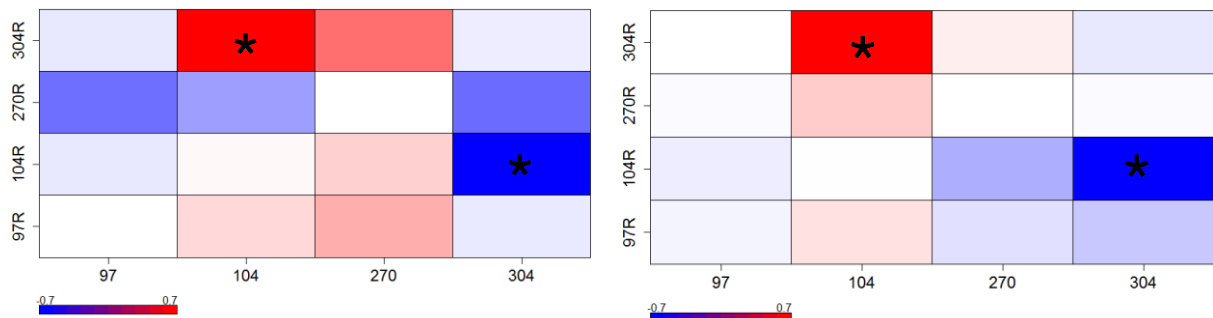


Figure 5-4: Fitness des souches transformées dans les différentes compétitions. A gauche, la fitness à 24h ; à droite l'écart de fitness entre 24 heures et 8 heures. Les étoiles indiquent les fitness significativement différentes de zéro.

D'après les valeurs estimées dans les données de phénotypage, la souche 104 a un taux de mortalité estimé de 0.19, la souche 270 de 0.14, la souche 97 de 0.08 et la souche 304 de 0.02. La souche 104 a donc le taux de mortalité le plus élevée et est donc désavantagée lorsque la survie est sélectionnée.

De la même manière, on s'attendrait à ce que la souche 304 ait une meilleure valeur sélective. Cependant la tolérance à l'éthanol de la souche 304 est 10 fois plus faible que celles des trois autres. On aurait donc un effet inverse à celui de la 270 en phase exponentielle : la souche 304 arrêterait de se diviser plus tôt. Elle serait donc désavantagée le temps que les autres souches arrêtent aussi leur croissance. Les termes $r_f \Delta T_f$ et $T_{\Delta m}$ se compensent pour la souche 304 en compétition avec les souches 97 et 270, sa fitness à 24h est donc nulle. Pour tester cela, on pourrait regarder la fitness de la souche 304 à 48h, qui devrait être alors positive (le terme $T_{\Delta m}$ augmente, alors que $r_f \Delta T_f$ ne varie pas avec la longueur du batch).

Remarque sur la dynamique au-delà de 24 heures

Entre 24 et 96h, on observe une mortalité des cellules et une diminution de la taille de la population. A partir de 48h certaines souches ont un nombre de cellules nulles. Dans ce cas il

n'est plus possible de mesurer la fitness. De plus la fitness est beaucoup plus variable à partir de 48 heures. La table 5-2 illustre ces effets qui pourraient être liés à la variance du nombre de cellules lors du prélèvement. Nous n'allons donc pas étudier la fitness aux points 48 et 96h, d'autres études devront être menées pour comprendre l'origine de cette variabilité et ses conséquences.

Fitness souche 304R contre 104	Répétition 1	Répétition 2	Répétition 3
À 24h	0,39	0,67	0,52
À 48h	-0,12	0,69	0,59
À 96h	-0,10	(souche 104 à 0 cell/mL)	(souche 104 à 0 cell/mL)

Table 5-2 : Fitness de la souche 304 transformée mise en compétition avec la souche 104 à différents temps dans les trois répétitions

III. Conclusion du chapitre

Dans ce chapitre nous avons complété la caractérisation des souches évoluées à partir de données de phénotypage en monoculture faite dans (Spor 2009; Spor et al. 2014) grâce à l'exploitation du modèle dynamique donnant accès à des paramètres moins intégrés. Dans cette précédente analyse, la concentration d'éthanol (un caractère fortement intégré) était l'un des traits qui discriminait le plus les souches ayant évolué dans différents environnements. Grâce aux traits moins intégrés estimés par ajustant du modèle, je montre que les différences observées seraient liée à la fois à une différence de rendement éthanol/glucose, de tolérance à l'éthanol et à la vitesse de consommation de celui-ci. Les souches ayant été sélectionnées sur un environnement pauvre en glucose et renouvelé fréquemment consomment l'éthanol plus rapidement, en produisent moins et sont moins résistantes à celui-ci. De manière similaire les environnements riches en glucose ont sélectionné la tolérance à l'éthanol plus fortement que le milieu avec une concentration initiale en glucose plus faible.

Par ailleurs, nous avons estimé la fitness relative de souches ayant des traits d'histoire de vie différents en analysant les compétitions entre les ancêtres de l'expérience d'évolution au laboratoire (Spor et al. 2014). Nous avons tout d'abord montré que le coût du marquage était nul ou négligeable, permettant d'utiliser ces souches ancêtres marquées en compétition avec des évoluées issues de l'évolution expérimentale. Ce travail qui n'a pas pu être réalisé durant mon travail de thèse sera poursuivi par la suite au sein de mon équipe de recherche.

Nous avons aussi détecté de différences de fitness cohérentes avec les attendues sachant les valeurs de paramètres estimés à partir des cultures isolées, validant le modèle comme un outil d'estimation des paramètres individuels des souches pertinent. A partir de cela, nous avons pu classer les souches. La souche issue de la forêt (104) a un taux de mortalité plus faible que la souche 97 ; la souche issue de fruit (270) a un taux de croissance plus faible que la souche 97 ; et enfin la souche clinique (304) a un taux de mortalité plus faible, mais une plus faible tolérance à l'éthanol que la souche 97. Ces quatre souches issues de quatre milieux différents présentent donc bien quatre stratégies d'histoire de vie différentes.

Ces données ont également mis en évidence une forte variabilité de la fitness après 48h, probablement liée au faible nombre de cellules vivantes que l'on détecte. La dynamique serait alors fortement influencée par la dérive, ce qui n'est pas pris en compte dans notre modèle. On rediscutera de ce point précis plus en détail dans la discussion.

Bibliographie :

- Raffoux, Xavier et al. 2018. "High-Throughput Measurement of Recombination Rates and Genetic Interference in *Saccharomyces Cerevisiae*." *Yeast*. <http://doi.wiley.com/10.1002/yea.3315> (May 5, 2018).
- Spor, Aymé. 2009. "Adaptive Strategies to Contrasted Environments in *Saccharomyces Cerevisiae*." Université Paris-Sud XI.
- . 2009. "Niche-Driven Evolution of Metabolic and Life-History Strategies in Natural and Domesticated Populations of *Saccharomyces Cerevisiae*." *BMC evolutionary biology* 9: 296.
- . 2014. "Phenotypic and Genotypic Convergences Are Influenced by Historical Contingency and Environment in Yeast." *Evolution* 68(3): 772–90.

DISCUSSION

Discussion

I. Synthèse des résultats

Le but de cette thèse était d'étudier l'évolution des traits d'histoire de vie dans un environnement saisonnier et de comprendre l'effet des traits eux-mêmes et de l'environnement sur les pressions de sélection. Pour cela, par une approche de modélisation de la dynamique évolutive de souches de levures dans une évolution expérimentale en culture batch, j'ai notamment proposé une décomposition de la valeur sélective en fonction des saisons et mis en évidence quels traits sont sélectionnés selon les saisons et comment l'évolution de traits influencent les composantes de la valeur sélective. Cette décomposition permet de prendre en compte l'effet de l'environnement, via les longueurs de saison sur la valeur sélective d'une souche par rapport à une autre. Dans cette première partie, j'ai ainsi mis en évidence les boucles éco-évolutives présentes dans ce type de dispositif expérimental, c'est-à-dire le fait que la dynamique des fréquences des souches (liée à leur valeur sélective) va influencer la dynamique des saisons et donc les traits soumis à sélection qui vont à leur tour influencer la dynamique des fréquences des souches. Dans un second temps, j'ai montré que, du fait de ces boucles éco-évolutives, l'existence d'une contrainte entre deux traits pouvait entraîner au cours de la dynamique évolutive l'émergence de relations entre d'autres traits historiquement indépendants.

Enfin, pour calibrer le modèle et mettre en regard des données biologiques et les prédictions du modèle, j'ai analysé des données de phénotypage de souches de levures ayant évolué dans un dispositif expérimental de cultures successives en batch, identique à celui modélisé. À travers un ajustement du modèle, j'ai pu déterminer pour toutes les souches les paramètres mathématiques du modèle et en étudier la diversité. Deux types de phénotypage ont été analysés : des monocultures de souches évoluées et ancestrales et des cultures en compétition entre des souches ancêtres.

Le phénotypage des souches seules nous a permis de mettre en évidence les stratégies sélectionnées dans chaque environnement de l'évolution expérimentale. Ces stratégies dépendent principalement des traits liés à la respiration et à la dynamique de l'éthanol, précisant les résultats du doctorat d'Aymé Spor (Spor 2009; Spor et al. 2014). L'analyse des

souches en compétition a permis de mettre en évidence quels traits influencent le plus la valeur sélective et les différences entre les souches étudiées (les ancêtres de l'évolution expérimentale). Ces différences sont liées aux taux de croissance, de mortalité, mais aussi à la tolérance à l'éthanol ou au temps de latence. Ces différences entre souches et leurs impacts observés sur la valeur sélective sont cohérents avec la décomposition de la fitness proposée à partir de notre modèle.

Dans cette partie, je discuterai en premier lieu de mes résultats concernant les composantes de la valeur sélective. Dans un second temps, je discuterai de l'impact des boucles éco-évolutives sur i) l'émergence potentielle de relations entre traits, ii) la coexistence. Enfin, je discuterai des limites de mon travail.

