Numéro National de Thèse : 2017LYSEN055

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée par
l'Ecole Normale Supérieure de Lyon

Ecole Doctorale N° 340
BIOLOGIE MOLECULAIRE, INTEGRATIVE ET CELLULAIRE (BMIC)

Spécialité de doctorat : Génomique Fonctionnelle
Discipline : **Biologie**

Soutenue publiquement le 29/09/2017, par :
Jérôme SALIGNON

Genomics of fitness in periodic stress

Génomique de la prolifération cellulaire en stress périodique

---

Devant le jury composé de :

Pelet, Serge	Assistant Professeur	Université de Lausanne	Rapporteur
Sicard, Delphine	Directeur de Recherche	Université de Montpellier	Rapporteure
Chevin, Luis-Miguel	Chargé de Recherche	Université de Montpellier	Examinateur
Friedrich, Anne	Maître de Conférences	Université de Strasbourg	Examinatrice
Yvert, Gaël	Directeur de Recherche	Université de Lyon	Directeur de thèse

Thesis prepared in LBMC laboratory UMR 5239 - ENS DE LYON / CNRS

Prepared in the team *Genetics of Intra-Species Variations*

Under the supervision of Dr Gaël Yvert





Acknowledgements

Obtaining my PhD feels like a milestone on my longstanding objective to become a researcher. I would like to express my gratitude to the BMIC doctoral school and the ARC Foundation that made it possible by their financial support.

Furthermore, I want to thank the people that made it pleasant by great human interactions, and a good working environment. Particularly, I thank the current and past member of our team: Sébastien for our geeky discussions, Gérard for his brilliant jokes, Uschi for sharing her passion for science, Etienne for his easygoingness, Fabien for sharing his PhD experience with me, Claire for her support at late hours, Hélène for her kindness, Florent for his advice in computer science, Magali for her mentoring and contribution to the project and Chewbacca for a fruitful collaboration despite our numerous fights. It was also my pleasure to work with an efficient and cheerful administrative staff.

Additionally, I was fortunate to mix with awesome people – Axel, Clémence, Manon, Delphine, Flore, Marion, Charlotte, Wallis, Caro, Soizic, Vincent, Séverin, Sonia, Angélique, Sophie, Ludo, Thibault, Fabrice, Élodie, Julietta, Nicolas, Anissa, Éléonore, Julie, Lama, Brian, Éloise, Steven and others – and I thank them for the great times we had. I am surely gonna miss the pseudogame team.

I am grateful to Zhou for his wise advice on how to lead a PhD. I also acknowledge the contribution of Dr. Julien Gagneur for his precious scientific help and involvement in this project, and of Dr. Abderrahman Khila for his participation in my Thesis Advisory Committees.

Thanks are also due to Dr. Delphine Sicard and Dr. Serge Pelet for the evaluation of my thesis manuscript and defense, and to Dr. Luis-Miguel Chevin and Dr. Anne Friedrich for joining as well my Thesis Examination Committee.

I would like to extend my thanks to the members of my family for their support and their continuous attempts to try to catch the subject of my thesis. In addition, I owe to my father and my mother for instilling in me respectively curiosity and tenacity, two key skills for a researcher.

My utmost appreciation goes to Dr. Gaël Yvert whose unceasing mentorship shaped my scientific project and my critical thinking. I have been impressed and inspired by his professionalism, scientific skills and sharpness for communicating scientific results.

Last, but far from least, I am, indeed, indebted to my partner, Charlotte, for her unconditional support and love throughout my thesis. I also thank her for bearing with the uncertainty of going out with a researcher, and with my mood that sometimes fluctuated with my results.

Genomics of fitness in periodic stress

Abstract

Organisms live in dynamic environments. However, most experimental approaches study the function and selection of genes in steady environments. Therefore, natural selection acting on fluctuating environments remains poorly understood. The objective of my project was to determine if some genes are especially important for fitness (growth rate) of yeast cells in oscillating environments. A genomic screen, based on an automation of micro-cultures and on a multiplexing of sequencing libraries, allowed me to measure fitness of thousands of null mutants in periodic stress conditions. I found that predictability of fitness in periodic stress, from fitness in steady environments, varies depending on the specific genes and conditions considered. This way, some mutants have similar growth in steady conditions, and different growth in dynamic conditions. Curiously, some genes play a bivalent role: they strongly favor growth during slow fluctuations, and reduce it during fast fluctuations. I also observed many mutants with higher growth than expected at the highest frequencies of fluctuations. This effect can be partially explained by a loss of environmental sensitivity of those mutants, that continue to divide quickly despite the presence of a stress. Those results show how natural selection can act on mutations in fluctuating environments. They open the door to mechanistic studies of the predictability of fitness in periodic environments.

Génomique de la prolifération cellulaire en stress périodique

Résumé en Français

Les organismes vivent dans des environnements dynamiques. Or la plupart des approches expérimentales étudient la fonction et la sélection des gènes dans des environnements statiques. De ce fait, la sélection naturelle agissant en environnements fluctuants reste mal comprise. L´objectif de mon projet a été de déterminer si certains gènes sont particulièrement importants pour la fitness (taux de croissance) de cellules de levures en environnements oscillants. Un crible génomique, basé sur une automatisation de micro-cultures et sur un multiplexage de banques de séquençage, m´a permis de mesurer la fitness de milliers de mutants nuls en conditions de stress périodique. J´ai trouvé que la prédictibilité de la fitness en environnements périodiques, à partir de la fitness en environnements statiques, diffère selon les gènes et les conditions. Ainsi, certains mutants présentent des croissances similaires en conditions statiques mais différentes en conditions dynamiques. Curieusement, quelques gènes jouent un rôle bivalent : ils favorisent fortement la croissance lors de fluctuations lentes et ils la défavorisent lors de fluctuations rapides. J´ai également observé de nombreux mutants avec une croissance plus élevée qu´attendue aux fréquences de fluctuations les plus rapides. Cet effet s´explique partiellement par une perte de sensibilité environnementale de ces mutants, qui continuent à se diviser rapidement malgré la présence d´un stress. Ces résultats montrent comment la sélection naturelle agit sur les mutations en environnements fluctuants. Ils ouvrent la porte à des études mécanistiques de la prédictibilité de la fitness en environnements périodiques.

Table of Contents

ABSTRACT....7
RÉSUMÉ EN FRANÇAIS....9
1 INTRODUCTION....15
I Yeast genetics....15
A Introduction to yeast biology....15
B The Yeast Deletion Library and powerful approaches to measure fitness in yeast....16
C Yeast response to osmotic stress....20
II Adaptation to a single environmental change....21
A Genetic adaptation and phenotypic plasticity....21
B Tracking the dynamics of adaptive mutations in microbes....22
III Adaptation to fluctuating environments....23
A Strategies of adaptation to fluctuating conditions....23
i Main strategies of adaptation to fluctuating environment....23
ii Bet-hedging....24
A Different types of bet-hedging....24
B Examples....25
C Mechanisms....26
iii Types of environmental changes determine the optimal strategies of adaptation....27
B Mechanisms of transcriptional anticipation and memory of environmental changes....28
i Transcription during a first stress in anticipation of a second stress....29
A The Environmental Stress Response and transcriptional anticipation....29
B Yeast screens for genes expressed in anticipation of future stresses....31
ii Mechanisms of memory of the first stress....34
A Inheritance of long-lived memory factors....35
1 Proteins....35
2 Other potential inherited memory factors....37
B Memory in the nucleus....38
1 Propagation of chromatin states....38
2 Genetic circuits and feedback in signaling pathways that generate stable states....39
3 Memory at the nuclear periphery – reinduction memory....41
IV Artificial fluctuations to characterize biological dynamics....42
A Experimental evolution in periodic fluctuations to characterize the evolution of genetic variance....42
B Applying methods from engineering to predict biological dynamics....44
i Introduction to Frequency Response Analysis and System Identification....45
ii Application of System Identification to characterize network dynamics....47
V Genomics of yeast fitness in periodically fluctuating environments....49
A What can we learn by experimentally studying fitness in periodically fluctuating conditions?....49
B Low throughput measures of fitness in fluctuating conditions....51
C My project: yeast genomics of fitness non-additivity....56

2 METHODS.....58
3 RESULTS.....60
I Submitted publication.....60
II Summary of the project.....60
III Four genomic screens.....62
A Experiment.....62
i Beginning of the project.....62
ii Description of the experiment.....63
iii Growth of populations during fluctuation experiments.....64
iv Solving an apparent issue of PCR amplification of the barcodes.....67
B First, sequencing results.....70
i Quality of sequencing.....70
ii Reads parsing.....71
iii Quality controls.....73
A Spike-in: reliable quantification of mutants in a large range of counts.....73
B Filtering data.....74
C Data normalisation.....75
D Computation of Fold Changes.....77
E Expected observations for well-known mutants.....78
F Correlations between replicates: highest reproducibility for the Salt and Sodium Metabisulfite experiment.....79
iv Principal Component Analysis (PCA): days and conditions discriminate populations of the salt experiment.....80
C Resequencing.....82
i Resequencing increased the power to detect mutants with subtle effects.....82
ii Data filtering.....83
iii Correlations between replicates.....84
iv Number of paired replicates of *steady* conditions used to compute fitness.....85
D Generalised Linear Model (GLM).....86
E Genomic analysis.....88
i No general relationship between expected and measured fitness at the 6 hours period 88
ii No general relationship between Antagonistic Pleiotropy and fitness *nonlinear*ity..90
iii Fitness nonlinearity as a function of fluctuating period.....91
F Detailed plots of some interesting nonlinear mutants.....93
i Biases that can lead to false positives.....93
A Mutants with aberrant counts in the steady conditions at day 0.....93
B Mutants with aberrant counts in the steady conditions after day zero.....94
C Mutants with very low counts or no counts at day 0.....95
D Mutants with very high growth in all conditions.....96
E Mutants with very slow growth in all conditions.....97
F Conclusions.....98
ii *vhr1*: pleiotropy but condition-dependent response to oscillations.....99
iii *thi2* in sodium Metabisulfite: strong antagonistic nonlinear effect across periods 102
iv *met12* in Methionine: nonlinear loser at short periods.....103
IV Secondary Screen: single-mutant measurements in periodic salt stress.....104

A Flow-cytometry assays: fitness relative to Wild-Type....104
B Absolute fitness....107
i Optical density measurements suggest that absolute fitness may not be transgressive 107
ii Absolute fitness estimated by flow-cytometry....110
C Negative and non-reproducible results....113
i *hog1* mutant showed non-reproducible results across experiments....113
ii Haploid strains were not transgressive....114
iii Some complementations did not rescue the transgressive phenotype....116
D Generality of transgressive effect for the *pde2* and *tom7* mutants in other environmental fluctuations....118
V Analysing co-fitness data to group mutants with similar phenotypic profiles....120
A *tom7* group: a potential common function related to mitochondria....122
B *trm1* group: no obvious cellular function....124
C *pde2* group: only one strain highly co-fit....126
D YCR group: mutations in a region of the genome results in a strong transgressive winner phenotype....127
E Loser group: no obvious cellular functions....130
VI Deviation between expected and measured fitness: costs and benefits of ignoring environmental stress....131
4 DISCUSSION....135
I Environmental insensitivity can be an adaptive strategy in artificial periodic salt stress 135
II Low data quality in other experiments prevent from drawing general principles about fitness in periodic environments....136
III Hypothesis to explain the behavior of selected *nonlinear* mutants....137
A Mutants with a frequency-dependent phenotype....137
A Vhr1....138
B Thi2....139
B Met12: a gene that may favor fitness in periodic methionine stress....140
C Groups of nonlinear mutants: implication of mitochondria?....141
D Interactions between time and environment for selected mutants: the Pde2p protein has a long half-life....145
IV Perspectives....147
A Single-cell approaches to characterize nonlinear mutants....147
B Perspectives for the exploration of the impact of environment and time on phenotypes 148
i Periodic conditions can help to set up models more robust to environmental dynamics....148
ii How to improve prediction of fitness in periodic conditions from fitness in steady conditions?....149
iii Linearity, homogeneity and additivity of fitness – towards novel assays to detect environmental interactions?....151
5 CONCLUSIONS....153
REFERENCES....155
RÉSUMÉ DE CHAQUE PARTIE EN FRANÇAIS....169

1 Introduction

I Yeast genetics

A Introduction to yeast biology

Saccharomyces cerevisiae is a species of yeast. This eukaryotic unicellular organism of 5 to 10 µm of width does not divide symmetrically, but by making "buds"-like shapes, which is why it is called the "budding" yeast. Even though *S. cerevisiae* is invisible to the naked eye, it has been used for thousands of years for the fabrication of beers, wine and for baking. This is due to its top fermenting abilities. More recently, the budding yeast has been used in biotechnologies for the bioproduction of compounds, as engineered yeast can be more stable than bacteria, and for research as a model organism.

Yeast is a popular model organism since it is a unicellular organism with a small generation time (typically less than 2 hours in standard medium) that can easily be manipulated; allowing powerful experimental techniques. In addition, it is a eukaryote and thus shares many similarities with human cells, such as genes, cellular conformation, metabolism… As a result, research in *S. cerevisiae* already culminated in 4 nobel prices in Medicine in this millennium: Hartwell, Hunt, and Nurse in 2001 for their work on cell cycle regulators; Blackburn, Greider, and Szosta in 2009 for their work on telomeres and telomerase, Rothman, Schekman, and Südhof in 2013 for their work on vesicle trafficking, and Yoshinori Ohsumi in 2016 for his work on autophagy. *Candida albicans* and *Schizosaccharomyces pombe* are two other popular yeast model organisms. In this thesis, if not specified otherwise, we will refer to *S. cerevisiae* when writing the word "yeast".

S. cerevisiae has been the first eukaryotic organism to be sequenced, back in 1996 (Goffeau et al., 1996). This choice was motivated by a small genome size: 12 million base pairs with 6275 genes (but only ~ 5800 are functional) organised in 16 chromosomes. Yeast Open Reading Frames (ORF) are named with the form Y$$###$, with dollars being letters and hashtags being numerals. The first letter indicates the chromosome (from A to P), the second letter the right or left arm of the chromosome (R or L), the numerals indicate the order on the chromosome arm (irrespective of strand), and the last letter indicates the Watson or Crick strand (C or L). The community of researchers working in yeast benefits from a comprehensive website; the Saccharomyces Genome Database (SGD) that gathers various types of data (DNA sequence, proteins, phenotypes, literature…) in an easily searchable fashion (Cherry et al., 2012a).

As mentioned above, yeast shares many conserved genes with humans: 31% of yeast genes have human homologs (Botstein et al., 1997). Strikingly, a recent study showed that replacing essential yeast genes (genes that yeast needs to survive in standard conditions) by their human orthologs could rescue about half of them (Kachroo et al., 2015). Moreover, membership to a conserved pathway was as important, if not more, as the level of protein conservation, for the success of the rescue. However, essential genes are more evolutionarily conserved than other yeast gene categories, as ~70% have a

human homolog. And yeasts have specificities related to their lifestyle.

Wild *S. cerevisiae* cells are living in ripe or rotten fruits or in insects, at a temperature of 30 to 35 degrees. This yeast is a sessile organism and thus it relies on insects, animals or natural elements to disperse. This also implies a higher degree of exposure to adverse conditions, and thus strong stress response programs to face fluctuating stressful conditions. On the contrary, *C. albicans* that lives within humans or other animal hosts shows a reduced gene expression stress response (López-Maury et al., 2008). *S cerevisiae* has the rare ability to be able to do both fermentation and respiration. Interestingly, when there is ample oxygen and fermentable sugars (maltose, fructose and especially glucose) it will choose to use fermentation, even though it produces much less energy than respiration. This paradox is resolved when looking at the yeast ecology: during fermentation yeast produces ethanol which kills most competitor micro-organisms. Indeed, yeast can survive high ethanol concentration and even metabolize ethanol as a carbon source after exhaustion of better carbon sources. The switching from a rapid fermentative growth on a rich carbon source to a slower growth by aerobic respiration using ethanol is called the diauxic shift.

Yeast can live as diploids or haploids. As a haploid, budding yeast has two sex types that can mate together: MAT a and MAT alpha. Under stressful conditions, a diploid cell may enter meiosis and sporulate to generate four haploid cells. This allows to generate genetic diversity that can increase the chances of survival to new adverse conditions.

B The Yeast Deletion Library and powerful approaches to measure fitness in yeast

“Darwinian” fitness, “selective value”, or just fitness, (often denoted w or ω**)** can be defined as a quantitative measure of the ability of an organism/a genotype to reproduce in a given environment. For asexual organism, it is straightforward to assign a fitness value to a genotype in a controlled environment. However, for sexual organisms, genotypes are recombined at each generation. In this case, an alternative strategy is to assign fitness values to alleles by looking at the reproductive success of all individuals bearing them. Fitter alleles will increase in the population over time due to Natural Selection.

An important distinction should be made between absolute and relative fitness. Absolute fitness indicates the absolute increase or decrease in the abundance of a genotype, while relative fitness is the increase or decrease relative to another genotype. Thus, absolute fitness does not reflect the same information as relative fitness. For instance, in case of a drought, one plant might be more resistant than another neighbour reference plant. The resistant plant would thus have a high relative fitness. However, its absolute fitness would probably decrease. Relative fitness is informative in terms of evolution as it directly indicates the change in genotype frequency due to Natural Selection. The selection coefficient is another metric that takes the value of one minus the relative fitness.

Fitness of yeast cultures can be measured using different techniques with different characteristics as shown in Table INT1. Plating and Colony Forming Units (CFU) counting consists in spreading a liquid culture on Petri dishes at a low-enough concentration so that each yeast is on a distinct location. After

few days, colonies appear and their number indicates the concentration of yeast in the population. Optical density consists in determining how much a liquid culture of yeast absorbs light. This value is then compared to a range of reference values to determine the concentration of yeasts in the culture. For flow-cytometry, cells are first fixed and then aspired by a flow-cytometer. This machine contains small pipes where yeast cells are individually transported, subjected to different lasers and can be sorted. This provides information about the size, granularity and fluorescence (several colors can often be measured) for each cell. Microfluidic designates a set of machines that also contain cell sized small pipes were yeasts can be sorted or tracked over time. Those systems are often coupled with automated high-throughput microscopes, and image processing software. This allows to culture yeasts in very precise and dynamic conditions, and to monitor their growth rate and their fluorescence activity over time.

Technique	Throughput	Precision	Time to get fitness values	Total cost
Plating & CFU counting	Low (IA)	Very High	Few days	low
Optical density	Medium (IA)	Low	Immediate	low
Flow-Cytometry	Medium (IA)	High	Few days	medium
Microarrays	High (PA)	Medium/High	Few weeks/months	high
BarSeq	Very high (PA)	Medium/High	Few months	high
Microfluidics	Low (IA)	Very high	Few weeks	medium

Table INT1. Advantages and Drawbacks of techniques to measure yeast fitness. Abbreviations: PA = Pooled Assay; IA: individual

A common approach in yeast biology to determine the function of a gene is to measure the fitness of a loss of function mutant strain in various conditions. A change in fitness indicates that the gene has a function important in this condition. In 1996, an international effort was launched to systematically disrupt all yeast genes. This resulted in the generation of 4 Yeast Deletion Libraries of more than 4000 viable strains each: a homozygous, a heterozygous (only one copy of the gene was deleted), the MAT a and MAT alpha libraries (Giaever and Nislow, 2014; Giaever et al., 2002a). They were constructed in the s288c background in order to be consistent with the genome sequencing data. These Yeast Deletion Libraries became widely used in the last decades due to their simplicity of use and great power to assess systematically most genes of the yeast genome. A two-steps PCR experimental strategy was used for constructing the strains. For each strain, one ORF was completely removed and replaced by a cassette of resistance to Kanamycin (KanMX), flanked by two barcodes (the uptag and the downtag). Each barcode is unique in the library and can be amplified using a common set of primers (U1-U2 for the uptag, D1-D2 for the downtag). Most pooled assays using a YDL follow the same pipeline: pooled growth of a library in specific conditions, extraction of genomic DNA, PCR amplification of either the uptag, downtag or both, and quantification of barcodes through either Oligonucleotide hybridization of barcodes (microarrays) or Barcode Sequencing (BarSeq).

Historically, microarrays with oligonucleotides complementary to the strains' barcodes were synthesized. Hybridizations of the barcodes to the oligonucleotides allow to quantify the abundance of each strain in a population. Pooling strains in this type of experiments decreases greatly the costs and time required to do experiments (per strain). In addition, it adds consistency to experiments as all

strains are subjected to the same condition. BarSeq was developed later and take advantage of the decreasing cost of sequencing to make the technique more and more accessible (Smith et al., 2009). In addition, it is possible to further increase multiplexing (number of conditions tested per experiment) by incorporating another barcode during the PCR amplification step (Smith et al., 2010). This second barcode is present in one of the two primers and indicates the population of origin of the mutant. Then, amplicons from different populations can be pooled and sent together to sequencing. Both the mutant and the population barcodes can then be read in the DNA sequences.

However, there are also some drawbacks that one should be aware of when working with such deletion libraries. First, and most importantly, there can be interactions between strains in a pool. This can include physical interactions between strains (i.e. biofilm production, cell to cell contact), production of toxins, or exchange of metabolites. Another issue is the presence of secondary mutations or aneuploidies in strains of the pool, which can partially or completely change the phenotype of the mutants. One study found that most of the haploid deletion mutants acquired at least one secondary mutation (Teng et al., 2013). Interestingly, when deleting a gene in replicate wild-type strains, the acquired secondary mutation was often the same gene or a gene with a related function, with homologs of human tumour suppressor genes often targeted. In another study, over 187 haploid yeast null mutants, with a fitness defect in rich liquid medium (YPD), were evolved during 400 generations in YPD, with four replicates each (Szamecz et al., 2014). The authors found that 68% of lineages acquired secondary mutations that reduced their fitness defect. Secondary mutations were acquired quickly and could compensate defects in a broad range of cellular processes. In particular, mutants with the strongest fitness defects were more likely to acquire secondary mutations. Four replicates were evolved per mutant. Replicate lines showed a lack of convergent evolution for both secondary mutations and phenotypes (i.e. growth) in various environments. Thus, it appears that secondary mutations are common in mutants of the Yeast Deletion Libraries, especially when strong fitness defects are observed.

Due to those issues, complementation assays and individual strain cultivation are needed when one wants to definitely confirm the role of a mutation in a given condition. A complementation assay consists in inserting, in a mutant, a wild-type copy of the gene that had previously been deleted, at another unrelated locus. If the fitness effect disappears, then it confirms the role of the gene and not of potential aneuploidies, secondary mutations or perturbation of the regulation of other genes in cis of the gene deletion. Another issue with the deletion library is the presence of hundreds of “sick” mutants, that are growing so slowly that they are in practice impossible to study (Giaever and Nislow, 2014).

In general, the homozygous library is considered of better quality than the haploids and heterozygous library. For instance, the haploid library has about 8% of strains with aneuploidies while the homozygous library has less than 0.3% of tetraploids (Giaever and Nislow, 2014). However, due to the simplicity of genetic analysis, most studies were performed in the haploid deletion. The heterozygous library has lower fitness effects than the homozygous library (Manna et al., 2012). Thus, differences between conditions are less marked, and more difficult to detect. Additionally, (and taking this effect into account), the heterozygous library has much lower replicability than the homozygous library. For instance, some replicates of the same conditions have anti-correlated fitness values (Manna et al., 2012). Finally, the genetic background of those deletion libraries contains several auxotrophic selective markers, that were included to improve genetic manipulations. However, most studies using the yeast deletion library are interested in traits related to metabolism (such as growth on different nutrients, aging, cell cycle, growth/metabolism…). Those studies should preferential be performed with

prototroph strains in order to be able to catch subtle metabolic effects. For those reasons, different prototrophic libraries were constructed recently (Gibney et al., 2013; Mulleder et al., 2012).

The construction of those YDL was a huge advance in genetics and allowed to address several fundamental questions, such as: how many genes are essential? Can we find a function to all genes? How many genes show a dosage-dependent effect (copy number) or a dominant effect on fitness? Can we identify new genes in well-characterized pathways using systematic approaches? For a review on the yeast deletion libraries, please refer to (Giaever and Nislow, 2014). Briefly, it was initially found that most genes have no growth phenotype in standard laboratory conditions, which was unexpected. Later, studies of "chemogenomics" exposed the YDL to banks of thousands of small molecules (Hillenmeyer et al., 2008; Hoepfner et al., 2014; Lee et al., 2014), and revealed that when the number of tested conditions increase, almost all mutants show a growth phenotype.

An issue with using the homozygous library for screening banks of drugs is that most targets of drugs are essential genes. For this reason, chemogenomics assays usually screen the heterozygous library, in an experiment called HaploInsufficiency Profiling (HIP) (Giaever and Nislow, 2014). This allows to identify genes that are direct targets of the drug. HIP also have the advantage of testing the specificity of drugs to a gene target, and evaluate the number of potential off-targets. A library called DAmP (Decreased Abundance by mRNA Perturbation) in which the expression level is decreased to 10% of the wild-type was constructed to increase the range of expression level tested (Schuldiner et al., 2005). In addition, to HIP, chemogenomics assays perform HOmozygous Profiling (HOP). This allows to identify genes that buffer the effect of the drug. This assay is particularly useful when the drug has no direct gene target (i.e. genes involved in multi-drug resistance). Thus, both assays are complementary and are conveniently named HIP-HOP.

An interesting output of chemogenomics assays is the generation of co-fitness data. Those are simply produced by the correlation of the matrix of fitness of all strains and all drugs. Then, one can rank strains (resp. drugs) that are the more similar to another strain (resp. drug). Since there were thousands of strains (resp. drugs) tested, the most similar strains (resp. drugs) often share the same functions (resp. targets) or are involved in the same pathway. Even though this approach is extremely powerful to determine the potential function of a gene, in some cases, it occurs that certain strains are highly "co-fit" because they share the same secondary mutations.

Synthetic Genetic Array (SGA) is a procedure initially developed in 2001 that allows to create huge libraries of double mutants (Tong et al., 2001). It consists in systematically crossing a query gene deletion mutant with a whole Yeast Deletion Library (or a part of it). This is achieved thanks to an automated protocol where crossing, sporulation, selection of double mutant haploids and fitness measurements are performed in an automated way, using robotic platforms. Then, an epistatic interaction is detected if fitness is higher (genetic enhancement) or lower (synthetic lethality) than average fitness of both individual mutants. An epistatic interaction can indicate that two genes are part of the same or different pathways, or have similar functions (as for chemogenomics assays). More than 23 million double mutants were constructed recently (Costanzo et al., 2016). The authors detected ~550,000 negative and ~350,000 positive Genetic Interactions (GI). This study showed that essential genes were enriched in GI. It also showed that negative GI are predictive of shared genes functions, while positive GI indicated general regulatory connection among gene pairs. Most well-known cellular pathways and their interactions could be reconstituted using this approach.

C Yeast response to osmotic stress

Osmosis designates a phenomenon whereby when a solvent (usually water) is separated by a semipermeable membrane, it starts diffusing from the most concentrated side in osmoles to the least concentrated side. The osmole is a unit of measure similar to molarity, that indicates how many molecules are present in water. For cells in suspension, when there is a lot of osmoles in the surrounding medium they start losing water by osmosis, shrink and eventually die. In order to prevent such hyperosmotic stress, cells produce compatible osmolytes: compounds that do not inhibit cellular processes and can increase the intracellular osmolarity. Hypo-osmotic stress is the opposite phenomenon: when the extracellular milieu is less concentrated in osmolytes than the intracellular milieu. This induces an increase in cell volume (swelling), and eventually apoptosis.

Yeasts can grow up to glucose 40%, thus they have high resistance to hyperosmotic stress (Saito and Posas, 2012). They can produce different osmolytes such as trehalose, amino acids, ions, but the most effective one to mediate osmo-resistance is glycerol. Yeast response to osmotic stress starts by an arrest of the cell cycle, an adjustment of transcription and translation and an increased production of glycerol. Common osmotic stress used in yeast experiments include sodium chloride (NaCl), potassium chloride (KCl), or the sugar alcohol sorbitol. Ionic solutions induce an additional ionic stress that includes different detoxifying mechanisms.

In response to osmotic stress, yeasts trigger the High Osmolarity Glycerol (HOG) pathway. This pathway has been intensively studied, partly due to its high conservation in other species, including humans. For instance, replacing the core gene of the pathway in yeast, the HOG1 Mitogen-Activated Protein Kinase (MAPK), by the human homolog p38 MAPK rescues the sensitivity to osmotic stress (Saito and Posas, 2012). MAPK pathways are very conserved across eukaryotes, and regulate major decisions such as mitosis, apoptosis, survival or differentiation. MAPKs pathways share the same structure: a sensor(s) at the cell's plasma membrane and a cascade of 3 kinases (MAPKKK, MAPKK, MAPK) that are activated one after the other, until the last kinase, the MAPK, gets activated and regulates the expression of hundreds of genes. In addition, some MAPK pathways share components, which result in crosstalk and interactions between pathways.

There are two branches in the Hog pathway that are named by their osmosensor name: Sho1 and Sln1. Both branches result in the activation of the MAPKK PBS2, which activates the MAPK HOG1. Once activated some Hog1p proteins have cytosolic targets, and some move to the nucleus and activate transcription factors which induce the expression of hundreds of genes. As cells become adapted, negative feedback mechanisms deactivate Hog1p and export it from the nucleus. Upon mild-step osmotic stress, Hog1p gets activated in about 5 minutes and then gradually deactivates until basal level in about 30 minutes. Negative feedback mechanisms include glycerol accumulation, protein phosphatase activity, phosphorylation of elements upstream of Hog1p, or crosstalk with other MAPK pathways (Saito and Posas, 2012).

Despite being functionally similar, the Sho1 and Snl1 branches have different properties. The Sln1 branch is considered as more critical for survival to very high or low osmolarities (O’Rourke et al.,

2002). In addition, only the Sln1 branch is basally active in standard conditions (Macia et al., 2009). A basal activity could be a general property of MAPKs that allows a faster reaction to stimuli (i.e. there is no time needed to "initiate" the pathway). Accordingly, a microfluidic experiment where cells were stimulated with periodic osmotic stress showed that the Sln1 branch more faithfully follows changes in osmotic stress concentrations than the Sho1 branch, and thus has a shorter reaction time (Hersen et al., 2008).

II Adaptation to a single environmental change

A Genetic adaptation and phenotypic plasticity

When faced with new environmental conditions, individuals can adopt one of the following strategy: dispersal to follow their favoured environment or evolutionary adaptations (Berg et al., 2010; Chevin et al., 2010; Kokko and López-Sepulcre, 2006). Dispersal plays contrasting roles on the evolutionary potential of a species: on the one hand, it allows the propagation of beneficial alleles in different niches which also increases genetic diversity and the rate of evolution. On the other hand, it reduces the fitness of individuals that were well-adapted to their previous niche (Schiffers et al., 2013). Also in many situations populations are not able to track their favoured environmental niche, due to their limited dispersal abilities, or to the scarcity of this niche. In these situations, evolutionary adaptation, through genetic adaptation and/or phenotypic plasticity, is needed. Interestingly, evolved traits could be dispersal abilities, evolvability or plasticity themselves.

Phenotypic plasticity refers to the ability of a given genotype to produce different phenotypes in function of the environment in which it develops/grows. Organisms that rely poorly and heavily on phenotypic plasticity are respectively called specialists and generalists (see part IIIAi). There has been much work on determining the costs and limits of plasticity (DeWitt et al., 1998; Murren et al., 2015). The cost of plasticity usually refers to the cost on fitness of maintaining a plastic genotype. That is, the fitness difference for the same value of a trait between a specialist and a generalist. However, since few or no significant costs of plasticity have been detected despite many studies, some authors suggested that the cost of being a specialist might outweigh the cost of plasticity in most cases (Murren et al., 2015). The limits of plasticity correspond to the maximal phenotypes that a species can produce. That is the difference in trait level between a specialist and a generalist (Auld et al., 2010; Murren et al., 2015). Phenotypic plasticity can be either adaptive or non-adaptive (maladaptive) depending on whether it provides a fitness advantage or not. A simple measure of phenotypic plasticity is the coefficient of variation of the phenotypes of one genotype tested in different environments (Valladares et al., 2006). The reaction norm is the phenotypes/traits of one genotype across different environments. Another measure of plasticity is the slope of a linear reaction norm (Chevin et al., 2010).

Physiological adaptation or adaptation through phenotypic plasticity refers to the organism non-genetic changes happening during an individual lifetime/generation in order to adjust to its current environment. Physiological adaptation occurs on very short timescales, (~generation time).

In contrast, genetic adaptation usually requires at least a hundred generations, as has been shown by experimental evolutions of microbes. Determining the time needed for genetic adaptation to occur is a fundamentally complex issue as it depends on both population and environmental factors that are dynamic across time. Main population factors are population size, standing genetic variation, target size (the fraction of the genome in which mutations will increase adaptation), mutation rate, evolvability or non-additive effects such as gene redundancy and epistasis. Environmental factors include selection pressure of the environment, temporal or spatial environmental heterogeneity, competition with other strains, type of environments (steady, periodic, stochastic, ramp increase...), speed of environmental changes, or level of fluctuating selection.

The current and historical consensus in evolutionary biology is that long-term adaptation is purely genetic, with the apparition and fixation of mutations as the key components. However, recently, the role of epigenetic mechanisms to long-term adaptations has been hotly debated (Laland et al., 2014). Indeed, it has been argued that plasticity and non-genetic transgenerational inheritance mechanisms (see part 1-III-B) are key players and not just by-products of evolution.

B Tracking the dynamics of adaptive mutations in microbes

Microbes have become the organisms of choice for the study of the dynamics of genetic adaptation. Indeed, their short generation times allow for parallel experimental evolution assays in reasonably short time-scales with a high degree of control over population genetic parameters. And their small genome size allows for Whole Genome Sequencing (WGS) in order to find the genetic determinants of adaptation. In addition, it is possible to "save" all time points of an experiment for further molecular analysis or even "resuscitation", through frozen fossil records. Performing WGS has become a routine in many laboratories in the last decade due to the decreasing costs of sequencing. New techniques of pooled sequencing were developed, such as Bulk Segregant Analysis (BSA) (Duveau et al., 2014). In a BSA assay, individuals in a population are sorted, and individuals with extremely high or low trait value are separated into two groups. Pooled sequencing of the DNA of those two groups can result in a differential enrichment between the two groups, of genes important for the sorted trait.

Experimental evolution experiments revealed valuable results about genetic evolution (reviewed in (Lang and Desai, 2014)). For instance, few examples of epistasis were found among beneficial mutations. Accordingly, mutations in the same pathways often produce the same effect on fitness. Importantly, phenotypic convergence (parallel evolution) is common while genotypic convergence is rare. This illustrates that there are multiple possible genetic roads to reach a given phenotype. In fact, considering pathways instead of single genes can largely improve analysis of WGS data: it can improve the rate of parallel evolution from 2% to 30% (Lang and Desai, 2014). Most common mutations in haploid microbes are loss-of-function mutation, as expected as many single point mutation can cause them, followed by gene duplication and sometimes gene translocation. The concept of diminishing return epistasis indicates a negative epistasis phenomenon whereby the beneficial effect of beneficial mutations is smaller in the presence of other beneficial mutations (Chou et al., 2011; Martin et al., 2007). This effect could be stronger for large effect beneficial mutations. Finally, phenotypic evolution can be predictable on short time scales, but not on long time scales where epistasis is important and

chance mutations will start combining to produce unexpected phenotypes (Lang and Desai, 2014).