II. Composantes de la valeur sélective et longueur des saisons

A partir de la décomposition de la valeur sélective proposée dans le chapitre 2, j'ai identifié deux types de trait sous sélection : les traits d'histoire de vie reliés à la croissance ou la mortalité et des traits de transition.

Les traits reliés à la croissance et la mortalité sont classiquement étudiés et utilisés comme proxy de la valeur sélective (Chevin 2011; Wiser and Lenski 2015). On les retrouve également dans les données de compétition que j'ai étudiées (chapitre 4) : la souche 270 a une valeur sélective négative en raison de son taux de croissance plus faible, et la souche 104 en raison de son taux de mortalité plus fort que les autres souches ancêtres.

La décomposition de la valeur sélective met également en évidence l'effet de traits de transition. Dans le cas du modèle proposé ici, on retrouve plusieurs traits étudiés dans la littérature comme l'efficacité à utiliser les ressources lorsque celle-ci sont rares (K_f et K_r dans notre modèle) (Kremer and Klausmeier 2013; Vasi, Travisano, and Lenski 1994), la tolérance à une toxine (E_m dans notre modèle ou celle aux métaux lourds par exemple dans (Gorter et al. 2018)), ou la transition entre les deux phases (K_c) (Spencer, Bertrand, et al. 2007). Cet effet des traits de transition sur la valeur sélective peut expliquer la valeur sélective négative de la souche 304 malgré son taux de mortalité plus faible que celui des autres souches.

D'autres traits de transition, non intégrés au modèle proposé ici, ont été étudiés et identifiés dans d'autres études comme des composantes de la valeur sélective, par exemple le temps de latence au départ de la culture dans la Long Term Experimental Evolution (LTEE) (Vasi, Travisano, and Lenski 1994).

L'intensité de la sélection sur les différents traits dépend de la longueur de la saison associée permettant d'expliquer la fréquence dépendance de la valeur sélective, mise en évidence dans certains travaux (Ribeck and Lenski 2015; Yi and Dean 2013). De plus, j'ai montré au chapitre 3 que la valeur optimale d'un trait sous sélection, lorsqu'elle existe, dépend notamment des longueurs de saison. Ces stratégies optimales dépendent également des autres traits entraînant des boucles éco-évolutives discutées dans le paragraphe suivant. De plus nous avons montré qu'une stratégie optimale existe, s'il existe un trade-off convexe entre deux traits. Ce résultat avait été montré dans le cas d'un environnement fixe (Doebeli 2002; Levins 1962), par simulation dans le cas d'un environnement batch (Spencer, Saxer, et al. 2007). Dans le cas général, il a été montré que la forme du trade-off influence la forme du point d'équilibre (Débarre, Nuismer, and Doebeli 2014; Mazancourt and Dieckmann 2004), et des méthodes ont été proposées pour trouver un trade-off conduisant à n'importe quel équilibre (Ito and Sasaki 2016; Kisdi 2015).

III. Boucles éco-évolutives et impact sur les relations entre traits

Comme dit précédemment, la valeur optimale des traits dépend de la longueur des différentes saisons, qui dépendent elles-mêmes des traits, via les boucles éco-évolutives. Il y a donc des relations entre les valeurs optimales des traits. La figure 6-1, commentée en Supplementary Information de l'article du chapitre 2 – voir annexe) résume les relations entre les traits et la sélection sur ces traits. Les longueurs de saisons influencent aussi le poids de chaque trait sur la valeur sélective. Ces relations ont donc une influence à la fois sur les valeurs optimales et l'intensité de la sélection sur les différents traits. On retrouve quatre traits sous sélection (cases blanches) et trois traits non sélectionnés (cases bleues). Les trois traits restants K_f , K_c et K_r ne sont pas présents sur le schéma. Ces traits sont sous sélection, car ils influencent les longueurs efficaces de saison. Ils sont même parmi les principaux traits qui distinguent les souches S et L de *E. coli* (Spencer, Bertrand, et al. 2007) ou certaines souches de

phytoplanctons (Edwards, Klausmeier, and Litchman 2013). Cependant dans notre étude, ces traits ont été fixés à des valeurs faibles, suite à une recherche dans la littérature. On observe donc peu d'effet dans les différentes analyses de sensibilité que nous avons réalisées (chapitre 2) par rapport à celui des taux de croissance.

On peut identifier trois types de relations :

- Des boucles ou « feedbacks » : un trait influence sa propre valeur optimale (en orange sur le schéma). Ces feedbacks permettent de modifier les types d'équilibres évolutifs possibles. Si il n'y a pas de feedback, le branchement est par exemple impossible, et la stabilité n'est déterminée que par la forme du trade-off (cas analogue à celui de (Levins 1962) en environnement fixe).
- Des relations entre deux traits d'histoire de vie (relation entre deux cases blanches sur le schéma). Comme chaque trait est sélectionné, ils ont chacun un certain nombre de valeurs optimales. Il y aura donc un nombre fini de points d'équilibres, si les deux traits sont indépendants.
- Des relations entre un trait sélectionné et un trait non sélectionné. Dans ce cas, on prédit un continuum de stratégies optimales si les deux traits sont indépendants.

On peut envisager de tester cette prédiction expérimentalement : deux populations qui évoluent indépendamment atteindraient deux équilibres différents le long de ce continuum. On pourrait alors observer une relation entre les deux traits en observant le phénotype de plusieurs populations ayant évoluées indépendamment dans des environnements similaires. On s'attend notamment à observer cela dans les évolutions expérimentales où plusieurs répliques sont faits. Cela n'a cependant pas été observé dans des évolutions expérimentales en environnement saisonnier (Jasmin, Dillon, and Zeyl 2012; Novak et al. 2006). Dans la littérature, les relations entre traits non-sélectionnés sont peu documentées. Dans la plupart des études d'évolution expérimentale, seuls les taux de croissance sont mesurés, car utilisés comme proxy de la valeur sélective. On mesure cependant régulièrement le rendement (Orr 2010; Wiser and Lenski 2015) ou des traits sélectionnés par le stress imposé (Dhar et al. 2011; Gorter et al. 2016). Afin de valider l'existence de ce continuum d'équilibres, il faudrait étudier durant une évolution expérimentale deux traits réellement indépendants.

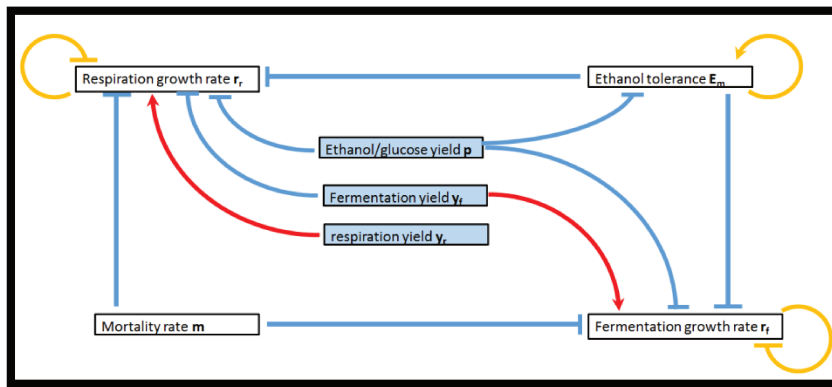


Figure 6-1 : Résumé des différentes rétroactions dans le modèle. En orange, les « feedback » stricts : un trait influence sa valeur optimale et son poids dans la fitness ; en bleu (resp. rouge), les relations négatives (resp. positives) : un traits diminue (resp. augmente) la valeur optimale et le poids d'un second trait lorsqu'il change sous l'effet de la sélection. Les fonds bleus indiquent les traits non-sélectionnés et les fonds blancs les traits sous sélection.

Il est cependant rare que deux traits soit indépendants. Parmi les traits utilisés dans le modèle, aucun trait non-sélectionné n'est indépendant des autres (Table 1-2). Ce résultat est donc difficile à tester directement.

Un autre moyen de tester cela serait de modifier artificiellement la valeur des traits. On pourrait par exemple augmenter ou diminuer la consommation des ressources, en modifiant leur concentration par exemple. On s'attend alors à observer des points d'équilibre différents, voire des points d'équilibres de type différents (stable, branchement ou instable). Une évolution *in silico* semble aller dans le sens de ce résultat (Rocabert et al. 2017). Elle a en effet montré que le branchement est plus probable en batch qu'en chemostat, où les ressources sont consommées plus vite à cause du renouvellement du milieu. De nombreuses autres différences existent entre les deux dispositifs, on ne peut donc pas conclure de façon certaine, il semble cependant que modifier la vitesse de consommation des ressources a modifié le type d'équilibre évolutif.

La forme du trade-off peut aussi impacter le nombre de souches qui peuvent coexister. Par exemple dans le cas particulier d'un trade-off linéaire, une infinité de souches peuvent coexister. Dans ce cas, l'équilibre est simplement défini par la pente du trade-off ($=T_f/T_r$). L'équilibre est donc atteint dès que l'environnement est optimal. A partir de là, toutes les souches ont une valeur sélective nulle et coexistent donc, quel que soit leur nombre ou leur

phénotype. Nous n'avons cependant pas déterminé le nombre maximal de souches pouvant coexister. Cette question est cependant importante, notamment lors des évolutions expérimentales. En effet, si plusieurs souches coexistent, il devient nécessaire de prélever plus de cellules pour connaître tous les phénotypes qui composent la population. Cette coexistence peut également conduire à une dynamique cyclique des souches qui coexistent, comme dans le cas d'un jeu pierre-feuille-ciseaux (chapitre 2 Figure 2-2 et (Paquin and Adams 1983)). Dans ce cas, connaître les phénotypes présents dans la population et leurs fréquences requiert différents points de mesure. Il serait intéressant d'étudier ces questions à l'aide du modèle et des longueurs de saisons, afin de pouvoir les prédire.