In a landmark study, Sasha Levy *et al.* developed a protocol that allows to tag ~500,000 haploid yeast cells by transforming a plasmid library containing millions of unique barcodes in a yeast population (Levy et al., 2015). Using barcode sequencing, they were able to track the abundance of each cell lineage across time, and thus to measure the apparition of new mutations and their fate: disappearance or fixation. They found that initial small effects mutations are deterministic, while rare large effect mutations become fixed and generate variability between replicates. However, they explain in a follow up study, that introduction of barcodes induced diploidization of most cells (~80% of them) (Venkataram et al., 2016). This technique, once perfected, could allow even more powerful yeast genetics. For instance, a later potential application could be to construct a yeast library with a much higher genome saturation, for more comprehensive genomic analysis. This might be done by barcoding a huge isogenic population, and subjecting it to DNA-damaging agents at small concentrations. This might result in a barcoded library that would contain (probabilistically speaking), for any single point mutation, one strain with only this mutation.

III Adaptation to fluctuating environments

No organism lives in a constant environment. Thus, it is important to consider how adaptation occurs in fluctuating conditions, even though it is usually more challenging to study experimentally. In this section we will first describe the different strategies of adaptation to fluctuating environments, and then the mechanisms of anticipation and memory (priming) that can be advantageous in fluctuating conditions.

A Strategies of adaptation to fluctuating conditions

i Main strategies of adaptation to fluctuating environment

When considering adaptation to “real” fluctuating conditions, several (sometimes related) strategies have been described, such as: generalism, specialism, priming, sensing, bet-hedging, plasticity, stochastic switching, stochastic sensing, phase variation. Many theoreticians are developing models to determine the optimal strategy in various environmental contexts. However, few experimental studies have confirmed those predictions due to inherent difficulties in studying fluctuating environments. Importantly, there is no doubt that organisms in the wild are not restricted to a specific strategy but combine different strategies to some degree. However, understanding which conditions favour which strategies is a key issue in evolutionary biology.

Generalists designate “all-rounder” genotypes that are well fit in many different environments. In contrast, **specialists** are genotypes that are highly fit in one or few environments but poorly fit in other environments. **Plasticity** is the ability of a genotype to develop different phenotypes according to the environment. Plasticity is often related to generalism. And it is often assumed that plasticity has a

fitness cost that has to be paid to keep the ability to produce different phenotypes. **Priming** (or deterministic/anticipatory maternal effects or transgenerational phenotypic plasticity) is a strategy where organisms react to a first stimulus in anticipation of the later apparition of another related environment (see next 1IIIB). **Sensing** designates organisms that rely on their sensors to detect that the environment changed and to adapt their physiology. This strategy is well characterized molecularly as it corresponds to plasma membrane receptors that sense and transmit signals to MAPKs or other cell signalling proteins. Sensing is usually opposed to bet-hedging (or randomizing maternal effects): a strategy that aims to reduce the risks of extinction during extreme environmental conditions (see below).

ii Bet-hedging

A Different types of bet-hedging

Bet-hedging is defined as a decrease in both arithmetic mean fitness (across environments) and its variance and an increase in geometric mean fitness (Philippi and Seger, 1989; Sæther and Engen, 2015; Seger and Brockman, 1987). This concept was first developed by Dan Cohen in 1966 (Cohen, 1966). The rationale behind lies in the observation that fitness in fluctuating (real) environments is more accurately estimated by geometric than by arithmetic mean fitness. Indeed, extreme environments can wipe off entire populations, and have more weights on geometric mean fitness (since it is multiplicative) than on arithmetic mean fitness.

Different forms of bet-hedging have been described: conservative bet-hedging, diversifying bet-hedging, and adaptive coin flipping. **Conservative bet-hedging** indicates a strategy where a phenotype is produced that is neither optimal nor detrimental in any environment. A common example to illustrate bet-hedging strategies is the size of eggs at birth (Olofsson et al., 2009a; Philippi and Seger, 1989). Small eggs have less chances to survive to harsh conditions, and big eggs require more resources to produce. In this case a conservative bet-hedging strategy would consists in producing eggs of intermediate size at each clutch. **Diversifying bet-hedging**, or within generation bet-hedging, is a strategy where a single genotype produces different phenotypes. This allows to have, in an isogenic population (e.g. for microbes), individuals that can resist to different types of stressful environments. This way, risks are spread among the population. This is the same principle as the one used in economics and finance, resumed by the adage: “don't put all your eggs in the same basket”. Coming back to the bird example, a diversifying bet-hedging strategy would consist in producing eggs of various sizes at each clutch.

Stochastic switching, or phase variation in bacteria, is a type of diversified bet-hedging where individuals stochastically switch their phenotypes at rates much higher than mutation rates (Hallet, 2001; Salathé et al., 2009; Salaun et al., 2003; van der Woude, 2011). Stochastic sensing indicates a combination of sensing with stochastic switching (New et al., 2014). In **adaptive coin flipping**, or between generation bet-hedging, the strategy is randomized at each generation (i.e. production of progeny; for instance once a year for plants, or once at every clutch for birds) (Cooper S. and Kaplan H., 1982; Hopper et al., 2003). Thus, there is a high between-generation heterogeneity but a low

within-generation heterogeneity. For example, an adaptive coin flipping strategy would consists in producing eggs of different size at each clutch, with all eggs having the same size within a clutch.

Bet-hedging is a difficult concept to prove experimentally since it requires showing that the bet-hedging phenotype induces a better fitness in dynamic environments (J. Ripa et al., 2009; Olofsson et al., 2009a; Rees et al., 2010). To clarify research in this field, in 2011 Pr Andrew Simons made a comprehensive review in which he classified over 100 studies across 16 phyla where evidence of bet-hedging was claimed. He observed that most studies focused on diversifying bet-hedging strategies and few on conservative bet-hedging (he did not mention adaptive coin-flipping). Candidate bet-hedging traits were classified in six groups of increasing and cumulative evidences. In order to classify a trait in the two highest groups, a proof should be made that the trait is adaptive under fluctuating conditions. Results of his classification showed that most studies fell into the weakest categories of evidence (II and III) and just 12 studies were in the two highest categories. Therefore, despite a high number of studies claiming evidence of bet-hedging, few have shown that variability in the trait is adaptive. This study underlines that revealing heterogeneity in the values of a trait in an isoclonal population is not sufficient for classifying it as "purely" a bet-hedging. This heterogeneity should also be adaptive (i.e. increase fitness) in fluctuating conditions.

B Examples

Among the cases of diversifying bet-hedging are many examples of dormancy of seeds (Childs et al., 2010; Simons, 2009; Venable, 2007), insects (i.e. diapause) (Danforth, 1999; French et al., 2014), fungus (Graham et al., 2014), bacteria (i.e. persistence) (Balaban et al., 2004; Kussell and Leibler, 2005; Kussell et al., 2005; Ratcliff and Denison, 2010; Zhang and Rainey, 2010) or variabilities in egg clutches (A. Sarhan and H. Kokko, 2007; J. Ripa et al., 2009; K. Thumm and M. Mahony, 2002; Olofsson et al., 2009b; Rees et al., 2010). In all those cases heterogeneities in dormancy durations or in egg size/types allow to face unpredictable and harsh conditions (such as a drought, invasion by predators...). For instance, a landmark study by Balaban et al used a microfluidic device to show that persistence of bacteria to antibiotics could be due to heterogeneities in the population, with a sub-population of slow growing highly resistant cells (Balaban et al., 2004). Cells were switching between two phenotypes (high growth and low resistance versus the opposite) stochastically. This example showed that random phase variations have important biomedical consequences.

Another type of bio-medical consequence of diversifying bet-hedging is illustrated by two studies on cancer cells (Dannenberg and Berns, 2010; Roesch et al., 2010; Sharma et al., 2010). In 2010, Roesch et al identified a subpopulation of slow growing cells in a melanoma (Roesch et al., 2010). These cells can be distinguished since they are JARID1B (a H3K4 demethylase) positive. After knockdown of JARID1B, the population starts growing very fast and soon become exhausted. This shows the important role of this sub-population in tumor maintenance. Moreover, those slow-growing cells can switch to fast-growing cells (JARID1B negative) and *vice versa*. In a related paper Sharma et al analysed various tumor cell lines and consistently identified a subpopulation of cells that showed an increased resistance to drugs of more than 100 times (Sharma et al., 2010). As in Roesch et al, they found that the resistant cells were rare, transient, and reliant on JARID1A (a paralog of JARID1B (KDM5 family)), but also on IGF-1 receptor signalling.

Phenomenons related to bet-hedging have also been observed in budding yeasts. Most of those experiments focused on the resistance to different stresses, such as metal stresses in the laboratory of Pr Simon V. Avery (Bishop et al., 2007; Holland et al., 2014; Howlett and Avery, 1999; Sumner and Avery, 2002), heat stress (Attfield et al., 2001; Levy et al., 2012), nutrient limitations (Breker et al., 2013; Petrenko et al., 2013; Vardi et al., 2013). In 2012, Levy et al set up a single-cell automated microscopy assay to study yeasts growth heterogeneities. Trehalose is a sugar known to be important for resistance to several types of stresses. The authors found that the abundance of Tsl1p, a trehalose-synthesis regulator, is negatively correlated with growth rate and positively correlated with longevity and heat shock survival. Their results are qualitatively similar to those in the bacteria and cancer studies cited above, with a sub-population of slow-growing and highly stress resistant cells (however they cannot switch stochastically to a fast-growing stress sensitive phenotype in this case, as trehalose accumulation is correlated with age). Thus, it appears that diversified bet-hedging is present in many single-cell species. And heterogeneity in the population can be maintained by a simple physiological parameter such as cell age. However, within a population there does not seem to be many different cell types each pre-adapted to a different environment, but just two cell types with one being rare, able to resist to many different types of stresses, and with a slow growth phenotype (Geisel et al., 2011).

In order to be classified as applying a diversifying bet-hedging strategy, a trait should satisfy several conditions. Mainly it should be both heterogeneous in the population, and this heterogeneity should be adaptive. There have been several reviews trying to determine how such “noise” could be adaptive (Richard and Yvert, 2014; Veening et al., 2008; Viney and Reece, 2013). One type of traits that could be particularly subject to such bet-hedging strategies is plasma membrane proteins levels. This include sensors of different stresses, and protein transporters of various nutrients. In order to face unpredictable conditions, it could be advantageous for genotypes to produce cells with various levels of these proteins. One computational study supports this idea (Zhang et al., 2009). They showed that plasma membrane proteins have higher gene expression variability than other gene categories. Using simulations, they compared two genotypes with similar and suboptimal mean gene expression levels but different noises. They found that noise could be adaptive in unpredictable conditions, as predicted by theory. Related to that, an experimental study found that two plasma membrane proteins were trans-regulators of the expression variability of another gene (Fehrmann et al., 2013).

C Mechanisms

An interesting question to better characterize bet-hedging strategies is “How are heterogeneities in different traits generated?”. There are many sources of heterogeneities or noise. Some are genetics, other epigenetics and other physiologicals. It is often assumed that the main sources of noise in a population are physiological factors, such as: cell cycle stage, ultradian rhythms, growth rate or cell age (Avery, 2006; Sumner and Avery, 2002). For this reason, studies trying to find the genetic sources of noise usually include in their design a way to control for those factors. Epigenetic mechanisms that can generate noise include: prions (Alberti et al., 2009; Halfmann et al., 2010; Newby and Lindquist, 2013), alternative histone variant (Richard and Yvert, 2014), DNA methylation or gene network structures (Satory et al., 2011). Genetic sources of noise include protein copy number, retrotransposons

(Specchia et al., 2010), TATA Box and transcriptional bursts (López-Maury et al., 2008; Richard and Yvert, 2014). A low protein copy number increases cellular noise simply because of reduced chances of interactions with other molecules (Niepel et al., 2009). Accordingly, there is a known negative correlation between the number of proteins/genes and their variability (Newman et al., 2006).

Noise or variability in gene expression can be either extrinsic or intrinsic. Intrinsic noise indicates noise that originates from cis regulations of the gene. While extrinsic noise indicates variability that has sources that are extrinsic to cis-regulation (in trans), such as physiological parameters. An elegant experiment allowed to distinguish these two types of noise: introducing, at two distant loci in the genome, two different fluorescent proteins under the same promoter (Elowitz et al., 2002; Raser, 2004; Rinott et al., 2011). Then, on a Figure where each axis represents the level of one fluorescent protein, the spreading of measurements on the identity (resp. orthogonal) axis indicates the amount of extrinsic (resp. intrinsic) noise.

iii Types of environmental changes determine the optimal strategies of adaptation

Predictable environments include ramp increases, periodic environments, or more generally highly autocorrelated environments. **Ramp increase** is a gradual and continuous change from one environment to another. A prominent ecological example is climate change, that impacts most species. **Periodic environments** are also widespread, with: diurnal fluctuations, circadian cycles, ultradian rhythms, and seasons. **Autocorrelation** indicates the correlation of a signal with a delayed copy of itself, as a function of delay. Spatial autocorrelation can be illustrated by a storm that affect a whole region. Temporal autocorrelation can be illustrated by a series of bad years, or drought. Thus, a highly (resp. poorly) autocorrelated environment is predictable (resp. unpredictable) (Hallsson and Björklund, 2012). Plasticity is predicted to be favoured in environments that are both predictable and heterogeneous (Dey et al., 2016; Reed et al., 2010). Unpredictable, poor/deteriorating or heterogeneous environments favour bet-hedging or generalist genotypes (Dey et al., 2016; Tuljapurkar and Istock, 1993). These environments correspond to sudden events that could not be anticipated: such as apparition of a physical or chemical lethal stress (i.e. a predator), disappearance of a food source, or more generally extreme and poorly autocorrelated environments.

There is a substantial body of theoretical work on trying to determine the optimal strategy of adaptation according to the predictability of various environments (Chevin and Lande, 2015; Donaldson-matasci et al., 2008; Ezard et al., 2014; Geisel et al., 2011; Lande, 2009; Lof et al., 2012; Reed et al., 2010; Salathé et al., 2009; Svardal et al., 2015; Tuljapurkar and Istock, 1993; Yamamichi et al., 2011). For instance, a recent study mathematically confirmed the theory that unpredictable and heterogeneous environments are favouring a bet-hedging strategy, where the heterogeneous trait is plasticity itself (Frankenhuis et al., 2016). The authors also predicted that this “differential plasticity” can only emerge if the cost of being mismatched to the environment exceeds the benefits of being well matched. More generally, many studies try to determine which parameters (in addition to predictability) influence the choice of the optimal strategy between stochastic switching (or bet-hedging) and sensing. For instance, studied parameters include: cellular parameters (switching rate and sensor precision in (Wolf et al., 2005)), population parameters (composition and size in (Arnoldini et al., 2012)), environmental

parameters (fitness landscapes and selection pressure in (Kobayashi and Sughiyama, 2015; Salathé et al., 2009), range of changes in (Donaldson-matasci et al., 2008), rate of changes in (Kussell and Leibler, 2005), autocorrelation and noise in (Wolf et al., 2005)). Finally, fluctuating selection in the wild is assumed to generate diversity and to decrease the strength of natural selection (Bell, 2010; Simons, 2009).

There are several beautiful experimental studies on the strategies of adaptation in fluctuating environments (Dey et al., 2016; Hallsson and Björklund, 2012; Ketola et al., 2013; Kvitek and Sherlock, 2013; New et al., 2014; Venail et al., 2011). As often, experimental data do not always confirm theories. For instance, in 2016 Dey et al evolved a *C. elegans* strain in fluctuating normoxia-anoxia conditions. As predicted by theory, they observed the evolution of anticipation of glycogen stocks in periodic conditions. However, they did not observe the evolution of bet-hedging in unpredictable conditions of anoxia-normoxia fluctuations. Lastly, Kvitek and Sherlock did pooled sequencing on yeast populations evolving in constant environments (Kvitek and Sherlock, 2013). They observed that in such environments, evolution of a specialist strategy is reproducibly traded against a lower level of sensing. Indeed, most mutations occurred among three major signalling networks in yeasts that govern growth: glucose signaling, cAMP-PKA and HOG.

B Mechanisms of transcriptional anticipation and memory of environmental changes

Triggering a transcriptional stress response is a very costly cellular process, that involves the coordinated expression of hundreds of genes. However, in nature, fluctuating conditions are the norm rather than the exception. Environmental fluctuations can be either stochastic, periodic, autocorrelated or sequentially correlated. For instance, when the rain stops during a rainy day, there are high chances that it will start raining again later in the day. Launching a stress response every time the rain strikes again would be an energy sink for cells. Since living organisms are smart, they developed efficient strategies to face such fluctuating conditions: to anticipate that rain could occur again, or to keep in memory that it occurred. Importantly, not all environments are independent from one another: one environment/stimulus can increase the probability of occurrence of other types of environments. For instance, if rain keeps preys hidden, then their predators will be forced to diet until weather improves. Thus, organisms interpret environmental cues as changing the probabilities of occurrence of potential future environments. Those probabilities are specific to each species, as it is linked to their ecology and evolutionary history.

Specifically, the term **priming** has been proposed to describe a situation where an organism exposed to a first stress (priming stimulus) will launch epigenetic 'memory' mechanisms that will increase its protection to a later second stress (triggering stimulus), as compared to a naive organism, (Hilker et al., 2016). The **memory time** or **memory phase** refers to the time delay between the two stresses during which the cell “remembers” the priming stress. Cis-priming and trans-priming indicate that the first stress is respectively similar or different to the second stress. Transgenerational memory can be achieved through transposon mobility, or DNA rearrangement (Hilker et al., 2016). However, this is not considered as priming since the memory effect was not caused by regulated epigenetic modifications.

Transcription is a highly dynamic process by which cells respond quickly to different stress. Transcriptional responses are increasingly low-cost and simple to analyze thanks to decreasing costs of sequencing and improved bioinformatic methods. For these reasons, transcriptome profiles in response to single stresses have been generated. During the last decade, researchers have started to unravel the molecular determinants of transcriptional priming. In the following parts, we will describe two different aspects of priming through transcription: expression of stress response genes after a priming stimulus in anticipation of a possible triggering stimulus, and the molecular mechanisms that allow cells to remember a first priming stimulus and launch a faster transcriptional response during a second triggering stimulus.

i Transcription during a first stress in anticipation of a second stress

A The Environmental Stress Response and transcriptional anticipation

Once subjected to a stress, *Saccharomyces cerevisiae* cells regulate the expression of hundreds of genes. Some of those genes are condition-specific genes that are important for growth in the current stressful condition. However, most genes are not condition-specific: any stress induces the upregulation of ~300 genes and the downregulation of ~600 genes (Gasch and Werner-Washburne, 2002; Gasch et al., 2000). Upregulated genes are involved in stress responses, while downregulated genes encode for ribosomal protein or ribosome biogenesis. Thus, there seems to be a cellular trade-off between stress response and growth. This common stress response has been termed the General Stress Response (GSR) or the Environmental Stress Response (ESR).

It is partly regulated by the paralog Transcription Factors (TF) Msn2, Msn4 and Yap1 (Berry and Gasch, 2008; Gasch et al., 2000). Indeed, nascent protein synthesis is needed to survive a severe secondary stress (~20 to 60 minutes after the first stress) but not a mild primary stress (Berry and Gasch, 2008). Accordingly, mutants for msn2 or msn4 show no defect in primary mild stress resistance but a strong defect in secondary severe stress resistance, which is consistent with their reduced gene expression response to primary stresses (Berry and Gasch, 2008). A later study showed that a double mutant msn2 mns4 has no defect in severe secondary stresses (Zakrzewska et al., 2011). Thus, the influence of msn2 msn4 on the ESR is condition-specific, and other regulatory factors are likely involved. In addition, the large set of genes expressed during a stress is quite different from genes needed to survive it (Giaever et al., 2002b; Gibney et al., 2013; Zakrzewska et al., 2011). Those results show that most regulated ESR genes are not needed for surviving the initial priming stress, but to prepare the cell to potential future stresses. Interestingly, a similar ESR also exists in *Escherichia coli* and in *Schizosaccharomyces pombe* but it is reduced/absent in *Candida albicans* or in cells from multicellular organisms (Battesti et al., 2011; López-Maury et al., 2008). The latter cells, within living hosts, are less directly exposed to environmental variations, which could explain why they show this reduced gene expression response to stress.

Importantly, a recent study showed that the common signature of ESR genes could in fact, only reflect the distribution of cells over different cell cycle phases (O'Duibhir et al., 2014). They found that slow growing strains – either defective mutants, or wild-type strains grown in non-optimal conditions – had a common transcriptional signature due to an increased number of cells in the G1 phase of the cell cycle. This study underlines that the trade-off between growth rate and stress resistance commonly observed in many organisms/strains, also occur within the cell cycle, with arrested-G1 cells being more resistant than dividing cells. Still, the stress resistance phenotype of G1-arrested cells represents a form of anticipation of future stresses, that is expressed at the molecular level by over-expression of hundreds of stress related genes.

Genes expressed during the ESR can be seen as a form of molecular memory and anticipation. Molecular memory since the transcripts and proteins of those genes have a given half-life and so they will be maintained for some time. Molecular anticipation since the expression of those genes will protect the cell against the same type of stress if it occurs (cis-priming), but also of different types of stress (trans-priming or cross protection). Trans-priming is not always symmetrical. This behaviour likely reflects the evolutionary history of each strain. If it has been used to encounter more frequently a given stress after another specific stress, it might evolve anticipation. Thus, studies on priming or anticipation, when applied on many conditions, could potentially allow to adopt a reverse-engineering approach to determine the ecology of micro-organisms. Two studies beautifully showed such anticipation features in micro-organisms (Mitchell et al., 2009; Tagkopoulos et al., 2008). Those studies consisted in measuring fitness and gene expression profiles of microbes in environments that mimicked the sequential changes of their habitats.

In 2008, Tagkopoulos *et al.* mimicked *E. coli* transition from the outside environment to the oral cavity, where the environment switch from less than 30 degrees to 37 degrees, followed by entry in the gastrointestinal tract, where the oxygen drops from ~21% O_2 to ~0% O_2 (Fig INT2a). Genes expressed during an increase in temperature (entry in oral cavity) or an increase in O_2 (exit from host) were highly correlated/overlapping with genes expressed in the next phase of the cycle (respectively decrease in O_2 or decrease in temperature). They then evolved *E. coli* cells for 100 hours in an environment that oscillated in the opposite direction than they are used to in their ecology: high temperature (37 degrees), accompanied 40 minutes after by high oxygen (21% O_2) for a certain duration (randomized to avoid periodic selection), then a switch to low temperature (25 degrees) accompanied 40 minutes after by low oxygen (0% O_2) during a certain duration, before the next cycle. Comparing evolved to parental strains, they observed that genes expressed in anticipation of future environments were largely reduced in the evolved strains. This showed that disappearance of anticipation can occur in less than a hundred generations.

In 2009, Mitchell *et al.* mimicked the wine production process (growth in grape must medium, followed by fermentation and respiration) for *S. cerevisiae*, and the passage in the mammal intestine (lactose, followed by maltose) for *E. coli* (Mitchell et al., 2009) (Fig INT2b). Results showed that priming provided a strong fitness advantage as long as the priming stimulus was appearing before the triggering stimulus in the ecology of the organism. Studying in more detail the trans-priming of heat shock (during fermentation) followed by oxidative stress (during respiration) in yeast they observed a set of 300 genes that was highly induced by heat shock and later by oxidative stress. This trans-priming effect was not observed when testing a different priming stimulus (osmotic stress) or a different triggering stimulus (YPD).

In 2013, Dhar *et al.* performed an experiment similar to Tagkopoulos *et al.*: they evolved yeast strains in an environment that oscillated between salt and oxidative stress (3 replicates) and in steady controls (salt, oxidative stress, no stress; 6 replicates each) (Dhar et al., 2013) (Fig INT2c). They observed an asymmetric trans-priming fitness gain where evolution in steady oxidative stress protected against salt stress (and oxidative stress) but not *vice-versa*. Importantly, in as few as 15 cycles (20 generations per cycle) they observed signs that gene expression anticipation evolved in the cyclic environment. This indicates that in addition to disappear fast (Tagkopoulos et al., 2008), anticipation can quickly evolve. However, they explain that it is very difficult to disentangle the effects of anticipation and trans-priming in a fluctuating environment since it requires to know precisely the physiological states (or phenotypic adaptation status) of cells, which are highly dynamics.

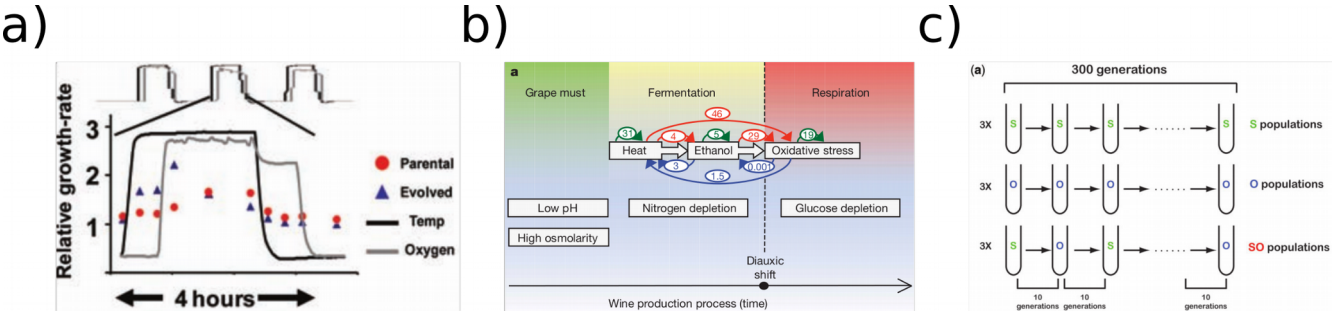


Fig INT2. Selection regimes of three studies working on anticipatory gene transcription. a) Figure from Tagkopoulos *et al.*, showing their selection regime: 37 °C then 21% O_2 then 25 °C then 0% O_2. **b)** Figure from Mitchell *et al.*, showing the steps of the wine production process. The numbers within arrows indicate the fold change protection of being exposed to a mild stress (arrow beginning) before a severe stress (arrow end). **c)** Figure from Dhar *et al.*, showing the selection regime in the three evolved populations. O is oxidative stress. S is salt stress.

However, one potential problem of studies on gene anticipation is that gene expression analysis is often performed on bulk populations. Thus, an alternative hypothesis for putative anticipation phenomenon of all cells, is that few cells in a population are becoming highly resistant to future stresses. This subpopulation of cells could be slow growing and highly resistant. Single cells studies could help in differentiating anticipation from such diversifying bet-hedging strategies.

B Yeast screens for genes expressed in anticipation of future stresses

There have been several yeast screens for finding genes important for trans-priming effects with different stresses. Those studies used sequencing and/or microarrays to quantify the abundance of mutants in the Yeast Deletion Library (YDL) after subjecting cells to two different stresses.

In 2011, the Gasch laboratory exposed the YDL to one of 3 different priming treatments (salt, Dithiothreitol (DTT) or heat shock) and one triggering stress (hydrogen peroxyde stress (H_2O_2)) (Berry et al., 2011) (Fig INT3b). Those priming stimulus were chosen at specific conditions (concentration and induction time) that induce a similar increase in H_2O_2 resistance in the wild-type. Unexpectedly, they found that mutants in which H_2O_2 resistance was lost/decreased were quite different depending on

the priming stress. For instance, they found only 28 genes to be important in all 3 conditions, and no Gene Ontology enrichment was significant in this set of genes. Those results suggest that different molecular routes launched through different priming stimulus can lead to resistance to the same secondary stress.

Zakrzewska *et al.* tested one priming condition (2 hours mild stress; acetic acid, heat (38 degrees), cold (10 degrees)) and 4 triggering stimuli (10 minutes severe stress; H_2O_2, heat (48 degrees), acetic acid) (Zakrzewska et al., 2011) (Fig INT3a). They found that all priming conditions induced increased survival to severe stress, except cold followed by heat shock. They observed a strong negative correlation (r=0.7) between the mutants' growth rate just before the lethal heat stress and their survival to this stress. Moreover, they observed that growth rate reduction was a critical factor for stress tolerance acquisition. They thus concluded that the reduction in growth rate is one of the key determinants of trans-priming. The authors used linear regression to correct the survival of mutants by their growth rate. After this correction, a Gene Ontology analysis indicated that genes involved in transcription and epigenetic mechanisms seem to play an important role in cross-protection.

In another Study, Gibney *et al.* screened a haploid prototrophic YDL for cis-priming from mild to lethal heat stress (37 to 50 degrees) (Gibney et al., 2013) (Fig INT3c). Death rates were computed as the slope of a linear regression of the fold changes of the mutants' barcodes over the course of the experiment. They made a test to find which strain have significantly higher heat sensitivity/death rate than the rest of the population. They found 65 and 10 mutants significantly sensitive after priming at a p-value of respectively 0.05 and 0.01. Unfortunately, they didn't exploit their data on priming: they stopped their analysis when they discovered that most genes that are significantly sensitive to heat after priming, are also sensitive to heat without priming. However, a better analysis could have been to divide the death rate with priming by the death rate without priming. And to test if this normalized value is significantly higher than the rest of the population. As found in other studies they observed a lack of correlation between genes expressed at mild stress (28–36 °C), and gene deletions sensitive to severe stress (30–50 °C).

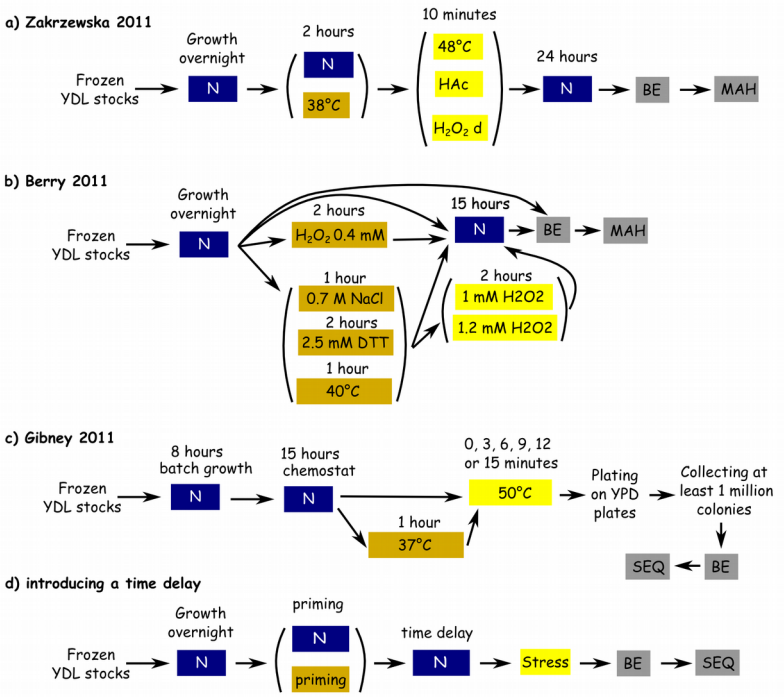


Fig INT3. Screens for acquired stress resistance in yeast. a,b) N (no stress) medium was YPD 30°C (generation time ~= 1h30). Experiments were performed in duplicates. In **b)**, cells were always maintained in exponential phase by frequent washing and dilution in fresh medium. **c)** N was a minimal medium with supplements (generation time ~= 5h). Experiments were performed without replicates. **d)** Potential design of future experiments to screen for yeast genes important for memory of stress. Abbreviations: BE: Barcodes Extraction; MAH: MicroArrays Hybridization; SEQ: Sequencing.

These different studies all independently confirmed that there is a lack of correlation between the set of genes expressed in response to a stress and the genes important for surviving it, even if the degree of correlations varies depending on the conditions (Berry et al., 2011). Multiple biological processes can explain this effect. For instance, proteins levels might be different from transcripts levels, post-translational modifications of proteins might be important for their functions, and low basal levels of a protein might be sufficient for it to perform its function. These studies also confirmed that growth rate is a key factor for yeast survival to lethal stress, with slow growing strains being more resistant. It is not clear if it is the most important factor for heat resistance as results were incoherent on this matter (Gibney et al., 2013; Zakrzewska et al., 2011). However, the importance of growth rate in stress resistance is consistent with the gene expression trade-off between growth and stress resistance observed in the yeast ESR. In fact, upon a stress, cells usually experience a lag phase during which the growth rate is reduced/absent for some time. This time is used by cells to make major physiological changes that will allow them to become stress resistant. Indeed, energy is limited and cell must optimize it to grow as fast as possible but also to survive to stressful and unpredictable conditions.

An important question is to determine if trans-priming effect are anticipatory or just “side effects”. That is, are genes important for a trans-priming effect expressed only on the purpose of priming to other stresses, or are they important for the priming stress and also for the triggering stress? In the latter case trans-priming would be a side effect of the response to the primary stress. In the former it would be a regulated mechanism for cells to be more fit in their ecology, as proposed in (Mitchell et al., 2009; Tagkopoulos et al., 2008). Dedicated experiments to answer this question should determine the

importance on fitness of all genes expressed in the priming stress on both the priming stress and the triggering stress with or without priming.

An inherent difficulty in working with the YDL is that slow growing strains generally get depleted in the population. This problem is more pronounced when studying stressful conditions due to two factors: A) “sick” cells can be even sicker in this condition, and B) slow-growing cells are generally more resistant, as described above. In the case of lethal stress, one has to get rid of dead cells before determining barcode abundances. To this end, in Berry *et al.*, Zakrzewska *et al.*, Gibney *et al.,* cells were grown after the lethal stress for, respectively, 24 hours, 10 generations, until the apparition of colonies. However, this procedure of amplification can introduce a bias and it is not clear how to normalize for it. The study of Berry *et al.* was potentially the only one that could correct for this bias since the authors included two conditions for this purpose in their design: the initial sample, and the initial sample after amplification (INT3-b).

It is important to consider that all those screens involved application of a secondary stress right after the first stress. This experimental design indicates genes that are important for acquiring stress protection after a first stress. However, it cannot indicate genes that are important for remembering that a stress occurred. Performing the same type of experiments but including a time delay (memory time; for instance 1 hour, 10 hours, 1 day, 1 week) between the two stimuli has not been done yet, as far as I know (Fig INT3d). This could allow to find genes important for memory of stress over long timescales. Finally, genes important for priming effects may vary with different time delays.

ii Mechanisms of memory of the first stress

Some molecular mechanisms of memory of stresses exposures have recently been discovered. Those memory mechanisms can be advantageous for cells in nature, since stressful conditions can be fluctuating. Many studies showed that this molecular memory can even be transmitted to daughter cells for several generations. In fact, molecular memory of a stress can last up to months, as has been shown in some plants and trees (Hilker et al., 2016). Thus, we can distinguish long-term memory that can be transmitted to daughter cells from short-term memory that is not transmitted. Knowledge on how memory is generated, maintained and transmitted to daughter cells is currently limited. However, proteins and RNAs with a high stability could be one of the most prominent vector of memory. Some important questions could now be addressed, such as: how much of the memory of previous stress exposure is transmitted from one generation to the next? Which types of molecular memory mechanisms are important for short-term and long-term memory? Are there some types of memory that are never transmitted to daughter cells?