IV. Limites de notre approche

i. Interprétation des longueurs efficaces de saison

Au chapitre 2, la longueur efficace de saison est définie comme une intégrale qui dépend des ressources. Ces longueurs efficaces de saison ne sont donc pas mesurables directement. Il pourrait être utile de trouver un proxy de ces longueurs mesurables expérimentalement. En mesurant ces longueurs de saison, il serait possible de prédire quels types de souches peuvent envahir un environnement donné, éventuellement un environnement imposé par un résident ayant un intérêt (industriel, ...). Dans plusieurs travaux expérimentaux (notamment (Ribeck and Lenski 2015; Spencer, Saxer, et al. 2007; Yi and Dean 2013)) la croissance de la population est proche d'une croissance exponentielle et le modèle simplifié dans lequel les longueurs de saison correspondent à la dynamique des ressources peut être utilisé. Il est alors possible d'avoir une bonne approximation des longueurs de saisons et ainsi pouvoir prédire qualitativement l'issue de la compétition entre deux souches. Il est aussi envisageable de transformer la dynamique temporelle en une dynamique spatiale avec l'aide d'un fermenteur multi-étagés (Baake, Cordero, and Hummel 2018). Dans ce dispositif expérimental, quatre chemostats sont connectés les uns aux autres et chacun nourrit le suivant. On pourrait donc avoir un fermenteur reproduisant l'état de fermentation et un autre celui de respiration. En étudiant la fréquence des souches à l'état stationnaire du système on pourrait mieux comprendre le rôle exact de chacun des paramètres, notamment ceux liés à la respiration.

ii. Effet de la dérive

La dynamique évolutive d'une population de levure au cours d'une évolution en cultures batch successives dépend notamment du hasard. L'effet du hasard intervient à trois niveaux : lors de l'apparition de nouvelles souches et dans la dynamique de population en cours de batch et au moment du repiquage.

1. Effet de la dérive en cours de batch

Le modèle de dynamique de population au cours d'un batch que j'ai développé et exploité au cours de ma thèse est un modèle déterministe, c'est-à-dire que l'on ne prend pas en compte l'effet du hasard sur la dynamique des populations. Dans ce cas, la survie des souches est surestimée car elles ne peuvent pas s'éteindre par dérive. Cette hypothèse est valable pour de grandes tailles de populations, mais ne permet pas de modéliser correctement la stochasticité liée à la très faible fréquence initiale de souches mutantes. Un certain nombre de modèles stochastiques existent aujourd'hui pour décrire la dynamique de populations bactériennes, utilisant des équations différentielles stochastiques (par exemple, (Biliouris et al. 2012)), des équations aux dérivées partielles (Jourdain, Méléard, and Woyczynski 2012), des graphes ancestraux (Baake, Cordero, and Hummel 2018) ou des processus de branchement (Casanova et al. 2016) et pourraient constituer des pistes à adapter à notre approche. Concernant une culture en batch, j'ai calculé la variance du nombre de cellules mutantes au moment du repiquage dont dépend la probabilité d'extinction. Des résultats préliminaires montrent que cette variance ne dépend que de la taille de population du mutant et non de sa fitness (résultats non présentés ici). Ainsi, quelle que soit la valeur sélective d'une souche, il faut un taux de croissance suffisamment fort pour ne pas s'éteindre. Ceci pourrait conduire à un conflit entre les souches qui ont une faible probabilité d'extinction (donc une dynamique avec une forte croissance) et celle qui ont une fitness élevée. A l'inverse, des souches ayant une fitness positive pourraient ne pas émerger du fait d'une forte probabilité d'extinction.

2. Effet de la dérive au moment du repiquage

Pour le modèle utilisé au chapitre 3, le repiquage a été simulé en tirant la fréquence des souches dans une loi multinomiale dont les paramètres sont les fréquences à la fin du batch

précédent. Ce modèle prend donc en compte l'effet d'échantillonnage au moment du repiquage (on parle de « bottleneck ») étudié dans (Handel and Bennett 2008; Wahl and Zhu 2015). On peut donc calculer la probabilité d'extinction d'une souche mutante dû à cet effet de « bottleneck » pour une souche présente en début de culture ou apparue au cours de la culture (résultats non présentés ici).

Pour les souches présentes au début de culture, cette probabilité dépend de la fitness du mutant et de sa fréquence initiale. Plus une cellule a une fitness élevée, moins sa probabilité d'extinction au moment du repiquage est élevée. De même, plus sa fréquence est élevée, plus sa probabilité d'extinction est faible.

Pour une souche apparaissant au cours de la culture, la probabilité d'extinction dépend aussi du moment de son apparition au cours du batch (Wahl and Zhu 2015). Notons que ces calculs prennent uniquement en compte l'effet de « bottleneck » et non les effets aléatoires de la dynamique de population au cours du batch commentés plus haut.

iii. Choix de modélisation et adéquation aux données

Bien que l'adéquation du modèle choisi aux données soit acceptable, certains points pourraient être améliorés. Par exemple la forme de la fonction choisie pour la transition entre la fermentation et la respiration ne permet pas d'ajuster précisément la dynamique de la concentration d'éthanol et de glucose. De même la mortalité semble être surestimée. D'autres formes de fonction auraient pu être utilisées pour mieux ajuster ces données. Cependant, les données disponibles n'auraient pas permis d'estimer correctement un nombre plus important de paramètres et de comparer différents modèles. De plus, nous n'avons pas d'explication satisfaisante pour expliquer la forte mortalité que l'on observe. Elle pourrait être liée à une carence en azote, un effet d'un autre produit de la fermentation (comme l'acétate), à la variation du pH, ou à un effet osmotique lié à la forte concentration en glucose. Les données disponibles ne permettent pas de privilégier une de ces hypothèses. D'autres données seraient donc nécessaires pour affiner le modèle et mieux expliquer la dynamique de population.

Ces modifications ne devraient pas affecter le comportement qualitatif du modèle. En effet, elles pourraient modifier les longueurs de saisons, mais le point d'équilibre pourrait toujours être défini à partir de celles-ci. Si d'autres paramètres étaient pris en compte, il pourrait être ajouté au schéma 6-1 représentant les relations entre paramètres.

Une modélisation plus fine de la dynamique éco-évolutive pourrait expliquer des différences plus marquées entre les stratégies apparues dans les différents environnements d'évolution. L'ajustement du modèle sur les données de phénotypage a mis en évidence des différences liées à la respiration. Ces données doivent être consolidées car ces paramètres étaient difficiles à estimer. Cependant, d'autres auteurs ont montré que l'adaptation n'est pas liée principalement à la fermentation (Gerstein and Otto 2011). Il serait donc intéressant d'investiguer les mécanismes adaptatifs sur les phases de respiration et de mortalité ou sur les traits de transition et de pouvoir raffiner le modèle sur ce point.

La compétition des ancêtres et de leurs descendants ayant évolué dans les différents milieux pourraient apporter des réponses. Mais, les données de compétition présentées au chapitre 4 montrent la difficulté d'étudier cette phase pour laquelle la fitness est très variable d'un réplicat à un autre après 24 heures. D'autres développements expérimentaux seront nécessaires pour évaluer si cette variabilité est due à un problème expérimental de mesure ou si la dérive a une part importante dans la dynamique de population.

Bibliographie

Baake, Ellen, Fernando Cordero, and Sebastian Hummel. 2018. "A Probabilistic View on the Deterministic Mutation–selection Equation: Dynamics, Equilibria, and Ancestry via Individual Lines of Descent." *Journal of Mathematical Biology*: 1–26.
<http://link.springer.com/10.1007/s00285-018-1228-8> (May 5, 2018).

Biliouris, Konstantinos, David Babson, Claudia Schmidt-Dannert, and Yiannis N Kaznessis. 2012. "Stochastic Simulations of a Synthetic Bacteria-Yeast Ecosystem." *BMC Systems Biology* 6(1): 58. <http://bmcsystbiol.biomedcentral.com/articles/10.1186/1752-0509-6-58> (May 5, 2018).

Casanova, Gonzalez A., Noemi Kurt, Anton Wakolbinger, and Linglong Yuan. 2016. "An

- Individual-Based Model for the Lenski Experiment, and the Deceleration of the Relative Fitness." *Stochastic Processes and their Applications* 126(8): 2211–52.
- Chevin, Luis-Miguel. 2011. "On Measuring Selection in Experimental Evolution." *Biology letters* 7: 210–13.
- Débarre, F., S. L. Nuismer, and M. Doebeli. 2014. "Multidimensional (Co)Evolutionary Stability." *The American Naturalist* 184(2): 158–71.
<http://www.journals.uchicago.edu/doi/10.1086/677137>.
- Dhar, R. et al. 2011. "Adaptation of *Saccharomyces Cerevisiae* to Saline Stress through Laboratory Evolution." *Journal of Evolutionary Biology* 24(5): 1135–53.
- Doebeli, M. 2002. "A Model for the Evolutionary Dynamics of Cross-Feeding Polymorphisms in Microorganisms." *Population Ecology* 44(2): 59–70.
- Edwards, Kyle F., Christopher A. Klausmeier, and Elena Litchman. 2013. "A Three-Way Trade-Off Maintains Functional Diversity under Variable Resource Supply." *The American Naturalist* 182(6): 786–800. <http://www.journals.uchicago.edu/doi/10.1086/673532>.
- Gerstein, Aleeza C, and Sarah P Otto. 2011. "Cryptic Fitness Advantage : Diploids Invade Haploid Populations Despite Lacking Any Apparent Advantage as Measured by Standard Fitness Assays." *PloS one* 6(12).
- Gorter, Florian A., Mark M. G. Aarts, Bas J. Zwaan, and J. Arjan G. M. de Visser. 2016. "Dynamics of Adaptation in Experimental Yeast Populations Exposed to Gradual and Abrupt Change in Heavy Metal Concentration." *The American Naturalist* 187(1): 110–19.
- Gorter, Florian A, Mark G M Aarts, Bas J Zwaan, and J Arjan G M de Visser. 2018. "Local Fitness Landscapes Predict Yeast Evolutionary Dynamics in Directionally Changing Environments." *Genetics* 208(1): 307–22.
<http://www.ncbi.nlm.nih.gov/pubmed/29141909> (May 4, 2018).
- Handel, Andreas, and Matthew R. Bennett. 2008. "Surviving the Bottleneck: Transmission Mutants and the Evolution of Microbial Populations." *Genetics* 180(4): 2193–2200.
- Ito, Hiroshi, and Akira Sasaki. 2016. "Evolutionary Branching under Multi-Dimensional