Transcriptional (reinduction) memory usually refers to the ability of cells to induce a faster transcriptional response during the triggering stimulus. This effect can last for few generations (less than 10) and relies on proteins with long half-lives. However, other types of memory have been studied as well. For instance, people studying the GAL network observed persistent memory, which corresponds to a long-term memory of the state of the network over more than 10 generations (Stockwell et al., 2014).

The main mechanisms that have been proposed to generate molecular memory are the inheritance of long-lived memory factors, the propagation of chromatin marks, network states that are maintained by feedback loops and targeting of genes to the nuclear periphery. In the following sections, we describe these different mechanisms.

A Inheritance of long-lived memory factors

1 Proteins

Transgenerational memory can be strongly mediated by transmission of proteins during mitosis. This type of memory depends on two key factors: protein abundance and protein stability. Most long-lived and highly abundant proteins are transmitted to daughter cells, and can thus form a type of transgenerational memory. Thus, the large number of genes expressed during a priming stress (i.e. the ESR) contribute to the memory of stresses.

In fact, most yeast proteins have a long half-life, with a median of about 8.8 hours which corresponds to about 3 cell generations (Christiano et al., 2014). Thus, protein transmitted through cell division contributes to a big part of the total proteome of a given yeast cell. This simple fact could explain why many studies of memory of stress in yeasts describe memories that last a handful of generations (less than 10 usually). To explore the subject, I made some basic analysis on the dataset of (Christiano et al., 2014). I found that only 5% of yeast proteins have a half-life higher than 18.3 hours. A Gene Ontology enrichment analysis (FunSpec version July 2011, p-value cutoff: 0.01, with Bonferroni correction) on this top 5% of proteins with the longest half-life revealed a strong enrichment in metabolic processes (especially arginine biosynthesis), in ubiquitin proteins and in protein localization in the cytoplasm. Thus, proteins involved in stress responses do not seem to have a particularly long half-life. However, these measures have been performed at steady states where metabolism is the main biological process. Stressful conditions might induce the addition of protein modification marks that could increase the stability of stress-response proteins. Such protein modifications could include: phosphorylation, sumoylation, ubiquitination, ADP-rybosylation, acetylation, methylation… The Ubi4p protein corresponds to ubiquitin in yeast, a mark that targets proteins to degradation through the ubiquitin-26S proteasome system. The Ubi4p protein is very stable, as its half-life is 24.5 hours. Thus, ubiquitination might be used for long-term memory effects in yeasts. Some protein modifications determine the localization of proteins (i.e. sumoylation) or their activity status (i.e. phosphorylation). Thus, transmission of proteins with post-translational modifications could be seen as a memory of a protein state.

Besides protein stability and activity, another important factor of memory through transmission of proteins is protein abundance. Upon stress, proteins and transcripts abundance of a stress-resistant gene usually both increase (even if the correlation between protein and transcription expression is not always high). The resulting very high level of some specific stress response proteins can stay high even after

few cell divisions, which results in transgenerational memory of stress and sometimes cross-protection.

For instance, in 2012 the Gasch laboratory tried to determine the determinants of cellular memory of H_2O_2 trans-priming by NaCl. They subjected yeast cells to a mild priming stress (NaCl 0.7 M) for 60 minutes, then transferred them to YPD. They observed that these cells had a strong resistance to H_2O_2 stress, even 5 hours after the initial priming event. They found that acquired stress resistance was decreasing at a similar rate as the percentage of cells in the population that experienced the priming stress. Using a reporter system of daughter cells, they showed that they have the same level of resistance to H_2O_2 stress as their mother cell. This showed that stress resistance was transgenerationally transmitted (and not due to an original resistant population), and that cell division was progressively diluting this effect. Cytosolic catalase is a key protein for detoxifying the cell during H_2O_2 stress. Moreover, it has a very long half-life (63rd longest half-life out of 3773 proteins in Christiano 2014, with a half-life of 108.8 hours), thus it became their primary suspect. They found that Ctt1abundance increased more than 100 times after salt induction, and stayed at very high levels for 5 hours; which is the same duration as the memory of H_2O_2 resistance effect. Moreover, expressing Ctt1 with an inducible promoter provided a similar memory of H_2O_2 resistance than priming with salt.

In another earlier study, they showed that a Ctt1 mutant had a strong defect in trans-priming from NaCl to H_2O_2 but not with Heat Shock or DTT as a primary stress (Berry et al., 2011). Then, they looked at the gene expression levels of Ctt1 upon these primary stresses and found that Ctt1 was highly induced only in the NaCl condition, explaining why its effect on cross-protection in the other conditions was negligible. Heat shock and DTT trans-priming effect were probably due to the expression of proteins involved in glutathione metabolism genes. Indeed, deletion of two such genes (Gsh1 or Glr1) resulted in defects in acquired stress tolerance in HS or DTT but not in salt. They could not clearly identify a protein responsible for this memory effect, although the protein levels of a gluthatione peroxidase (Gpx1) were twice higher after HS than after NaCl.

Another well characterized protein that mediate transgenerational memory through protein dilutions is the Gal1p protein from the Galactose network (Kundu and Peterson, 2010; Stockwell et al., 2014; Zacharioudakis et al., 2007). GAL1 and GAL3 are paralogs. However, Gal3p plays a more important role in the initial induction of the Gal pathway: mutants for GAL1 (resp. Gal3) needs hours (resp. days) to fully induce the pathway (Stockwell et al., 2014). However, Gal1p protein is highly induced as compared to Gal3p (1000 times vs 3 times). This, coupled with its high stability makes Gal1p an ideal memory device. This type of memory is called reinduction memory and can last up to 6-7 generations. During this time, the cell re-induces the Gal network much faster than naive cells (Stockwell et al., 2014). It was first described by the Brickner lab in 2007 (Brickner et al., 2007) (see section IIIB3). An elegant experiment illustrated the importance of Gal1p levels for reinduction memory, as compared to chromatin factors. The authors generated an heterokaryon in which the cytoplasm originated from a cell that had a recent galactose experience, while the nucleus originated from a cell with no galactose experience (Zacharioudakis et al., 2007). The memory was maintained, proving that reinduction memory of galactose is more strongly mediated by cytoplasmic factors than by chromatin state. More informations about reinduction memory are available in section 1-III-B-ii-B-3.

2 Other potential inherited memory factors

There are numerous factors that could potentially be transmitted to daughter cells during mitosis to mediate transgenerational memory of previous environments. We will describe some of them in this section. Obviously, mRNA play a role in memory effects, even though it should be more transient than proteins as mRNA stability is much shorter. For instance, the median half-life of mRNA at steady states is 11 minutes as compared to 8.8 hours for proteins (Christiano et al., 2014; Miller et al., 2014). Thus, many transcripts are synthesized and degraded several times during a cell cycle. Few transcripts are long-lived, for instance only 100 (resp. 10) transcripts have a half-life higher than 60 (resp. 120) minutes. However, as for proteins, it is possible that some stressful conditions induce the stabilization of mRNA. For instance, a mild osmotic shock induces broad destabilization of most mRNA and specific stabilization of osmotic stress mRNAs (Romero-Santacreu et al., 2009). RNA interference through microRNA (miRNA) or short interfering RNA (siRNA) play an important role in regulating mRNA levels. Accordingly, they are involved in regulating memory of stress, as has been shown in different studies in plants (Hilker et al., 2016). Another way by which mRNA could be transmitted is through the storage in specific vesicles/organelles such as P-bodies and stress granules. Those ribonucleoprotein bodies are built upon stress and contain mRNAs and RNA binding proteins. There is a link between those two bodies as stress granules are formed in P-bodies. However, the formation or function of those bodies is not well characterized (Saarikangas and Barral, 2016). P-bodies (or processing-bodies) are foci in a liquid state that are also important for mRNA degradation, and as such contain RNA degradation enzymes. Stress granules are in a solid state and are sites of storage of mRNA that have been translationally silenced (by RNAi), or are stalled in translation initiation (Saarikangas and Barral, 2016). These bodies could allow a transient fast adaptation to adverse conditions, through the fast release of stored mRNA upon stress. For instance, it has been shown that cells in stationary phase contain hundreds of polyadenylated mRNA that can rapidly be released upon stress. The authors hypothesized that P-bodies were responsible for this effect (Aragon et al., 2006).

Interestingly, a recent paper showed that P-bodies are unidirectionally transmitted to daughter cells during cell division (Garmendia-Torres et al., 2014). This fascinating result suggest that P-bodies could be involved in long-term molecular memory effects. It remains to be shown how long this effect can last, as RNA degradation could induce a fast turnover of mRNA. However, this effect probably lasts for at least one generation, as the same study showed that P-bodies are important for daughter cells growth under nutrient limitations. It is well-known that division is asymmetric in budding yeast, as well as in bacteria and high eukaryotes (Yang et al., 2015). However, some proteins may be more asymmetrically inherited than others. A higher asymmetrical inheritance of stress responses proteins could result in longer memory of stress.

In contrast, to P-bodies, some memory factors are not transmitted to daughter cells. For instance, in 2013 Caudron and Barral showed that yeast cells remember events of unsuccessful mating, but do not transmit this memory to their progeny (Caudron and Barral, 2013). This type of memory is mediated by Whi3, a RNA-binding protein that sequestrates Cln3 mRNAs in cytoplasmic foci. Whi3p, like many other RNA-binding proteins, contains Q/N rich domains called Prion forming Domains (PrD) that promote molten globule like structures (super-assembly). Upon pheromone exposure, Whi3 releases Cln3 mRNAs and adopts a super-assembled conformation. Strikingly, this super-assembled whi3 seems to be highly stable and could possibly last for the whole life of the cell; resulting in an indefinitely non-

transmissible molecular memory mechanism. The authors propose the name of “mnemon” to describe this type of memory. Two key features distinguish mnemons from prions: 1/ prions form stochastically with a very low frequency, while Whi3 superassembly is absent without pheromones and is induced in all cells upon 3 hours of pheromone exposure and 2/ prions are transmitted to daughter cells; they are “infectious”; while mnemons are restricted to mother cells. Interestingly, some proteins involved in P-bodies and nucleoporin contains PrDs and could be other examples of mnemons. Two ortholog proteins in fruit fly and Aplysia regulate long-term plasticity of neurons and behave like mnemons. They contain PrDs domains and their super-assembly regulate memory maintenance (Keleman et al., 2007; Si et al., 2003). This, coupled with the high number of proteins containing PrDs in yeast, as found by a bioinformatic approach (Alberti et al., 2009) indicate that mnemon might be a common mechanism of molecular memory.

Prions can also provide cytoplasmically-inherited memory (Shorter and Lindquist, 2005). Prions are “infectious” since proteins in a non-prion state change their conformation to adopt the prion state when they meet prion proteins. Thus, they can propagate very fast in a population. Accordingly, when one cell with a protein in a prion state mates with another cell with the same protein in non-prion state, the resulting progeny will have their proteins in the prion state. Proteins in a non-prion conformation change their conformation at low frequency. However, stress can increase this probability. Thus, the prion state could be a type of adaptive bet-hedging mechanism that allows cells to randomly produce diverse phenotypes in situations of stress (Newby and Lindquist, 2013). The adaptive prions conformation then propagates in all the population as long as the stress persists. Afterwards cells with proteins in non-prion conformations will randomly appear and their type will reach fixation in the population, as the prion state is costly in absence of stress.

B Memory in the nucleus

1 Propagation of chromatin states

There is a chicken-and-egg debate opposing chromatin states to genetic circuits states for the maintenance of epigenetic memory. These two types of memory are strongly entwined and both seem to be important for maintaining cellular memory long after initiation of the signaling event: genetic circuits through feedback loops, and chromatin states through their own maintenance. Some argued that chromatin modifications are very dynamic processes that are too volatile to be a true support for maintenance of epigenetic memory (Nicol-Benoît et al., 2012). While this can be true in some cases (see next section), some chromatin marks are highly stable and heritable and should thus be considered as long-term epigenetic memory devices. The modality of propagation of histone marks should thus determine if they are causal or consequential in maintaining cellular memory. Indeed, if a mark is lost right after mitosis, and re-written thanks to the activity of genes from a genetic circuit, then it cannot be considered as an epigenetic mark (Steffen and Ringrose, 2014).

The main types of chromatin states are DNA methylation, histone localization (through remodeling), histone modifications and histone variants. DNAm usually occurs at CpG islands and promotes gene silencing. It is the most stable and heritable mark, and also the one for which the inheritance is the best characterized. Propagation of DNAm during replication occurs by segregation of all marks on one strand, and template copying on the other strand (Chen and Dent, 2014). Propagation of histone modifications or histone variants are less well understood. As for DNAm a semi-conservative mechanism could allow histone variants and histone marks to be conserved on one strand (Chen and Dent, 2014). However, the marks on the other strand are not always replicated. It has been observed that in some cases a histone modifying enzyme gets bound with a histone that contains the modification it performs, and replicate the modification on the opposite strand after DNA replication (Chen and Dent, 2014). However, this process is likely not the general rule as some marks decrease, get maintained or increase after mitosis (Wang and Higgins, 2013 table 1) and the dynamics of histone marks re-apparition after replication varies between marks (Bar-Ziv et al., 2016). Thus, different mechanisms are probably involved, and some, but not all, histone modifications could be truly independent providers of epigenetic memory.

Even if not causal, epigenetic mechanisms can participate in mechanisms of memory of previous conditions. An interesting example is the GAL reinduction memory. As mentioned before Gal1p is a key factor for this type of memory when the time spent in galactose is long (6/7 generations or > 12 hours). In 2007, Kundu and Perterson found that after growth on non-inducing nonrepressing raffinose medium and switch to galactose medium, cells needed 40 minutes to fully adapt (Kundu et al., 2007). Then, after one hour in glucose and switching back to galactose, cells needed only 10 minutes to adapt. They found that Gal1p was not necessary for this short-term memory while Gal3p and SWI/SNF were important. This result is the opposite of the GAL long-term memory effect. These contrasting results might be due to the fact that Gal3p proteins are still present after just 1h in galactose (1 or 0 dilutions occured for most cells), and Gal3p is a much better initial inducer of the pathway than Gal1p. SWI/SNF function in this process is to remove (remodel) nucleosomes upstream of GAL genes, allowing their induction. The effect of SWI/SNF only for short-term memory is probably due to the fact that it only accelerates the speed of induction (Kundu and Peterson, 2010; Kundu et al., 2007; Stockwell et al., 2014). It is thus not causative of memory but participates to the speed of induction. Similarly, to SWI/SNF, incorporation of the histone variant H2A.Z seems to induce a faster induction of the downstream genes. Some studies proposed that it could play a role for maintenance of the reinduction memory. However, its pleiotropic roles make it hard to confirm this result (Brickner et al., 2007; D'Urso and Brickner, 2014; Stockwell et al., 2014).

2 Genetic circuits and feedback in signaling pathways that generate stable states

The concept that feedback-loops can amplify the response to an initial transient stimulus and lock a system into self-sustaining stable states has been around for a long time with the pioneering work of (Monod and Jacob, 1961). In this paper, they already mentioned that the lactose operon in *E. coli* or the adaptation to galactose in yeasts are examples of such genetic circuits. Later studies formalized that feedback loops can generate long-term memory storage that survive protein turnover (Lisman, 1985). Genetic circuits can be bistable (resp. multistable) meaning that 2 (resp. multiple) stable states can be

attained. The concept of attractors describes those stable states in Waddington's epigenetic landscape (Ferrell Jr., 2012; Nicol-Benoît et al., 2012). Examples of bistable switches are now numerous, especially in microbes (Dubnau and Losick, 2006; Norman et al., 2015; Satory et al., 2011; Veening et al., 2008), and include: oocyte maturation in Xenopus oocytes (Xiong and Ferrell, 2003), acquisition of competency in *B. subtilis* (Mugler et al., 2016), bacterial switching between different morphologies (Gallie et al., 2015; Hernday et al., 2013), memory of hormone exposure for faster egg yolk production (Nicol-Benoit et al., 2011).

The concept of feedback loops generating memory is better understood with simple systems containing few loops. A simple and common case of a bistable switch is the double negative-feedback loop of the lambda phage. In this system two genes, CI and Cro, repress each other and determine if the virus is dormant (lysogenic cycle) or active (lytic cycle). Once a gene becomes dominant it will increase its repression activity until reaching a stable state. Only strong rare fluctuations can then change the state of the system (Norman et al., 2015). Another similar positive-feedback loop is the activation of the lactose operon in *E. coli*: lactose (the inducer) negatively regulates LacI (a repressor) that negatively regulates LacY (the lactose permease). Pioneering work, done by Novick and Weiner in the laboratory of Jacques Monod, revealed this circuit (Novick and Weiner, 1957). They showed that subjecting isogenic *E. coli* cells to low concentrations of the inducer TMG (thiomethyl-,3-D-galactoside) results in a binary outcome with some cells fully activating the operon and others not at all. The operon state was heritable. Thus, uninduced cells become induced by stochastic bursting events of lac genes. The low basal level of LacI in cells (less than 1 transcript per cell per generation) can participate in generating cellular noise and thus cells with different fates (Satory et al., 2011).