- Evolutionary Constraints.” *Journal of Theoretical Biology* 407: 409–28.
<http://dx.doi.org/10.1016/j.jtbi.2016.07.011>.
- Jasmin, Jean-Nicolas, Marcus Dillon, and Clifford Zeyl. 2012. “The Yield of Experimental Yeast Populations Declines during Selection.” *Proceedings. Biological sciences / The Royal Society* 279: 4382–88.
- Jourdain, Benjamin, Sylvie Méléard, and Wojbor A. Woyczynski. 2012. “Lévy Flights in Evolutionary Ecology.” *Journal of Mathematical Biology* 65(4): 677–707.
<http://link.springer.com/10.1007/s00285-011-0478-5> (May 5, 2018).
- Kisdi, Éva. 2015. “Construction of Multiple Trade-Offs to Obtain Arbitrary Singularities of Adaptive Dynamics.” *Journal of Mathematical Biology* 70(5): 1093–1117.
- Kremer, Colin T., and Christopher A. Klausmeier. 2013. “Coexistence in a Variable Environment: Eco-Evolutionary Perspectives.” *Journal of Theoretical Biology* 339: 14–25.
- Levins, Richard. 1962. “Theory of Fitness in a Heterogeneous Environment. I. The Fitness Set and Adaptive Function.” *The American Naturalist* 96(891): 361.
- Mazancourt, Claire De, and Ulf Dieckmann. 2004. “Trade-Off Geometries and Frequency-Dependent Selection.” 164(6).
- Novak, Maja et al. 2006. “Experimental Tests for an Evolutionary Trade-Off between Growth Rate and Yield in *E. Coli*.” *The American Naturalist* 168(2): 242–51.
- Orr, H Allen. 2010. “Fitness and Its Role in Evolutionary Genetics.” *Nat. Rev. Genet.* 10(8): 531–39.
- Paquin, Charlotte E, and Julian Adams. 1983. “Relative Fitness Can Decrease in Evolving Asexual Populations of *S. Cerevisiae*.” *Nature* 306: 368–71.
- Ribeck, Noah, and Richard E. Lenski. 2015. “Modeling and Quantifying Frequency-Dependent Fitness in Microbial Populations with Cross-Feeding Interactions.” *Evolution* 69(5): 1313–20.
- Rocabert, Charles et al. 2017. 13 PLoS Computational Biology *Beware Batch Culture: Seasonality and Niche Construction Predicted to Favor Bacterial Adaptive*

Diversification.

- Spencer, Christine C., Melanie Bertrand, Michael Travisano, and Michael Doebeli. 2007. "Adaptive Diversification in Genes That Regulate Resource Use in Escherichia Coli." *PLoS Genetics* 3(1): 0083–0088.
- Spencer, Christine C., Gerda Saxer, Michael Travisano, and Michael Doebeli. 2007. "Seasonal Resource Oscillations Maintain Diversity in Bacterial Microcosms." *Evolutionary Ecology Research* 9(5): 775–87.
- Spor, Aymé. 2009. "Adaptive Strategies to Contrasted Environments in Saccharomyces Cerevisiae." Université Paris-Sud XI.
- . 2014. "Phenotypic and Genotypic Convergences Are Influenced by Historical Contingency and Environment in Yeast." *Evolution* 68(3): 772–90.
- Vasi, F, M Travisano, and R E Lenski. 1994. "Long-Term Experimental Evolution in Escherichia-Coli .2. Changes in Life-History Traits During Adaptation To a Seasonal Environment." *American Naturalist* 144(3): 432–56.
- Wahl, Lindi M., and Anna Dai Zhu. 2015. "Survival Probability of Beneficial Mutations in Bacterial Batch Culture." *Genetics* 200(1): 309–20.
- Wiser, Michael J, and Richard E Lenski. 2015. "A Comparison of Methods to Measure Fitness in Escherichia Coli." *PLoS ONE* 10(5): 1–11.
- Yi, Xiao, and Antony M Dean. 2013. "Bounded Population Sizes, Fluctuating Selection and the Tempo and Mode of Coexistence." *Proc. Natl. Acad. Sci. USA* 110(42): 16945–50.

ANNEXES

SUPPLEMENTS DU CHAPITRE 2

S1 Text: Model validation

The one strain version of the model was fitted to experimental evolution data obtained from a 6 month-experimental evolution described in details in [1]. Six different strains differing by their ecological niche were placed in four different batch cultures regimes differing by their glucose initial concentration (0.01 or 0.15 g/mL in a 3% YNB with amino acid) and/or the length of the batch (48 or 96 hours). Cultures were incubated at 30 °C. Over a 6 month-period, at the end of each batch, 10^6 cells/mL were sampled from the culture and used to start a new batch. For each ancestor and each of the 4 selection regimes, we replicated the experiment 3 times leading to 72 evolved and 6 ancestors strains. At the end of the experiment, all ancestors and evolved strains were characterized in the rich-culture medium (0.15 g/mL). We fitted the model to the data characterizing each strain independently. Kinetics of glucose concentration $G_{\text{obs}}(t)$, ethanol concentration $E_{\text{obs}}(t)$, cells concentration $Z_{\text{obs}}(t)$ and the survival rate $s_{\text{obs}}(t)$ were monitored over a 96h batch culture. $Z_{\text{obs}}(t)$ is the total number of cells per mL including dead cells. $s_{\text{obs}}(t)$ is the survival rate and is estimated by CFU (colony forming units).

From the model, the total density of cells $\mathbf{Z(t)}$ in the media is given by:

$$\frac{dZ}{dt} = r_f \frac{G}{K_f + G} \cdot \exp\left(\frac{-E}{E_m}\right) \cdot N + r_r \frac{E}{K_r + E} \left(\frac{K_C}{K_C + G}\right) N$$

Observed ethanol concentration after 16 hours was always lower than at times 12h and 24h and we decided to discard these observations in the estimation procedure. Negative concentrations are set to zero for parameters estimation.

Experimental data from monoculture of evolved strains and their ancestors were used to estimate the 7 parameters (J_f , r_f , m , E_m , p , J_r , r_r) of the model as well as the initial

glucose concentration G_0 and the initial population size N_0 for 78 different strains. The remaining parameters (K_f , K_r and K_c) were set at $5 \cdot 10^{-4}$ g/mL. We checked that our estimates were not sensitive to the chosen value.

To estimate the parameters $\theta = (J_f, r_f, m, E_m, p, J_r, r_r, G_0, N_0)$, we used ABC-SMC algorithm [2]. As in classical ABC methods, a prior distribution is used to determine a distribution given a certain distance. In ABC-SMC, distance is progressively decreased, which allow a faster convergence. The algorithm is:

4. $i=1$, $\varepsilon_1=4$.
5. a. Sample a vector of parameters θ from the prior distribution $\pi(\theta)$ described in Table 2.
 - e. If $d(X_{obs}, X_{pred}(\theta)) < \varepsilon_1$, the vector is added to the population P_1 .
 - f. If the size of the population is inferior to 10,000, return to step 1.a.
 - g. Compute ε_2 as the 10th percentile of $\{d(X_{obs}, X_{pred}(\theta)), \theta \in P_1\}$. Every vector has the same weight $\omega_1(\theta)=1/10,000$. Set $i=2$.
6. a. Sample a vector of parameters θ^* from the previous population P_{i-1} weighted by ω_{i-1} .
 - h. θ^* is perturbed by a kernel K to obtain θ^{**} . As the choice of the perturbation kernel has little effect on the speed of convergence (Toni et al), we used a uniform distribution $\theta^{**} \sim U(0.9, 1.1) \theta^*$.
 - i. If $d(X_{obs}, X_{pred}(\theta^{**})) < \varepsilon_i$ and $\pi(\theta^{**}) \neq 0$, the vector is added to the population P_i .
 - j. If the size of the population is inferior to 10,000, return to step 2.a.
 - k. Compute ε_{i+1} as the 10th percentile of $\{d(X_{obs}, X_{pred}(\theta^{**})), \theta^{**} \in P_i\}$.

- l. Compute the weight $\omega(\theta^{**}) = \frac{\pi(\theta^{**})}{\sum \omega(\theta_{i-1}^j) K(\theta^{**} | \theta_{i-1}^j)}$ where θ_{i-1}^j are the vectors of the population P_{i-1} . Normalize the weight.
- m. If $\epsilon_{i+1} - \epsilon_i > 0.05$, set $i=i+1$ and go to step 2.a.

The final population defined the posterior parameters distribution $\pi(\theta | d(X_{obs}, X_{pred}(\theta^{**})) < \epsilon)$.

The distance $d(X_{obs}, X_{pred})$ between data and model predictions was calculated as the Euclidian distance between observed values of $G(t)$, $Z(t)$, $E(t)$ and $N(t)=Z(t).s(t)$ at time 3, 6, 9, 12, 16, 24, 48, 72 and 96 hours ($s(t)$ has only been measured at 24, 48, 72 and 96 hours) denoted X_{obs} and the model predictions denoted $X_{pred}(\theta)$, normalized by the observed values and the number of observations (n_N , n_Z , n_E and n_G) :

$$d(X_{obs}, X_{pred}) = \frac{1}{n_N} \sum_{i=1}^{n_N} \left(\frac{N_{obs,i} - N_{pred,i}}{N_{obs,i}} \right)^2 + \frac{1}{n_Z} \sum_{i=1}^{n_Z} \left(\frac{Z_{obs,i} - Z_{pred,i}}{Z_{obs,i}} \right)^2 \\ + \frac{1}{n_E} \sum_{i=1}^{n_E} \left(\frac{E_{obs,i} - E_{pred,i}}{E_{obs,i}} \right)^2 + \frac{1}{n_G} \sum_{i=1}^{n_G} \left(\frac{G_{obs,i} - G_{pred,i}}{G_{obs,i}} \right)^2$$

Figure 1 in the main text provides an example of the fit for one strain.

The goodness of fit was evaluated by calculating the residuals for each strain on each time of observation on the different variables as the difference between predicted and observed value, standardized by the observed value. Standardized residuals for a variable x were computed as:

$$res_x = \frac{1}{n_x} \sum_{i=1}^x abs \left(\frac{x_{obs,i} - x_{pred,i}}{x_{obs,i}} \right)$$

The goodness of the fit depended on the time of observation (see Fig S2) and on the strain (see Fig S3). For most time points, the residuals are centered on 0 with residuals smaller than 22% in absolute values (respectively 5%, 36%, and 80%) for living cells (respectively total cells density, glucose concentration and ethanol concentration) except for specific time points described below. Overall, the goodness of fit using our model was reasonably satisfactory although two recurrent biases were observed in the predictions.

For all strains, the ethanol concentration at time 16h was a major outlier and was discarded from fitted data. The fit to the resource concentration data was not as good after 24h, especially when glucose was still present at the end of the culture (96h). The reason for this degradation in the fit comes from the lack of ethanol consumption: indeed, the model does not switch to respiration when glucose concentration is insufficiently low. In this case, the model could only agree with one of the two variables (ethanol or glucose), leading to high residuals. This switch before depletion of glucose could be due to more complex mechanisms than a simple glucose inhibition or to the lack of another resource in the medium.

The total density of cells is overestimated at the beginning of the dynamics and is underestimated at the end while the density of living cells is underestimated (from 24 hours to the end of the batch). This could be due to the fact that the model does not switch to respiration. As respiration does not start, the mortality is not compensated, and population decreases too rapidly. Underestimation of the density of living cells could also be due to the choice of a constant mortality rate over the batch. Observed mortality rate could be high before 24h and then decrease. A constant rate will therefore lead to the underestimation of living cells. But due to a lack of information, we kept it constant.

REFERENCES

- [1] Spor A, Kvitek DJ, Nidelet T, Martin J, Legrand J, Dillmann C, Bourgaïs A, de Vienne D, Sherlock G, Sicard D. Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. *Evolution*. 2014;68(3):772–90.
- [2] Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J R society Interface*. 2009;6:187–202

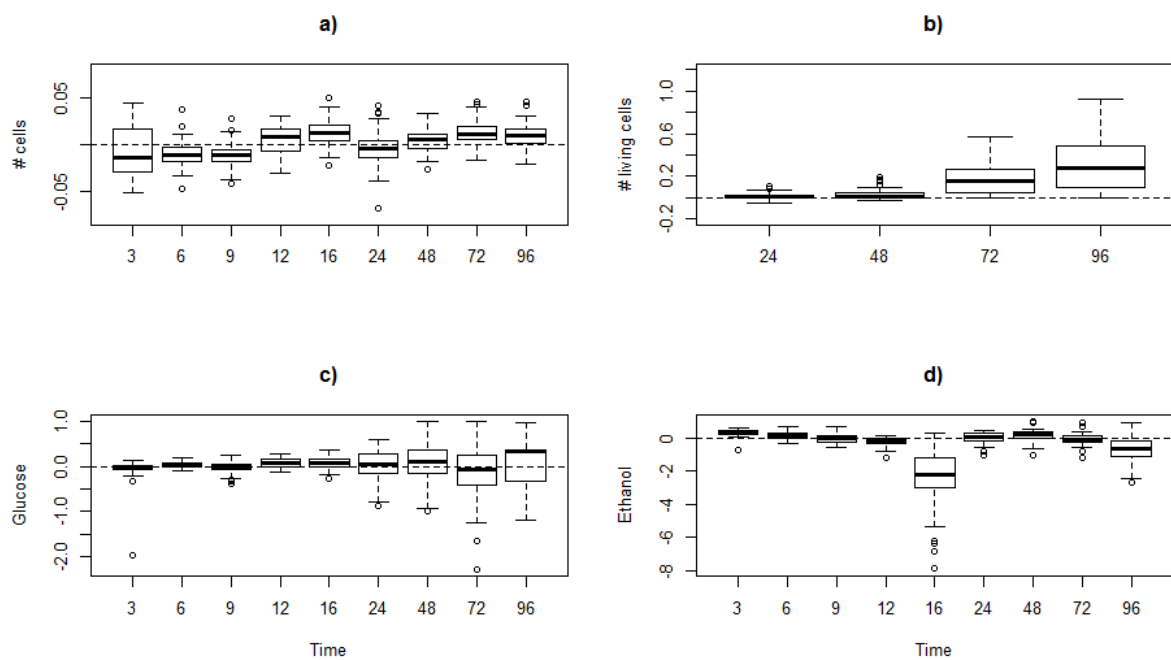


Figure S1: Distribution of residuals value depending on time. Residuals were computed as the difference between the observed value and the median of the predictions divided by the observed value. a) For the total density of cells, b) for the density of living cells, c) for the glucose concentration, d) for the ethanol concentration.

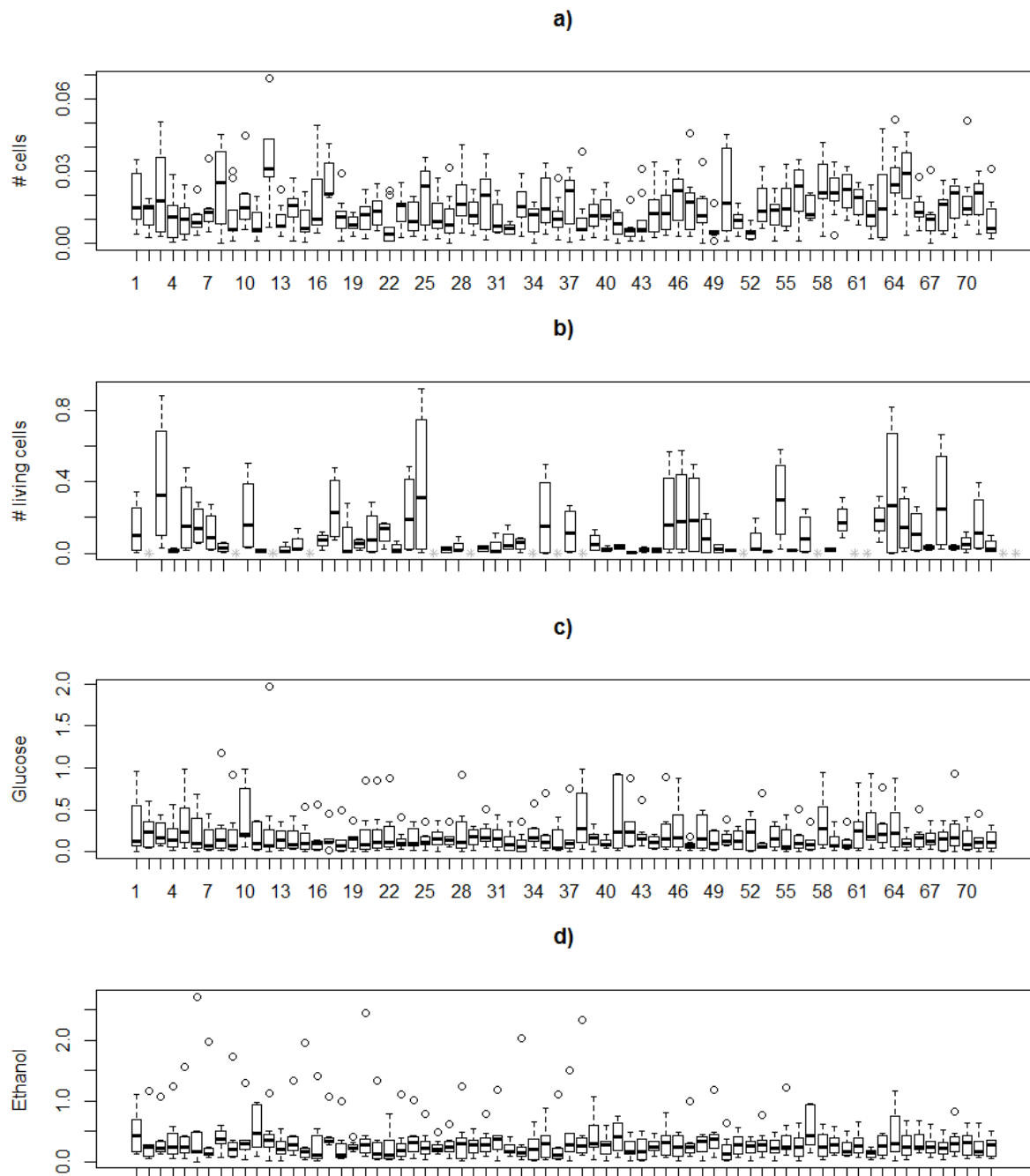


Figure S2: Distribution of the absolute value of residuals depending on strains.

a) For the total density of cells, b) for the density of living cells, c) for the glucose concentration, d) for the ethanol concentration. Gray stars indicate missing values.

Table S2: Bounds of uniform a priori distribution of parameters:

Parameter	m	J_f	r_f	E_m	J_r	r_r	p	K_f , K_r and K_c
Boundary set	0.01-0.2	10^{-11} - 10^{-10}	0.2-0.6	10^{-4} - 10^{-1}	10^{-11} - 10^{-9}	0.01-0.4	0-0.5	10^{-7} - 10^{-2}
reference		[1] [2] [3] [5] [6] [7]	[1] [2] [3] [4] [5] [6] [7]	[4] [7]	[1] [2] [3] [5]	[1] [2] [3] [5]	[1] [2] [3] [5] [6]	[3] [5] [6] [7]

m : mortality rate (h^{-1}), J_f : glucose consumption rate ($g \cdot h^{-1} \cdot cell^{-1}$), r_f : growth rate in fermentation (h^{-1}), E_m : tolerance to ethanol (g/mL), J_r : ethanol consumption rate ($g \cdot h^{-1} \cdot cell^{-1}$), r_r : growth rate in respiration (h^{-1}), p: ethanol/glucose yield (no unit), K_f : saturation constant for glucose, K_r : saturation constant for glucose, K_c : inhibition constant.