Acquisition of competence in *B. subtilis* is also a relatively simple system: the compK protein contains a fast-acting positive-feedback loop and a slow-acting negative-feedback loop on itself. This network structure describes an excitable system: only strong perturbations can allow the system to leave equilibrium (i.e. to become competent). Then, the system becomes activated for some time before going back to equilibrium (Norman et al., 2015).

We described above the reinduction memory of the GAL network, which is strongly dependent on Gal1p (or Gal3p) levels and fades over few generations. However, another type of memory exists in the GAL network: the persistent memory (Stockwell et al., 2014). This type of memory could last much longer, perhaps indefinitely, and is mediated by the numerous nested feedback loops in the GAL network. Persistent memory corresponds to the bimodal induction of the GAL genes of glucose primed cells exposed to a partially inducing medium (either intermediate galactose and low glucose, either low galactose). Inversely cells previously primed with a non-repressive medium (galactose or raffinose) get activated uniformly by the partially inducing medium. This type of memory can last for more than 27 hours and was first described by Biggar and Crabtree in 2001, followed by Acar et al in 2005 (Acar et al., 2005; Biggar and Crabtree, 2001). It is thought to depend on the several feedback loops in the GAL network: 3 feedback loops (Gal1p, Gal2p, Gal3p) and 2 negative-feedback loops (Gal1p, Gal80p) (Stockwell et al., 2014). The fact that only partially inducing conditions result in bimodal Gal1p expression indicates that GAL network is a bistable switch and partial inductions activate stochastically only certain cells in the population. Feedback loops then amplify this transient induction stimuli. Acar *et al.* showed that a strain with a modified inducible promoter for Gal3p lose this persistent memory, indicating that Gal3p feedback loop promotes this effect (Acar et al., 2005). Consistently, Gal80p showed the opposite effect with its negative feedback-loop reducing the strength of Gal1p activation.

Finally, most genetic circuits are probably much more complex than the ones described above, with for instance the genetic circuit controlling the white-opaque switch in *C. albicans*. *C. albicans* can adopt one of four phenotypes: gray, white, opaque or GUT (Gastrointestinally indUced Transition) (Pande et al., 2013; Scaduto and Bennett, 2015). Those states are highly stable and heritable, thus potentially caused by a multistable circuit. Moreover, they show high phenotypic divergence (despite similar genotype) in multiple traits such as: biofilm formation, mating, stress resistance, favoured niche and metabolism (Scaduto and Bennett, 2015). Interestingly, the white-opaque switching rates are very low at 25 degrees (10^{-4}) and increase drastically at high temperatures (10^{0}) and other stressful conditions. A recent systems biology study focused on 6 main regulators of the "all-or-none" white-opaque switch (Hernday et al., 2013). They observed multiple intertwined feedback loops among the main regulators or among their target genes: 3,225 for the opaque network and 36 for the white network. The authors thus propose that these loops could be responsible for maintaining the memory of the phenotypic state, and a high degree of redundancy could explain the resistance to perturbations. Interestingly, the white circuit seems to be included in the opaque circuit, on the contrary of lamda's Cro and CI circuits that are mutually exclusive. This show that network structures conferring different stable phenotypic states can evolve differently.

3 Memory at the nuclear periphery – reinduction memory

Reinduction memory, as mentioned above, is the ability during few generations to induce gene expression faster in response to a second identical or similar stimulus (cis or trans priming). Reinduction memory was first described in 2007 for the inducible inositol-1-phosphate synthase INO1 gene (~3/4 generations) and the GAL1 gene (~7/8 generations) (Brickner et al., 2007). Other types of reinduction memory were then discovered with similar durations: yeast trans priming of NaCl to oxidative stress (~ 4 generations) (Guan et al., 2012), interferon gamma (IFN-g)-induced class II major histocompatibility gene DR alpha (HLA-DRA) in Hela cells (> 3 generations) (Light et al., 2013). Reinduction memory was also observed in plants, sometimes with prolonged duration (Pastor et al., 2013). Protein abundance plays a role in this type of memory and could explain why it lasts for a number of generations realistic with protein number dilutions, as mentioned above for Gal1p and LacY. However, in the last decade, several studies revealed that cells employ additional sophisticated mechanisms conferring fast reinduction. After a priming stimulus, this memory consists of three features that are shared in the examples cited above: relocalization of the induced genes at the nuclear periphery through interactions with different Nuclear Pore Complexes (NPC), dimethylation of H3K4, and binding of poised RNA Polymerase II PreInitiation Complex (RNAPII PIC). The last feature allows genes to be "ready to fire" if the stimulus re-appears (D'Urso and Brickner, 2014; D'Urso et al., 2016). In fact, the presence of genes "ready to fire" in reinduction memory effects suggests that other genes, not involved in memory effects, might be regulated/primed the same way: with a faster induction than others thanks to similar regulatory mechanisms.

The mechanisms of INO1 memory are the best characterized so far. Upon activation, the DNA sequence of INO1 gets relocalized and interacts with Nup100p, a Nuclear Pore Complex (NPC) (Brickner and Walter, 2004; Brickner et al., 2007). Zip codes are Transcription Factors Binding Sites (TFBS) that, when bound by TF, promote targeting to the nuclear periphery and interaction with the

Nuclear Pore Complex (NPC). They are important for both reinduction memory and chromatin boundaries. Upon INO1 activation, two INO1 zip codes, called the Gene Recruitment Sequences (GRS), are bound by the transcription factors Put3p and Cbf1p, which result in DNA relocalization to the NPC Nup100p (Brickner and Walter, 2004; D'Urso and Brickner, 2014). Right after gene repression, another zip code, termed the Memory Recruitment Sequence (MRS), is bound by the transcription factor Sfl1p. Then, a modified version of the Compass/Set1p complex, without Spp1p, deposits H3K4me2 marks on histones at the INO1 locus. Those marks are subsequently maintained by Setp3 from the SET3C HDAC complex. H3K4me2 marks and Cdk8p binding are both needed to recruit inactive, poised RNA polymerase II PreInitiation Complexes (RNAPII PIC) at the promoter. Furthermore, the role of Cdk8p seems to be specific to memory since inactivation of Cdk8p did not affect INO1 induction but only reinduction and RNAPII poising (D'Urso and Brickner, 2014; D'Urso et al., 2016).

In 2012, Guan et al studied NaCl to Oxidative stress trans-priming in yeast and found a group of 77 genes that are quickly induced only in primed cells. They showed that Nup42p, but not Nup59p, was necessary for this memory effect. Nup100p could be involved too, but results were difficult to interpret since naive *nup100* mutants showed a reinduction level similar to the primed wild-types. Strikingly, they identified a motif in the upstream regions of those genes that was very similar to the Memory Recruitment Sequence of the INO1 gene. This suggests that transcriptional poising and reinduction memory could be a common cellular mechanism.

Finally, we can imagine that relocalization of proteins, and not only genes, or even organelles, might as well be important mechanisms of memory of previous environmental conditions. Proteins marked with localization tag could mediate such effects, but more complex regulations might also exist. For instance, the speed at which molecules propagate within the cell (e.g. transcription factors shuttling) might keep in memory regulatory events that occurred long ago.

IV Artificial fluctuations to characterize biological dynamics

A Experimental evolution in periodic fluctuations to characterize the evolution of genetic variance

Box1: genetic variance and heritability

Phenotypic variance indicates how much variations there is for one trait in a population. The genetic variance and environmental variance indicates how much genetic or environmental factors drive this variation. Genetic variance is the sum of three terms: the epistasis variance (interactions between loci), the dominance variance (dominant effect) and the additive genetic variance ("raw" individual effect of each loci independently of dominance or epistasis). In the 1930s, Fisher proposed that it corresponds to the rate of change in biological fitness (Fisher, 1930). His work opened the door to extensive quantitative genetic analysis, illustrated by those formulas:

$V_P = V_G + V_E$ $V_G = V_A + V_D + V_I$ $H^2 = V_G/V_P$ $h^2 = V_A/V_P$

$V_P$is the phenotypic variance, V_G the genetic variance, V_E the environmental variance, V_A the additive variance, V_D the dominant variance and V_I the epistasis variance, H^2 the broad sense heritability and h^2 the narrow sense heritability.

Heritability is a measure of how much heritable the phenotypic variance is. Broad-sense heritability includes non-additive genetic effects of epistasis or dominance, while narrow-sense heritability is restricted to purely-additive effects. Broad-sense heritability can be used for instance to determine the respective importance of genetics and environment for the apparition of a particular disease. While narrow-sense heritability can be helpful in determining how much a crop/livestock can be improved by directed evolution/breeding. The problem of missing heritability underscores the paradox that while numerous loci contributing to traits have been found their summed effect usually corresponds to a small proportion of the total narrow sense heritability. Thus, either many variants cannot be discovered due to a lack of power in Genome Wide Association Studies (GWAS), or narrow sense heritability is overestimated (or both) (Zuk et al., 2012).

Starting from the 1960s, a series of experiments in drosophila aimed at determining how additive genetic variance of various traits evolve in steady or heterogeneous conditions (Beardmore, 1961). If we reformulate, the question asked was: do heterogeneous conditions increase the speed of adaptation? Since drosophila generation time is typically longer than a week, those experimental evolutions experiments usually lasted one or two years, sometimes up to 5 years. Authors evolved several replicates of highly inbred or wild populations of drosophila in temporal or spatial heterogeneity, with a period of about 1 generation, or in steady conditions (Table INT5). Temporal heterogeneity consisted in periodic alternations between two media. Spatial heterogeneity consisted in periodically splitting the population in two different medium and joining the two separated populations at the end of the period. Many traits studied were related to fitness, such as: egg to adult survival, productivity, competitive ability. But some were partly related or unrelated such as: sternopleural chaeta (bristle) number, body size, body mass, wing shape.

The theory that is generally proposed is that the treatments leading to a higher genetic variance, and thus a higher speed of adaptation, are first spatial heterogeneity, then temporal heterogeneity and finally steady conditions (Huang et al., 2015; Mackay, 1981; Yeaman et al., 2010). However, results were often contradictory, with some studies claiming to confirm the theory and others to refute it. To illustrate that, the results of 3 experiments are shown in Table INT5 (design on table Table INT4). On this table, we can see that temporal heterogeneity generated the highest variance in the Mackay *et al.* study, while spatial heterogeneity generated the highest variance in the Huang *et al.* study, and no clear pattern emerged in the Yeaman *et al.* study.

A criticism that is commonly made to studies that do not find results in accordance with the main theory is that the traits that they measure are not, or vaguely, related to fitness. Thus, the link between the organism's adaptation and the trait under study is not clear. For instance, in the Yeaman *et al.* study, traits under study were all related to wing veins length. In contrast, Huang *et al.* focused on traits quite

related to fitness. They found a higher genetic variance in the spatial environment, as predicted by the theory. However, most of their results were not, or barely, statistically significant, due to a high measurement error. To conclude, the differences in results obtained in various studies is probably caused by differences in the exact experimental designs and traits measured. Future works could elucidate if fitness-related traits have indeed higher genetic variance in heterogeneous environments.

Study	Experiment duration	Period(s)	Medium A / B	Replicates
Mackay 1981	2 years	2 weeks, 8 weeks	no stress / ethanol 15%	2
Huang 2015	~2.5 year (45 generations)	1 generation	Salt / Cadmium	5
Yeaman 2010	116 weeks (~43 gen)	8 weeks	25°C and 16°C	5

Table INT4. Design of drosophila experiments aiming at measuring genetic variance of various traits in steady, temporally varying or spatially varying environments. Period indicates the period in the temporally heterogeneous condition. Environment A / B indicates the two growth environment (supplemented medium or heat) that alternate periodically in the temporally heterogeneous environment, or in which the drosophila were split ~every week in the spatially heterogeneous condition. Replicates indicates the number of replicates per condition.

Study	Trait	Steady	temporal	Spatial
Mackay 1981	SCN	1.4 (no stress)	4.7 (ST), **5.4 (LT)**	2.8
Mackay 1981	ABN	**5.8 (no stress)**	**5.8 (ST)**, 3.6 (LT)	5.2
Mackay 1981	BW	0.015 (no stress)	0.045 (ST), **0.048 (LT)**	0.041
Huang 2015	MMS	~0 (Cadmium), ~ $2.5*10^{-4}$ (Salt)	~ $3*10^{-4}$	**~ $4.2*10^{-4}$**
Huang 2015	MMC	~ $5*10^{-6}$ (Cadmium), ~ $2.5*10^{-5}$ (Salt)	~ $2.6*10^{-5}$	**~ $6*10^{-5}$**
Huang 2015	SS	~ 0 (Cadmium), ~ 0 (Salt)	~ 0	**~ $1*10^{-3}$**
Huang 2015	SC	~0 (Cadmium), **~0.027 (Salt)**	~ 0	~ 0.003
Yeaman 2010	CEN	26.82 (Cold), 24.81 (Hot)	26.37	28.15 (SP), **30.63(SLM)**
Yeaman 2010	ANG	**19.11 (Cold)**, 17.98 (Hot)	18.29	17.92 (SP), 17.98(SLM)
Yeaman 2010	LIN	0.17 (Cold), 0.19 (Hot)	**0.2**	0.14 (SP), 0.19 (SLM)

Table INT5. Genetic variance of various drosophila traits in steady, temporally varying or spatially varying environments. In bold are highlighted the conditions that had the highest genetic variance for the considered trait. Not all traits measured in each study are included in this table. Abbreviations of traits. SCN: Sternopleural Chaeta (bristle) Number, ABN: Abdominal Bristle Number, BW: Body Weight, MMS: Male Mass assayed in Salt, MMC: Male Mass assayed in Cadmium, SS: Survival assayed in Salt, SC: Survival assayed in Cadmium, CEN: Centroid, ANG: Angle7-8-9, LIN: Line9-10. CEN, ANG and LIN are traits related to wing veins length. Abbreviations of environmental regimes. ST: Short Term (2 weeks period), LT: Long Term (8 weeks period), SP: Spatial heterogeneity with panmixia, SLM: Spatial heterogeneity with Limited Migration. Values from Huang *et al.* 2015 were inferred from a Figure, hence the "~" sign.

B Applying methods from engineering to predict biological dynamics

i Introduction to Frequency Response Analysis and System Identification

Control systems are systems that can transform an input into a desired output (Åström and Murray, 2008). *Control theory* aims at mathematically determining what input to provide to a control system to get the desired output. "Open loop" control systems describe systems where the input does not depend on the output. Contrarily, systems where the input depends on the output through a sensor and a controller are usually called *closed loop* control system (or negative-feedback control, or automatic control). Control theory is useful in most engineering fields, since feedback-control can be required for many applications. In addition, to helping in the construction of feedback-control systems, system theory allows to build models on the system, make predictions, estimate the dynamics of the system or its noise filtering behaviours, and testing them. Control theory can only be applied to *Linear Time Invariant* systems (LTI).

LTI systems are a class of systems that respond in a certain way when subjected to an arbitrary input. Specifically, LTI systems are *linear systems* and time invariant systems (Åström and Murray, 2008; Zadeh and Deoser, 1976). Linear systems follow the properties of *homogeneity* and *additivity* (also called superposition). Homogeneity means that if the input is scaled by a given factor, the output will be scaled by the same factor. Additivity means that if input A gives output A' and input B gives output B', then input A+B gives output A'+B'. Linear systems have behaviours that are fully explained by individual effects of components of the system: there is no interactions between members of the system or between the system and the environment. *Time invariance* (resp. translational) means that applying the input at any later time will produce the same output. Similarly, "translation invariance" means that a shift in space doesn't change the output of the system.

Restrictions to define a system as LTI are so severe that almost no real world system meets them. Indeed, most systems contain nonlinearities. However, many complex systems can be approximated accurately by an LTI model, or they can be locally linear. Sometimes the system is nonlinear for certain input (wave) of interest. In this case linear rectifiers can be applied. The advantage of working with LTI systems is that they are solvable mathematically, which allows for deep analysis.

Control theory allows for extremely powerful predictions since, once the model has been determined, it can theoretically predict the output from any input. Simply, this is due to the fact that any signal can be decomposed into a multitude of sinusoids of different frequencies, amplitudes and phases, thanks to the Fourier transform. The model, called the *transfer function*, consists in a map between input frequency (or wave width) and output (phase and amplitude gain). Thus, an output can be predicted for each of the decomposed sinusoids (as long as it is in the range of the inputs fitted by the model). The property of additivity then allows to sum the individual effect of all decomposed sinusoids in order to predict the output of the signal.

The reason why signals are decomposed in sine wave and not in another types of waves such as square waves, exponential waves or different aperiodic signals, is that sinusoids are the only waveform that do not change shape when confronted to Linear Time Invariant (LTI) systems. This means that a sine wave retains its wave shape when added to another sine wave of the same frequency and arbitrary phase and magnitude, contrary to square waves for instance. This is due to the fact that sine waves respect the properties of LTI systems (time invariance, homogeneity and additivity). In mathematical terms, sine

waves are Eigenfunctions (a function that only multiply the input by a scale factor) of LTI Systems, which simplifies mathematical analysis.

Frequency Response Analysis refers to determining experimentally the output (amplitude, phase) of the system to a range of input frequencies (Ang et al., 2011; Åström and Murray, 2008). This generates *Bode plots*, which are scatter plots of amplitude (or phase) vs frequencies. Analysis of Bode plot are very instructive about the filtering behaviour of the system. If the output amplitude is amplified at some frequencies (which is called system *gain*), then the system is said to be resonant. And the most amplified frequency(ies) of a system is (are) called the *resonant frequency*. The output amplitude can also be filtered: in this case only some frequencies are preserved, the others are attenuated and almost disappear. There are different types of filters: *low-pass*, *high-pass* or *band pass filters* correspond to systems that keep respectively only low, high or intermediate frequencies. The *cutoff* frequencies (also called *corner* frequencies) are frequencies at which the output starts to be attenuated. Thus, there is 1 cutoff frequency for a low-pass filter or a high-pass filter and 2 cutoff frequencies for a band-pass filter. The frequencies where the system is not filtered are called the *bandwidth* of the system. For example, for a band-pass system, this corresponds to the frequencies between the two cutoff frequencies.