[1] Hagman A, Säll T, Compagno C, Piskur J. Yeast “make-accumulate-consume” life strategy evolved as a multi-step process that predates the whole genome duplication. PLoS One. 2013;8(7).

[2] Spor A, Nidelet T, Simon J, Bourgaïs A, de Vienne D, Sicard D. Niche-driven evolution of metabolic and life-history strategies in natural and domesticated populations of *Saccharomyces cerevisiae*. BMC Evol Biol. 2009;9:296.

[3] Lencastre Fernandes R, Carlquist M, Lundin L, Heins A-L, Dutta A, Sørensen SJ, et al. Cell mass and cell cycle dynamics of an asynchronous budding yeast population: experimental observations, flow cytometry data analysis, and multi-scale modeling. Biotechnol Bioeng. 2012 Mar;110(3):812–26.

[4] Antocea AO, Nămolesanu IC, Radoi FM. Comparative study regarding the ethanol resistance of some yeast strains isolated from Romanian vineyards. Rom Biotechnol Lett. 2011;16(2):5981–8.

[5] MacLean RC, Gudelj I. Resource competition and social conflict in experimental populations of yeast. Nature. 2006;441:498–501.

[6] Elbing K, Larsson C, Bill RM, Albers E, Snoep JL, Boles E, et al. Role of hexose transport in control of glycolytic flux in *Saccharomyces cerevisiae*. Appl Environ Microbiol. 2004;70(9):5323–30.

[7] Hoppe G, Hansford G. Ethanol inhibition of continuous anaerobic yeast growth. Biotechnol Lett. 1982;4(1):39–44.

S3 Text: Fitness function in the simple case:

In order to calculate fitness, the model has been simplified. The strains have a constant behaviour while a resource is available. They have an exponential growth at a constant rate ($r_f - m$) while glucose concentration is non-null, then an exponential growth at a constant rate ($r_r - m$) while ethanol concentration is non-null, and finally an exponential decrease at constant rate m . This case corresponds to the Equation 11, with $E_m \rightarrow \infty$, $K_r \rightarrow 0$, $K_f \rightarrow 0$ and $K_c \rightarrow 0$. T^{ferm} is the time where glucose is exhausted, and T^{resp} is the time where ethanol is exhausted. Therefore T^{ferm} is the length of the fermentation season, and $T^{resp} - T^{ferm}$ is the length of the respiration season. In this case,

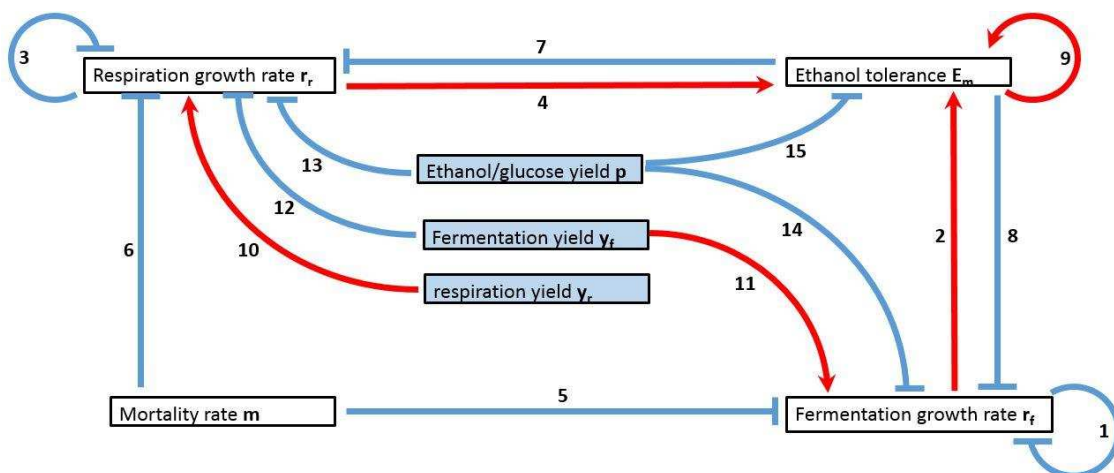
$$N_i(T) = N_i(0) \cdot e^{(r_{f,i} - m_i) \cdot T^{ferm}} e^{(r_{r,i} - m) \cdot (T^{resp} - T^{ferm})} e^{(-m_i) \cdot (T - T^{resp})}$$

So,

$$\begin{aligned} W_{1/2}(T) &= \ln \left(\frac{N_1(T)}{N_1(0)} \right) - \ln \left(\frac{N_2(T)}{N_2(0)} \right) \\ &= (r_{f,1} - r_{f,2}) \cdot T^{ferm} + (r_{r,1} - r_{r,2}) \cdot (T^{resp} - T^{ferm}) - (m_1 - m_2)T \end{aligned}$$

Figure S4: Feedbacks between selection and traits.

Traits in white boxes are selected traits. Traits in blue boxes are traits which are not directly selected. Connection between two boxes indicates that changes in one trait will affect strength of the selection on the other. The direction of the interactions is indicated by arrows. Red Arrows indicates that selection becomes more important if the trait is modified in the direction of the selection. A blue bar indicates that selection becomes less important if the trait is modified in the direction of the selection. For traits which are not selected, red arrows (blue bars) indicate a positive (resp. negative) correlation between the traits and the strength of the selection. Arrows are numbered in the text. Each arrow can be read as “selection tends to modify [first trait] in a given direction, which modifies a season length, and thus modifies the strength of selection on the [second trait]”. For instance, arrow 1 means that: selection tends to increase fermentation growth rate, which decreases fermentation season length, and so strength of selection on growth rate decreases. Similarly for arrow 5: selection tends to decreases mortality rate, which decreases fermentation season length, and so strength of selection on growth rate decreases.



This graph sums up the different feedbacks predicted by our model during evolution in batch culture. Interactions 1, 3 and 9 are direct feedbacks: the trait is selected and

affects its own selection strength. Interactions 2, 4, 5, 6, 7, 8 are indirect feedbacks: the trait is selected and modulates selection strength on another trait. Finally, interactions 10 to 15 are feedbacks due to non-selected traits: traits in blue box are not selected, but if they change (due to drift or links with other traits), they will affect selection strength on other traits. In each case, the selection strength is modulated via season lengths.

Table S5: Phenotypes of strains A, B and C.

	A	B	C
Mortality rate m	0.1	0.104	0.083
Fermentation growth rate r_f	0.6	0.4	0.5
Respiration growth rate r_r	0.3	0.4	0.25
Ethanol tolerance E_m	0.1	0.5	0.15
Inhibition constant K_c	$2 \cdot 10^{-3}$	$9 \cdot 10^{-4}$	$7 \cdot 10^{-3}$
Glucose consumption rate J_f	$7 \cdot 10^{-11}$	$2 \cdot 10^{-10}$	$2 \cdot 10^{-11}$
Ethanol consumption rate J_r	$4 \cdot 10^{-13}$	$6 \cdot 10^{-13}$	$4 \cdot 10^{-13}$
Ethanol/glucose yield p	0.4	0.4	0.1
Fermentation affinity K_f	$5 \cdot 10^{-4}$	$5 \cdot 10^{-4}$	$5 \cdot 10^{-4}$
Respiration affinity K_r	$5 \cdot 10^{-4}$	$5 \cdot 10^{-4}$	$5 \cdot 10^{-4}$

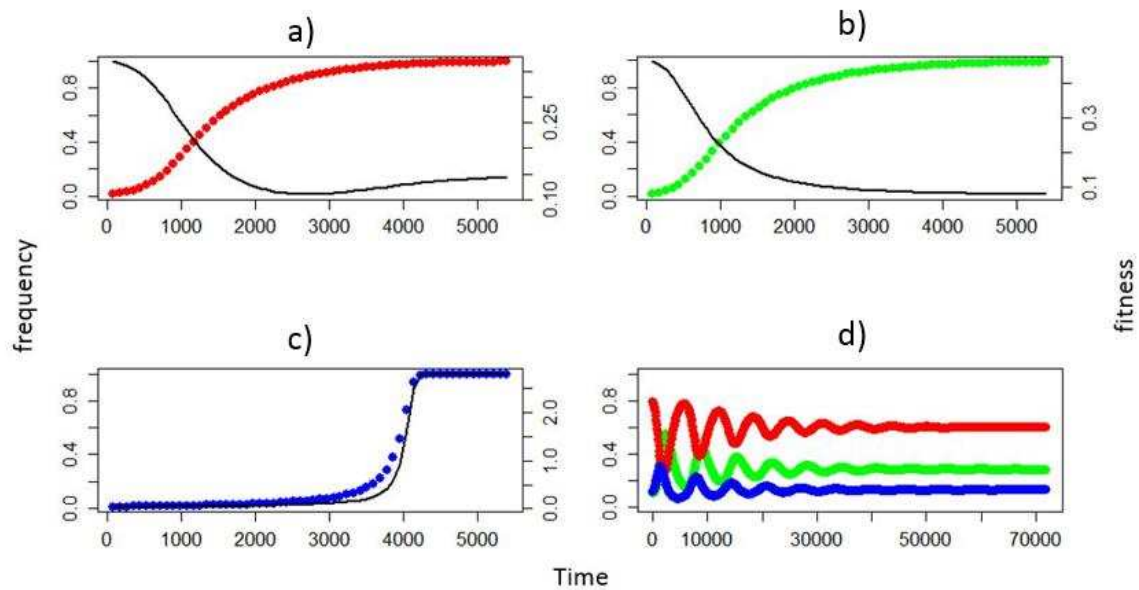


Figure S5: Competition between two or three strains in successive batches of 90 hours. Panel a) to c) represent initial frequency of a strain in each successive batch (pointed lines), and fitness of this strain (continuous lines) when two of the three strains are in competition. a) Frequency and fitness of strains B in competition with strain A, b) Frequency and fitness C in competition with strain B, c) Frequency and fitness of strains A in competition with strain C during 60 batches of 90 hours. d) A) Initial Frequency of strains A (blue), strain B (red), and strain C (green) at the beginning of each batch.

In the example illustrated in Figure S5, strain A is favoured compared to the other strains in fermentation but it leads to an environment where respiration is favoured whenever it is the resident strain. Strain B is favoured compared to the other strains in respiration but it leads to an environment where mortality is favoured whenever it is the resident strain. Finally strain C is favoured compared to the other strains in the mortality season but it leads to an environment where fermentation is dominant whenever it is the resident strain. As a consequence each strain is fitter than just one of the other two strains and thus no strain is “the fittest” (Fig S5 a-c).

Consider now what happens when the three strains are introduced together into the same batch. No strain can take over and eliminate the other two, instead one has co-existence of all the strains. Depending on the situation, this could translate into having an equilibrium whereby from one batch to the other the three strains would maintain the same frequencies (of course the frequency will vary within a batch) or one could have more complicated situations where from one batch to the next different strains would have the highest frequency. This is illustrated in Figure S5 d where we show an example of the frequencies of each of the three strains in successive batches; there the behaviour corresponds to a damped oscillation towards an equilibrium.

SUPPLEMENTS DU CHAPITRE 3

SI 1: Simplified case.

In order to solve the equation of the equilibrium, the model has been simplified. The strains have a constant compartment while a resource is available, and ethanol has no toxic effect on yeast cells. They have an exponential growth at a constant rate ($r_f - m$) while glucose concentration is non-null, then an exponential growth at a constant rate ($r_f - m$) while ethanol concentration is non-null, and finally an exponential decrease at constant rate m . This case corresponds to the Equation (1), with $E_m \rightarrow \infty$, $K_r \rightarrow 0$, $K_f \rightarrow 0$ and $K_c \rightarrow 0$. We also replace J_f by y_f in order to simplify the writing. However the yield is a metabolic trait too.

In this case, as the mutant does not influence seasons length, they can be calculated:

$$T_{ferm} = \frac{1}{(r_{f,res} - m)} \cdot \ln \left(\frac{y_{f,res} \cdot G_0}{N_0} + 1 \right)$$

$$T_{resp} = \frac{1}{(r_{r,res} - m)} \cdot \ln \left(\frac{y_{r,res} \cdot G_0}{y_{f,res} \cdot G_0 + N_0} + 1 \right)$$

The phenotype is described by two traits, y_f and r_f and we impose $m = f(r_f) = a \cdot r_f^c + b$, where a , b are real coefficients, and c is a positive real number not equal to 1. This form has been chosen to because the convexity of the function is determined by c : if $c > 1$ the function is concave, if $c < 1$, the function is convex, and if $c = 1$, the function is linear. Moreover, a and b are determined by the maximal and minimal values of m and r_f . So the trade-off by a single parameter, which correspond to the convexity.

According to Table 3: $f'(r_f) = \frac{T_{ferm}}{T}$.

$$\text{So } a \cdot c \cdot r_{f,opt}^{c-1} = \frac{T_{ferm}}{T}$$

$$\text{So } a \cdot c \cdot r_{f,opt}^{c-1} \cdot T = \frac{1}{(r_{f,opt} - a \cdot r_{f,opt}^c - b)} \cdot \ln \left(\frac{y_f \cdot G_0}{N_0} + 1 \right)$$

$$\text{And so } \boxed{y_f = \frac{N_0}{G_0} (e^{a \cdot c \cdot T r_{f,opt}^{c-1} (r_{f,opt} - a \cdot r_{f,opt}^c - b)} - 1)}$$

The stability is determined by the values of E and M .

$$\text{Here } E = -(c - 1) a \cdot c \cdot r_{f,opt}^{c-2} \cdot T \text{ and } M = \frac{\partial T_{ferm}}{\partial r_f} = - \left(\frac{1 - \frac{T_{ferm}}{T}}{(r_f - m)} \right) T_{ferm} \text{ at the singular points.}$$

E is negative if $c > 1$, and M is always negative $(r-m) > 0$, $0 < T_f \leq T$ (fermentation season cannot be longer than the batch).

So if $c > 1$, the equilibrium is convergence and evolutionary stable, so the singularity is an optimal strategy.

If $c < \frac{\left(1 - \frac{T_{ferm}}{T}\right) \frac{1}{r_f}}{\left(1 - \frac{f}{m}\right) \frac{1}{T}} < 1$, the equilibrium is convergence stable, but evolutionary unstable, so the singularity is a branching.

If $1 > c > \frac{\left(1 - \frac{T_{ferm}}{T}\right) \frac{1}{r_f}}{\left(1 - \frac{f}{m}\right) \frac{1}{T}}$, the singularity is a repellor.

In the range of the parameters we use, m/r_f is around 0.6 and T_{ferm}/T is around 0.5, and $T=48$. So branching is possible if $0.98 < c < 1$.

If the trade-off is linear ($c=1$), there is no optimal values of r_f . The singular strategy is $a = \frac{T_{ferm}}{T}$.

In this case, the evolution will stop as soon as the season lengths are optimal. In this case the composition of the population cannot be predicted: Many phenotypes can coexist, as long as the season length are correct. In this case, stability has no sense. This case is intermediate between selected and non-selected traits: no values is optimal, but selection still operate.

SI 2: Sign of M.

The aim is to prove that M is negative for the three cases, where two growth rates are linked by a trade-off.

trait 1	trait 2	M
r_f	r_r	$\frac{\partial W_{1/2}(T)}{\partial r_{f,res} \partial r_f}(r_f) = \frac{\partial T_{ferm}}{\partial r_{f,res}} + f'(r_f) \frac{\partial T_{resp}}{\partial r_{f,res}}$
r_f	m	$\frac{\partial W_{1/2}(T)}{\partial r_{f,res} \partial r_f}(r_f) = \frac{\partial T_{ferm}}{\partial r_{f,res}}$
r_r	m	$\frac{\partial W_{1/2}(T)}{\partial r_{r,res} \partial r_r}(r_r) = \frac{\partial T_{resp}}{\partial r_{r,res}}$

Case $m=f(r_i)$ (r_i can be r_f or r_r) :

$(r_{i,res}-m_{res})$ is the net growth rate of population during season i. So the higher the net growth rate is, the faster the population grows and the shorter the season is. As a consequence

$$\frac{\partial T_i}{\partial (r_{i,res}-m_{res})} < 0$$

$$\frac{\partial T_i}{\partial r_{i,res}} = \frac{\partial T_i}{\partial (r_{i,res} - f(r_{i,res}))} \cdot \frac{\partial (r_{i,res} - f(r_{i,res}))}{\partial r_{i,res}}$$

$$\text{So } \frac{\partial T_i}{\partial r_{i,res}} = \frac{\partial T_i}{\partial (r_{i,res}-m_{res})} (1 - f'(r_{i,res})). \text{ At the equilibrium, } f'(r_{i,res}) = \frac{T_i}{T}$$

So $\frac{\partial T_i}{\partial r_{i,res}} = \frac{\partial T_i}{\partial (r_{i,res}-m_{res})} (1 - \frac{T_i}{T})$. And $\frac{T_i}{T} \leq 1$ (the fermentation season is shorter than the batch).

$$\text{So } \frac{\partial T_i}{\partial r_{i,res}} < 0.$$

Case $r_r=f(r_f)$:

$$\text{In this case, } M = \frac{\partial T_{ferm}}{\partial r_{f,res}} + f'(r_f) \frac{\partial T_{ferm}}{\partial r_{r,res}} + f'(r_f) \left(\frac{\partial T_{resp}}{\partial r_{f,res}} + f'(r_f) \frac{\partial T_{resp}}{\partial r_{r,res}} \right).$$

$\frac{\partial T_{ferm}}{\partial r_{f,res}} = 0$, as fermentation happen before respiration.

$$\text{So } M = \frac{\partial T_{ferm}}{\partial r_{f,res}} + f'(r_f) \frac{\partial T_{resp}}{\partial r_{f,res}} + (f'(r_f))^2 \frac{\partial T_{resp}}{\partial r_{f,res}}$$

$$\frac{\partial T_{ferm}}{\partial r_{f,res}} < 0, \frac{\partial T_{resp}}{\partial r_{f,res}} < 0, \frac{\partial T_{resp}}{\partial r_{f,res}} < 0 \text{ and at the equilibrium, } f'(r_{f,res}) \left(= -\frac{T_{ferm}}{T_{resp}} \right) < 0.$$

So the sign of M cannot be determined.

Demonstration of three sufficients condition to have $M < 0$:

- If y_f is fixed, $\frac{\partial T_{resp}}{\partial r_{f,res}} = 0$ and so $M < 0$.
- It can be noticed that $M = a.(f')^2 + b.(f') + c$ with $a < 0$.

$$\text{So if } \left(\frac{\partial T_{resp}}{\partial r_{f,res}} \right)^2 < 4. \frac{\partial T_{resp}}{\partial r_{f,res}} \cdot \frac{\partial T_{ferm}}{\partial r_{f,res}}, \forall f'(r_{f,res}), M < 0.$$

$$\text{But we do not know if } \left(\frac{\partial T_{resp}}{\partial r_{f,res}} \right)^2 < 4. \frac{\partial T_{resp}}{\partial r_{f,res}} \cdot \frac{\partial T_{ferm}}{\partial r_{f,res}}.$$

We can suppose that the effect of a growth rate on it season is higher than the effect on the other season ($\frac{\partial T_{resp}}{\partial r_{f,res}} < \frac{\partial T_{ferm}}{\partial r_{f,res}}$) and that a season is more affected by the growth rate during

this season, than by the other ($\frac{\partial T_{resp}}{\partial r_{f,res}} < \frac{\partial T_{resp}}{\partial r_{r,res}}$). So the relation would be always true, and

$M < 0$. We do not found any parameters values that does not verify $\left(\frac{\partial T_{resp}}{\partial r_{f,res}} \right)^2 < 4. \frac{\partial T_{resp}}{\partial r_{f,res}} \cdot \frac{\partial T_{ferm}}{\partial r_{f,res}}$.

$$\text{For instance in case c), } \frac{\partial T_{resp}}{\partial r_{f,res}} + 0.25 \frac{\partial T_{resp}}{\partial m_{res}} = -24.7, \frac{\partial T_{ferm}}{\partial r_{f,res}} + 0.3 \frac{\partial T_{ferm}}{\partial m_{res}} = -25.4 \text{ and}$$

$$\frac{\partial T_{resp}}{\partial r_{f,res}} + 0.3 \frac{\partial T_{resp}}{\partial m_{res}} = -1.3 \text{ (and } \frac{\partial T_{ferm}}{\partial r_{f,res}} + 0.25 \frac{\partial T_{ferm}}{\partial m_{res}} = 0 + 0.25 \frac{\partial T_{ferm}}{\partial m_{res}} = 8.6).$$

$$\text{So } (-1.3)^2 < 4(-24.7)(-25.4)$$

- If $\frac{\partial \left(\frac{T_{ferm}}{T_{resp}} \right)}{\partial r_{f,res}} < 0$, $M < 0$ at the equilibrium.