System identification is a procedure that allows to obtain an accurate model of the input/output behaviour of a system (Ang et al., 2011). This procedure consists in fitting a model (or transfer function) to data from a Bode plot, which allows to solve the system for any frequency within the range of tested frequencies. System identification starts by measuring experimentally the response of the system to different input frequencies. Then, for each input Fourier filtering is applied, only at the input frequency, to get rid of the noise. This allows to generate Bode plots to which transfer functions can be fitted. A transfer function is a ratio of polynomials of a complex variable. The shape of the bode plot (number of cut-off and roll-off (the steepness of the decrease after the cutoff)) determine the degree of the numerator and denominator equation before fitting. The last steps are to add a linear rectifier if needed, and to validate the accuracy of the model through a testing set (experimentally testing frequencies not used for building the model). Finally, several iterations of different steps of the procedure can be repeated to improve the accuracy of the model.

Importantly, system identification can also be applied to square wave inputs (Ang et al., 2011). In this case, the formula for the Fourier series is slightly modified. A square wave is a highly non-linear input since there is a vertical increase in the input when the medium change, where one x can take several y values. This nonlinearity always results in a slightly higher errors near the beginning and end of squares, an effect known as the Gibbs' phenomenon. However, increasing the number of sine waves in the Fourier series can decrease this error close to zero.

Finally, models can be classified in three categories: white, gray or black-box models. “White-box” models or “physical” models describe models where only information about the internal structure of the system and different experimentally measured parameter values are used to build the model (Ang et al., 2011). This approach allows to test different system structures or parameter values. For instance, in biology it allows to test for different mechanisms within a pathway. The drawback is the difficulty to obtain accurate information about the system, and to model it correctly due to the complex nature of most systems. “Black box” models are using system identification procedures with no a priori knowledge to build the model. This approach offers robustness and predictive power, but no knowledge about the internal details of the system. “Gray-box” models are intermediate models where some knowledge about the system is incorporated, and other parameters are estimated using system

identification procedures. Ultimately, researchers want to understand the internal details of the system under study. Gray-box models can be used as an intermediate step towards this goal.

ii Application of System Identification to characterize network dynamics

Frequency Response analysis has crucial applications in digital signal processing (sound, image, wireless communications, electronic circuits). Systems identification has broad applications in engineering, but also in physics, economics, medicine, social systems or biology. This procedure allows to set up a predictive model of a system, with no a priori knowledge, by stimulating it at certain frequencies.

Periodic fluctuations can be applied to characterize pathways dynamics. This approach consists in periodically applying an activating input to a cell and recording the output, which reflects activation of the pathway. It requires using microfluidic devices that can trap cells and quickly change the input in the medium. These are coupled with microscopy tools for monitoring pathway activation. Both the input and output should be easily measurable and modifiable. Thus, this approach is limited to few well-characterized systems and has been applied to the physiological adaptation of yeasts to galactose or high osmolarity.

In fact, one could try to measure the physiological adaptation of cells after a single step of environment change. This usual approach is much less intensive experimentally. However, periodic fluctuations have two big advantages. First, they allow to reduce the noise by performing multiple measurements. Indeed, sampling many cycles improve the Signal to Noise Ratio (each cycle acting as a replicate). Though this can be difficult to implement for the lowest frequencies (longest periods), for which only one or few cycles are measured. Second, they open the door to frequency response analysis and system identification. Those methods can have broad implications for the fields of synthetic biology and systems biology.

Indeed, synthetic biologists try to adopt an engineering approach to biology. They wish to control the behaviour of organisms, by applying the principle of abstraction, design, modeling, modularity and standardisation to engineer living systems properties. Synthetic biology aims at re-wiring living organisms to produce new functions. For instance, this includes bioremediation (using microbes or plants to de-pollute certain sites), bio-production (using microbes to produce medicines, biofuels, or other valuable compounds). “Black box” modelling thus seems well fitted to achieve the level of control, prediction and standardisation targeted in synthetic biology.

System identification opens the door to advanced analysis on different aspects of pathways dynamics, such as giving boundaries to the rate of action of each component in a pathway, establishing the range of amplified or filtered frequencies, and determining rate-limiting components of a pathway (also called dominant processes). In the case of filtering, it is also interesting to determine the pathway's bandwidth, which is the range of frequencies at which outputs faithfully follows inputs. Outside this range the pathway is usually “blind” to the fluctuations, and consider that the environment is steady

with an input concentration that is the average of both alternating media. The dominant component(s) are acting at the same rate as the bandwidths, while other components operate at higher frequencies. Finally, the identification of resonance frequencies can reveal unexpected weaknesses of the pathway at specific frequencies, when it gets more activated than it should (Mitchell and Lim, 2016). This could have broad applications, such as targeting specific cell types (i.e. cancer cell, infected cells) that are particularly sensible to a given frequency of treatment. Identifying such failures in pathways activation (hyperactivation) in response to non-natural inputs (periodic fluctuations) could also be a new approach to understand the pathway, the resonant frequency being the frequency at which the pathway reaches its maximum activation level in the activating condition.

In order to use the methods of system identification, one must first determine if the system is linear, or at least locally linear at the frequencies of interest, and if not, try to find an appropriate linear rectifier. In the case of pathway response to square wave periodic fluctuations, the homogeneity principle implies that the pathway is twice as much activated when the input amplitude doubles. The additivity principle indicates that the activation of the pathway to input A+B should be equal to sum of its response to A and B.

There have been a handful of yeast studies that performed this type of experiment so far (Table X). Three of them focused on the High Osmolarity Glycerol pathway in periodic osmotic stress (Hersen et al., 2008; Mettetal et al., 2008; Mitchell et al., 2015), and one on the Galactose pathway in periodic glucose fluctuations and constant Galactose (Bennett et al., 2008). The three Hog studies used gray-box models; linear systems theory combined with knowledge from the literature. On the contrary, Bennet *et al*. constructed a white box model after analysis of the literature. This was possible since the GAL network had been extensively characterized. Due to the knowledge of internal details of the system, white box models allow to make predictions on specific mechanisms. Bennet *et al.* were indeed able to solve a discrepancy between their model and the data (the network was more sensitive than expected at low frequencies) by hypothesizing and experimentally validating a faster degradation rate in glucose than in Galactose of the Gal1 transcript.

Study	Type of model	Input: wave type	Input: Periodic stress	Input: Range of periods	Input: Range of frequencies in Hz	Measured Output	Type of filter	Resonant frequency in Hz (period)
Mitchell 2015	grey box	square	0.4M Kcl	0.5 min – 128 min	$1.3*10^{-4}$ – $3.3*10^{-2}$	∫ (Hog1-GFP NE)	Band-pass	$1*10^{-3}$ (16 min)
Mettetal 2008	grey box	square	0.2M NaCl	2 min – 128 min	$1.3*10^{-4}$ – $8.3*10^{-3}$	Hog1-YFP NE	Band-pass	$2.1*10^{-3}$ (8 min)
Bennet 2008	white box	sine	0.25% glucose	45 min to 4.5 h	$6.2*10^{-5}$ – $3.7*10^{-4}$	Gal1-yECFP WC	Low-pass	$6.2*10^{-5}$ (4.5 h)
Hersen 2008	grey box	square	1M Sorbitol	1 sec – 16 min	$1*10^{-3}$ – 1	Hog1-GFP NE	Low-pass	$1*10^{-3}$ (16 min)

Table INT6: Studies showing amplitude Bode plots in yeast. Hog1-GFP/YFP NE (resp. Gal1-yECFP WC) indicates the activity of the HOG (resp. GAL) pathway. Integral of Hog1-GFP NE indicates transcriptional output of the Hog pathway. In Mitchell *et al.*, other output were measured: generation time (cell growth, RP: 8 min), pixels (cell area, RP: 8 min), pStl1 (reporters of osmotic transcriptional response, RP: 16 min), pFre (reporter of the invasive-growth pathway , RP: 8 min). The grey box used in Mitchell *et al.* was adapted from Muzzley *et al*. Other studies build their own models. Abbreviations: RP: Resonant Period; NE: Nuclear Enrichment; WC: Whole Cell.

Three studies measured the activity of the HOG pathway in response to osmotic stress (Table I-IV-B). Interestingly, while applying different inputs, they all observed a resonant frequency at around 10^{-3}

hertz. In contrast, the Gal network seems to be acting at much slower time-scales with a resonant frequency potentially lower than $6.2*10^{-5}$. Those results suggest that oscillations in osmotic stress may be more frequent and critical to yeast survival than fluctuations in galactose levels.

V Genomics of yeast fitness in periodically fluctuating environments

A What can we learn by experimentally studying fitness in periodically fluctuating conditions?

Fitness is a special phenotype/output. Indeed, fitness is a measure of the selective advantage of an organism in a given environment. This selective advantage is important to understand the evolution of organisms. In addition, fitness is a useful phenotype to determine the genes that control the complex and key process of cellular division control, in various contexts. Fitness of microbes is usually measured in steady conditions. However, measuring microbes' cellular proliferation/fitness in binary periodically fluctuating conditions can provide much information regarding their dynamics of phenotypic adaptation through plasticity. It could help understanding the forces that shape selection in the wild, and the nonlinear effects of changing environments that cannot be predicted from steady conditions.

Yeast genes important in a given condition can be systematically determined by analysing the fitness of yeast null mutants in this condition. However, the evolutionary interpretations of fitness values in the wild, where environments change frequently, are not straightforward. We don't know yet if we can predict fitness in dynamic conditions from fitness values in steady conditions. Combinations of conditions in the wild are infinite. Thus, being able to extrapolate selective advantage in fluctuating conditions from fitness data in steady conditions would be a major advance (Cohen, 1985). A first step in this direction is to determine how much we are able to predict fitness in simple environmental changes such as binary periodic fluctuations. Although the answer to this question likely depends on the specific strains and conditions tested, general principles could be discovered.

Specifically, we can wonder: does fitness in a periodically-fluctuating environment equal the average of fitness in steady conditions? Is there a critical rate (or frequency) of environmental changes at which predictability degrades? If we have, for instance, steady state fitness values after several hours of growth in two condition, until which frequency of binary periodic fluctuations can we still predict fitness? Are there some genes that are more predictable than others? And if yes, which mechanisms can explain these differences? Can we discover the sources of nonlinear effects?

As described in chapter 3A, different strategies of adaptation to fluctuating environments have been described, and many studies made predictions on the conditions that favours one strategy over the others in the wild. Working with periodic fluctuations may allow to experimentally test those predictions and to determine how frequent those strategies are and their impact on fitness in conditions mimicking real environments (Fig INT7.3).

Unicellular organisms use environmental cues to adapt to changes in the environments. However, not all cells commit to launch a signaling pathway/stress response right after sensing a cue. Some cells will never commit: they are insensitive to this particular cue. Other cells will commit later: they wait for higher extracellular or intracellular levels of the cue. After commitment, cells usually experiment a phase of intracellular signaling that will allow them to become well-adapted to the new environment. During this phase, often called the lag phase, there is a reduction or an arrest of growth in order to rewire intracellular signaling. For those reasons, committing quickly is not always adaptive. For instance, a fast commitment can be maladaptive if the new environment quickly disappears, especially for long lag phases. Thus, cells often face a trade-off between fast-growth and adaptation to new conditions. Measuring fitness of many different strains in fluctuating conditions can allow to systematically identify candidate strains with a modified speed of adaptation (which include both the time to commitment and the lag phase duration) (Fig INT7.2). In fact, a short lag phase can be too small to be detected after a single environmental switch. In this case, performing multiple environmental switches, as in periodically fluctuating conditions, can allow to amplify the signal.

In addition, to changes in the speed of adaptation, fluctuating conditions can reveal memory mechanisms (Fig INT7.1). Indeed, epigenetic mechanisms sometimes allow cells to remember previous conditions for a certain duration (see chapter 1-III-B). This behaviour is especially advantageous for cells in a fluctuating environment where recent stresses have more chances to occur again, such as in periodic fluctuations. However, memory of stress can be a regulated process, and thus it can require energy that could be invested in optimizing growth in the current environment. Thus, as for the speed of adaptation, there can be a cost of maintaining a type of cellular memory. This is why memory of stress can be maladaptive if the memorized stress does not re-occur, or if it occurs too mildly over a short time.

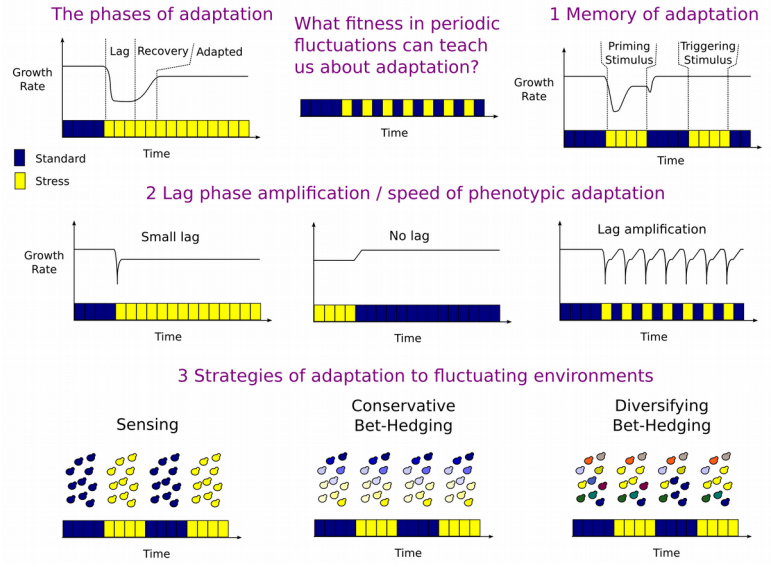


Fig INT7. What fitness in periodic fluctuations can teach us about adaptation?

B Low throughput measures of fitness in fluctuating conditions

In this section, we will describe some of the few studies that have measured fitness during or after binary periodic fluctuations. We will see that they observed some of the phenomenon described above of short and long-term memory of stress, lag evolution/amplification and different strategies of adaptation.

Stomp *et al.* studied chromatic adaptation, the ability of some micro-organisms to change their color in function of the prevailing light spectrum (Stomp et al., 2008). They exposed 3 cyanobacterial species – one red, one green and one that can change its color in ~7 days – to fluctuations in incident light color at 3 different frequencies (Table INT12). Despite the fact that the strain able of chromatic adaptation outgrew the 2 others in steady red or green light, they found that the rate of exclusion was higher in fluctuating light conditions. This was especially true for longer periods of fluctuations, owing to the full adaptation of the flexible strain. This study nicely illustrates that the benefits or phenotypic adaptation depends on the dynamics of environmental changes.

Beaumont *et al.* evolved Pseudomonas fluorescens populations with alternations of growth in steady or shaken microcosms (Beaumont et al., 2009). Their selection protocol at each round consisted in growing cells in one of the two conditions for 72 hours (one propagation). Then, they looked if cells with new heritable morphologies appeared. If not, cells were grown for another propagation in the same condition. Otherwise a new round of selection started in the alternative growth condition. It should be noted that this environment is not strictly periodic as rounds can have various length. After 6 and 9 rounds of selection (corresponding to 16 and 14 propagations) they observed that 2 out of 12 replicate lines evolved bet-hedging phenotypes, that were switching at each new round of selection between two different morphologies adapted to each condition. Genome sequencing revealed that 9 mutations were involved; with the last one being necessary and sufficient for the rapid switching phenotype. However, the previous 8 mutations were important to provide a fitness advantage to the switching phenotype, possibly through epistatic effects. This study shows that selection in a fluctuating environment can rapidly lead to the apparition of bet-hedging, generalists' genotypes with highly plastic morphologies.

In 2008, Acar et al engineered two yeast strains to stochastically switch between two phenotypes adapted to two different environments (Acar et al., 2008). They then measured the growth rate of these strains in an environment that changed periodically between those two environments at different frequencies. They found that, as expected, the fast switcher strain was growing faster than the slow switcher strain when fluctuations were rapid, and vice-versa. This suggests that phenotype switching can be an appropriate strategy of adaptation in fluctuating environments.

Bacterial resistance to antibiotics corresponds to the maximal concentration a bacterium can grow for an indefinite amount of time. Inversely, bacterial tolerance corresponds to the maximal duration a bacterium can survive, even at high antibiotics concentrations. Most studies so far focused on bacterial resistance. In a 2014 paper, Fridman et al evolved E. coli strains in conditions of daily administrations of antibiotics for different durations (3h, 5h, 8h; with two replicates lines each) (Fridman et al., 2014) (Table INT12). After evolution, the authors found that all evolved strains improved their survival to antibiotics (when exposed to the same antibiotic concentration and duration as in the evolution

experiment). This improvement was not due to an increase in resistance to antibiotics but to an increase in tolerance through a prolongation in the lag phase duration (Fig INT8a). Such "tolerance by lag" is highly medically relevant, as it is a non-specific mechanism of tolerance that could help bacteria to survive many different drugs or stresses. Moreover, improved tolerance allows bacteria to survive longer and thus gain more time to evolve resistance to antibiotics (as they showed in a recent paper (Levin-Reisman et al., 2017)). Interestingly, the mean lag time of evolved clones was very similar to the duration of antibiotics treatments (Fig INT8c). Whole genome sequencing and complementation assays allowed the authors to discover 3 genes involved in lag lengthening. Two of them were already known in the literature as playing a role in increased persistence. Persistent bacteria designate bacteria that can survive antibiotics due to the resistance of a small fraction of cells within the population (Levin-Reisman et al., 2017). Finally, a very interesting result was that the population mean and variance lag times were increasing together in evolved clones, with the clones evolving in the 8 hours condition having a very broad distribution of single cells' lag time (Fig INT8b). The periodic regime was unpredictable for the bacteria. As predicted from theory, this unpredictable situation leads to the apparition of a diversifying bet-hedging strategy where some cells keep short lag phases, while others have various lag phases, sometimes extremely long. This is advantageous for the organism as this strategy maximizes its survival regardless of whether the stress disappears, remains or increases.