Proof that if $\frac{\partial(\frac{T_{ferm}}{T_{resp}})}{\partial r_{f,res}} < 0$, $M < 0$ at the equilibrium:

$$\frac{\partial(\frac{T_{ferm}}{T_{resp}})}{\partial r_{f,res}} = \frac{T_{resp} \frac{\partial(T_{ferm})}{\partial r_{f,res}} - T_{ferm} \frac{\partial(T_{resp})}{\partial r_{f,res}}}{T_{resp}^2}.$$

$$\text{So } \frac{\partial(\frac{T_{ferm}}{T_{resp}})}{\partial r_{f,res}} < 0 \Leftrightarrow \frac{T_{resp} \frac{\partial(T_{ferm})}{\partial r_{f,res}} - T_{ferm} \frac{\partial(T_{resp})}{\partial r_{f,res}}}{T_{resp}^2} < 0.$$

$$\text{So } T_{resp} \frac{\partial(T_{ferm})}{\partial r_{f,res}} - T_{ferm} \frac{\partial(T_{resp})}{\partial r_{f,res}} < 0$$

$$\text{So } \frac{\partial(T_{ferm})}{\partial r_{f,res}} - \frac{T_{ferm}}{T_{resp}} \frac{\partial(T_{resp})}{\partial r_{f,res}} < 0 \text{ (as } T_{resp} > 0).$$

$$\text{As } \left(-\frac{T_{ferm}}{T_{resp}}\right)^2 \frac{\partial T_{resp}}{\partial r_{f,res}} < 0, \frac{\partial(T_{ferm})}{\partial r_{f,res}} - \frac{T_{ferm}}{T_{resp}} \frac{\partial(T_{resp})}{\partial r_{f,res}} + \left(-\frac{T_{ferm}}{T_{resp}}\right)^2 \frac{\partial T_{resp}}{\partial r_{f,res}} < 0$$

At the equilibrium, $f'(r_{f,res}) = -\frac{T_{ferm}}{T_{resp}}$, so $M < 0$.

Titre : Modélisation des dynamiques adaptatives de la levure de boulanger *Saccharomyces cerevisiae* dans un environnement saisonnier

Mots clés : adaptation, environnement saisonnier, traits d'histoire de vie, composantes de la valeur sélective, modélisation mathématique, levure, *Saccharomyces cerevisiae*, boucles éco-évolutives

L'adaptation des individus à un environnement dépend d'une combinaison de caractères adaptatifs, les traits d'histoire de vie, qui impactent la valeur sélective. Pour comprendre comment les organismes s'adaptent à leur environnement, on peut étudier quelles sont les traits composants la valeur sélective et comment ils dépendent de l'environnement biotique et abiotique. Au cours de cette thèse, je me suis intéressé aux composantes de la valeur sélective dans un environnement saisonnier et à ses conséquences sur la dynamique évolutive des traits quantitatifs.

Pour cela, j'ai utilisé une approche de modélisation mathématique d'une évolution expérimentale de l'espèce modèle *Saccharomyces cerevisiae* en cultures successives en batch. La levure de boulanger *S. cerevisiae* ici étudiée présente un cycle de vie respiro-fermentaire : en présence de glucose, elle le consomme par fermentation tout en produisant de l'éthanol, qui sera consommé dans un deuxième temps par respiration. Les souches de levures évoluent au cours de cycles successifs de fermentation-respiration. A intervalles de temps réguliers, des cellules sont transférées dans un nouveau milieu contenant du glucose où elles effectuent un nouveau cycle. J'ai développé un modèle mathématique d'équations différentielles pour étudier quels sont les traits sélectionnés dans les différentes saisons dans ce dispositif expérimental et comment l'environnement abiotique, l'environnement biotique et les relations entre les traits, impactent leur évolution.

Dans un premier temps, j'ai développé et paramétré un modèle d'équations différentielles décrivant la dynamique d'une population multi-souches au cours d'un batch (chapitre 1). J'ai ensuite proposé une décomposition de la valeur sélective et étudié quels traits sont sous sélection, et comment les pressions de sélection changent avec la composition de la population (chapitre 2). Deux types de traits sélectionnés ont pu être mis en évidence : les traits d'histoire de vie, liés au taux de croissance et à la mortalité, et les traits de transition, qui correspondent à la façon dont les souches réagissent aux changements de l'environnement. J'ai également montré que l'importance de chacune des composantes de la valeur sélective est liée à ces traits et à des traits non sélectionnés, via la longueur des différentes saisons. Au cours de l'évolution, ces composantes sont modifiées ce qui modifie la force de la sélection sur chaque trait. Ce phénomène de boucles de rétroaction éco-évolutives permet de mieux comprendre pourquoi la valeur sélective est fréquence-dépendante.

Dans un second temps, j'ai utilisé des simulations d'un modèle de dynamique adaptative pour montrer que l'existence d'un trade-off entre deux traits dans la population ancêtre pouvaient entraîner l'émergence d'autres relations entre un trait sélectionné et un trait non-sélectionné au cours de l'évolution (chapitre 3).

Enfin, pour mettre en regard les prédictions issues de modèles théoriques et des observations expérimentales, j'ai analysé deux jeux de données à travers le prisme de mon modèle mathématique (chapitre 4). Le premier jeu de données concerne le phénotypage de souches évoluées en batch successifs et leurs ancêtres. L'estimation des paramètres du modèle pour chacune des souches du jeu de données et leur analyse montrent que les traits liés à l'éthanol, sa consommation et sa production ont été principalement sélectionnés. Le second jeu de données, obtenu à partir de compétitions entre plusieurs couples de souches aillant des traits d'histoire de vie contrastés, a permis de mettre en évidence des différences de valeur sélective entre souches et de les relier avec des différences de traits phénotypiques, en cohérence avec les prédictions théoriques.

Title : Modeling adaptive dynamics of the yeast *Saccharomyces cerevisiae* in a seasonal environment.

Keywords : adaptation, seasonal environment, life-history traits, fitness components, mathematical modeling, yeast, *Saccharomyces cerevisiae*, eco-evolutionary feedbacks

Abstract : Adaptation of species to their environment involves combinations of traits, and in particular life history traits, that influence an organism's selective value. To understand the complexity of adaptation, it is appropriate to decipher the contributions of traits to fitness in the presence of different biotic and abiotic environments. In this thesis, I have investigated fitness components when the environment is seasonal, revealing how such components drive the evolutionary dynamics of quantitative traits.

My work is based on the mathematical modeling of experimental evolutions in successive batch cultures of *Saccharomyces cerevisiae* (baker's yeast). The life cycle of this yeast species is of the respiration-fermentation type: (i) in the presence of glucose, it grows by fermentation, transforming glucose into ethanol; (ii) once glucose has been consumed, it grows by respiration, consuming this time ethanol. This sequence corresponds to the two « seasons » in a batch culture and leads to a cycle of successive batches if cells are periodically transferred into fresh medium. By using differential equations for the time courses, my thesis work shows how growth dynamics and environmental features (abiotic or biotic) generate selection pressures on the different traits during these successive seasons, thereby determining evolutionary trajectories.

To describe batch dynamics, I first developed and calibrated a set of differential equations describing the growth dynamics of a population of yeast cells throughout a batch, allowing for one or multiple strains to be present (Chapter 1). Based on this model where cells divide without changing genotype, I then showed that a strain's fitness can be understood in terms of just a few components that are easily specified mathematically. I was then able to determine which traits were under selection and how the corresponding selection pressures were affected by the abundances of each strain in the yeast population (Chapter 2). Selected traits were found to be of two types: life history traits associated with growth and mortality rates, and “transition” traits that correspond to the way a strain reacts to environmental change. I also showed that the contributions of the different fitness components are tied to both selected and non-selected traits via the lengths of seasons. Thus, during population dynamics arising across successive batches, these components change, modifying the selection pressure on each trait. One therefore has a feedback loop, revealing why fitness is frequency-dependent in this system.

Next, using the fitness decomposition, I studied adaptive dynamics in successive batch cultures. In such a framework where genotypic changes were allowed, and assuming that there was a trade-off between two traits, I showed that adaptive evolutionary dynamics could lead to the emergence of new relations between selected and non-selected traits (Chapter 3).

Furthermore, in order to compare my theoretical predictions to experimental results, I used mathematical and statistical models to analyze two datasets (Chapter 4). The first dataset provides trait measurements in “evolved” strains, i.e., strains obtained after evolution across successive batches, as well as of those same traits in the “ancestral” strains at the origin of the experimental evolution. Parameters inference for the different strains showed that selection had operated mainly on ethanol-related traits (production and consumption). A second dataset was obtained from batch experiments putting strains in competition with one another; the analysis showed that my theoretical modeling well predicted the roles of the different traits for determining the relative fitness of the strains.