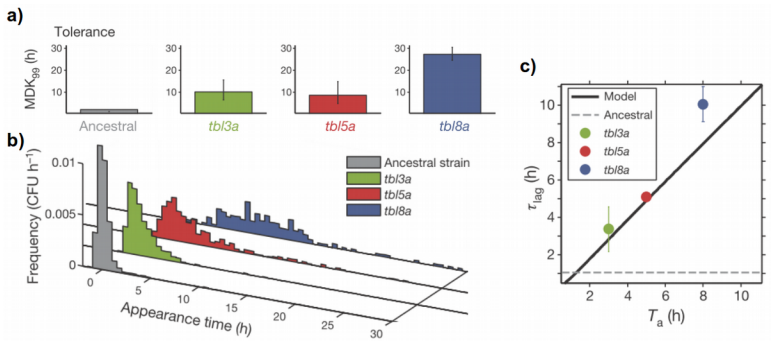


Fig INT8. Fridman et al. study: experimental evolution of tolerance by lag phase duration in E. coli. tbl3a (resp. tbl5a, tbl8a) were the evolved clones from the selection regime with antibiotic duration of 3 (resp. 5, 8) hours. **a)** improved tolerance of evolved clones. **b)** distribution of single cells lag time upon antibiotic treatment. **c)** mean lag-time of evolved clones corresponds to antibiotic duration of the corresponding selection regime. Reference: Fridman *et al.* 2014.

In 2014, Lambert et Kussel used a microfluidic device to grow *E. coli* cells in quickly-changing carbon sources (Lambert and Kussell, 2014) (Table INT12). They observed that, when transferred from glucose to lactose, cells with uninduced lac genes have a lag phase of about 1 hour. However, when cells were fluctuating between glucose and lactose at periods equal or lower to 8 hours, there was no lag phase after the first encounter with lactose (Fig INT9a,b). Thus, there seems to be a memory of lactose that lasts about 4 hours. There are three main genes in the lactose operon: LacZ, a beta-galactosidase, LacY, a permease, and LacA, a transacetylase. These genes are tightly repressed in glucose, and need to be expressed for growth in lactose. By using an over-expression plasmid, the authors found that the memory effect was strongly dependent on LacZ, and LacY, but not on LacA. When bacteria are faced with a carbon or amino acid stress they trigger the "stringent response", which induces the arrest of growth, reduction of translation, downregulation of metabolism and activation of alternative biosynthetic genes. The authors showed that memory of lactose in the 8-hours periodic regime could be simply explained by LacYp dilution over time. They found that LacYp levels above

0.4% of full LacYp induction prevented the lag time due to de-repression of lactose genes. When slightly induced, a positive-feedback loop quickly activates the full pathway. However, there was still a lag time due to induction of the stringent response. LacYp levels above 6.25% prevented the induction of the stringent response and thus could explain the memory effect in the 8-hours periodic regime. Finally, another memory effect on a shorter time scale was found: when cells are pre-grown in glucose, switched to lactose and then back to glucose, they continue to induce their lactose genes during 20 minutes (Fig INT9c). The authors propose that this effect is due to the time needed for the LacI repressor to get fully rid of bound inducer and to rebind to the operon's operator. This study illustrates how simple mechanisms can lead to memory of a carbon stress that have important consequences on fitness (or growth rate) in fluctuating conditions.

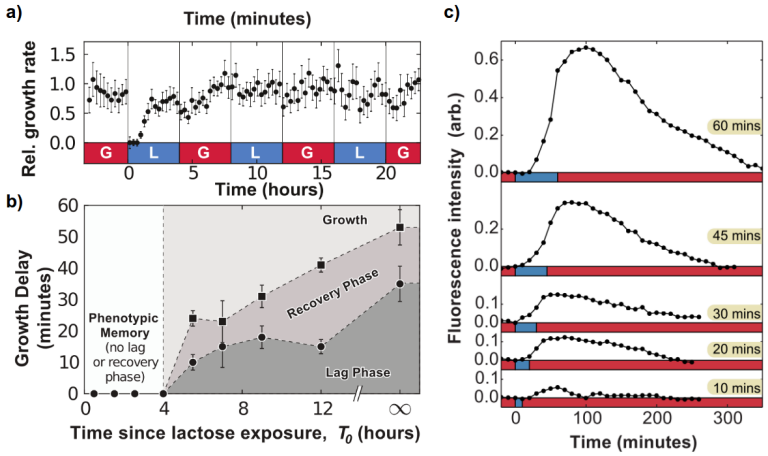


Fig INT9. Lambert and Kussel study: memory of exposure to lactose in *E. coli*. a) the lag in lactose medium appears only at the first cycle in an 8-hours periodic regime. **b)** Periods below 8 hours have no lag phase. Periods above 8 hours have increasing lag and recovery phases. Period = 2 x T_0. **c)** Fluorescence of the LacY-Venus fusion protein. Any pulse of lactose induces an increase in LacYp levels, that peak 40 minutes later. This delay includes ~14 minutes of fluorescent protein maturation, ~20 minutes of residual repressor (LacI) inactivated (bound) by the inducer (lactose), and ~6 minutes needed for the repressors to fully rebind the lac operator sites. Reference: Lambert and Kussel 2014.

Similarly, New *et al.* grew yeast strains in changing carbon conditions (New et al., 2014) (Table INT12). They found that wild isolates had large differences in their ability to gradually transition from glucose to alternative carbon sources such as low glucose (LG), LG + galactose or LG + maltose. Wild isolates' variability of fitness in gradually-changing carbon sources was correlated with their lag time in a sudden transition from glucose to maltose. Interestingly, the wild-type laboratory strain had one of the longest lag phase. Yeast growth on maltose depends on 3 maltose genes. When those genes were over-expressed in the wild-type, the lag phase disappeared. It is often assumed that a short lag phase is adaptive. However, as mentioned previously, this is not necessarily true in changing environments. The authors illustrated this by growing cells in glucose, then inducing them partially in maltose (~ half of the cells as measured by a fluorescent reporter gene), and then switching them back to glucose. Interestingly, cells that induced the maltose pathway had a higher doubling time than uninduced cells (Fig INT10a). This illustrates the cost associated with fast adaptation if the environment changes quickly. The authors then wondered if lag time could quickly evolve. They grew wild-type strains in glucose (20 hours) to maltose (3 days) periodic fluctuations for up to 8 cycles. Strikingly, evolved

clones had almost all a much shorter lag phase than the wild-type for a glucose to maltose transition, which was associated with a fitness benefit in this condition. Interestingly, several evolved clones improved their fitness of transition from glucose to LG + galactose, while the selection was only on maltose. This suggests that those clones became generalists. On the contrary, few clones evolved to become more specialized than the wild-type. Interestingly, there was a strong (R^2 = 0.61) negative correlation between maximal growth rate in glucose and geometric mean growth rate in changing carbon sources among the evolved clones, which reveals the trade-off between carbon specialists and generalists. Whole genome sequencing of evolved clones and complementation assays identified one mutation in the std1 gene, and multiple mutations in the Hxk2 gene (Fig INT10b). Hxk2 is pleiotropic gene involved in glucose sensing, glycolysis, Ras/PKA, Snf1 and nuclear repression. Unexpectedly, one Hxk2 clone became a specialist while others became generalists. This shows that mutations in a single hub gene can allow cells to finely tune the duration of their lag time. Interestingly, generalists hxk2 clones showed a bimodal distribution of maltose genes expression after dozens of hours of growth on maltose + glucose medium. This reveals a type of stochastic sensing, or bet-hedging strategy. In addition, this effect was dependent on the pre-growth medium, even after dozens of hours of growth, indicating that some memory mechanisms allow to maintain a memory of the pre-growth media. Finally, the authors make a model that predicts the winner of a competition between a glucose specialist and a generalist at different periodically fluctuating conditions. They verify their predictions experimentally using two evolved isolates and prove that the dynamics of environmental changes determine the winner of the competition (Fig INT10c).

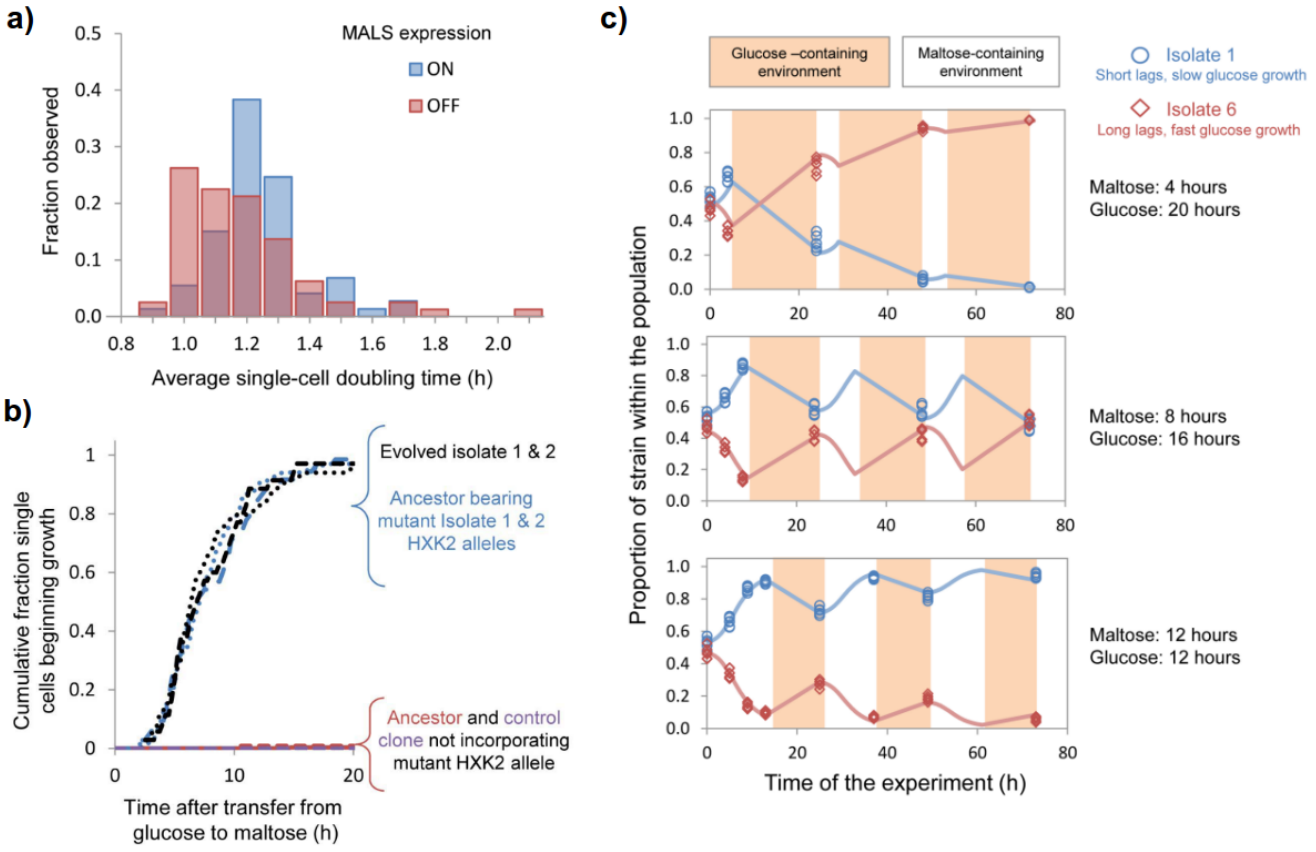


Fig INT10. New *et al.* study: Experimental evolution of lag phase in budding yeast. a) cells grown in glucose, partially induced in maltose and switched back to glucose. Cells that activated the MAL

genes (as shown by a fluorescent reporter of the MalSp) have a higher doubling time. **b)** Yeasts were evolved in a glucose(20h)-maltose(72h) periodic regime for 8 cycles. Whole genome sequencing on final isolates revealed several mutations in the Hxk2 gene. Introducing the Hxk2 allele of isolates 1 or 2 in the ancestral strain fully accounts for the reduced lag phase of evolved isolates. **c)** Two isolates from the evolution experiment evolved different strategies: glucose specialism (isolate 6), or generalism (isolate 1). When faced in competition, those two isolates perform as well in a fluctuation regime similar to the selective environment: glucose(8h)-maltose(16h). Changing the duration of maltose or glucose events favors one of the two strategies. Reference: New *et al.* 2014.

In 2013, Razinkov *et al.* made competitions experiments between two strains: a wild-type strain, and a strain that they previously engineered to have a decreased speed of phenotypic adaptation for the transition from galactose to glucose, by increasing the half-life of the Gal1 transcript (Baumgartner et al., 2011; Bennett et al., 2008; Grilly et al., 2007; Razinkov et al., 2013). Using a homemade microfluidic device, they performed co-culture assays between those two strains at different periods of fluctuations (4h, 8h, 10h) during about 200 generations (Table INT12). They found similar fitness values in steady conditions (constant galactose or glucose), and a very strong competitive advantage for the wild-type in fluctuating conditions (Fig INT11a). Moreover, this advantage was directly proportional to the number of environmental switches and not to the frequency of fluctuations (Fig INT11b). This study proves that a competitive advantage in fluctuating conditions can be directly related to molecular mechanisms that increase or decrease the speed of phenotypic adaptation. Moreover, it shows how periodic fluctuations can amplify the effect on fitness of a prolonged lag phase, that would have been missed with just a single environmental switch. This study also illustrates that fitness in periodically-fluctuating conditions cannot always be predicted from fitness in steady conditions.

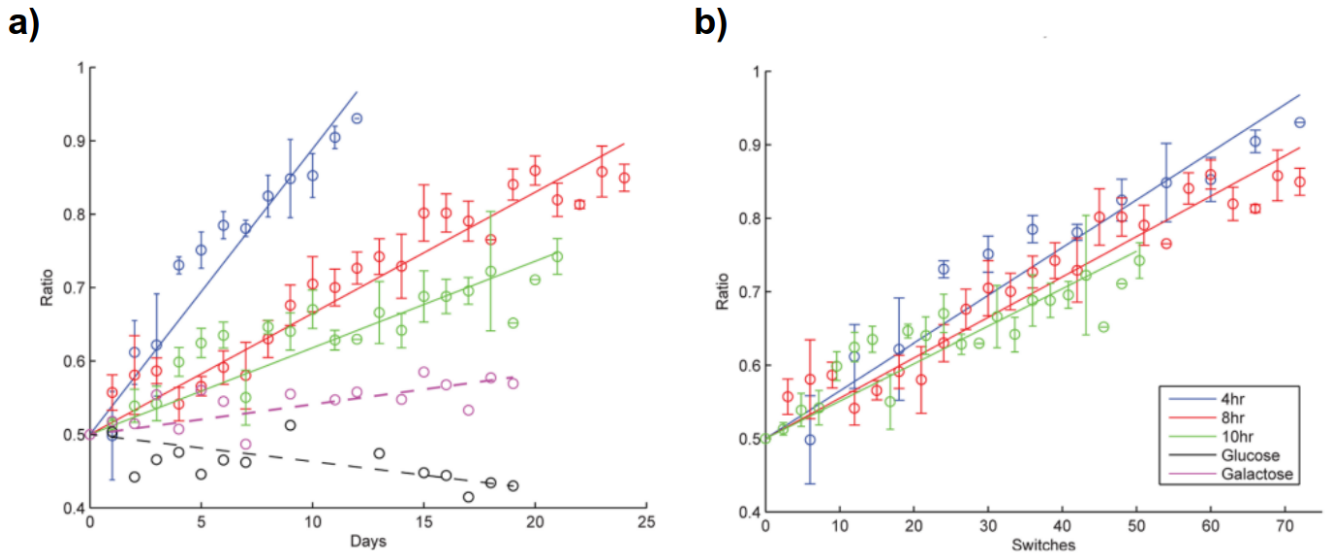


Fig INT11. Razinkov *et al.* study: Relationship between lag time and fitness in periodic conditions in budding yeast. y-axis is the percentage of wild-type cells in a competition against a mutant strain, with an increased galactose to glucose lag time. **a)** The wild-type wins the competition in fluctuating conditions. **b)** the number of galactose to glucose switches explains the competitive advantage of the wild-type strain. Reference: Razinkov *et al.* 2013.

Study	Organism	Fluctu tool	Fluctu type	Periods	Max # of cycles	Max duration	Organism doubling time	Max # of generations	Observed phenomenon
Lambert 2014	*E. coli*	microfluidics	Glucose / lactose	6-180 min (symmetric periods)	6	~18h	~1h	~18	Short term memory Transgenerational memory
Razinkov 2013	*S. cerevisiae*	microfluidics	Glucose/ Galactose	2:2h, 4:4h, 5:5h	75	~30 days	~1h25min (w/o gal.)	~500	Lag amplification
Fridman 2014	*E. coli*	manual	LB / LB + ampicilin	20:3h, 18:5h, 15:8h	10	10 days	25 min	~580	Lag evolution Strategies of adaptation
New 2014	*S. cerevisiae*	manual	Glucose / Maltose	12:12h, 16:8h, 20:4h, 20:72h	8	32 days	~1h15min (w/o mal.)	~120	Lag evolution Strategies of adaptation Lag amplification Long term memory
Stomp 2008	*Pseudanabaena*	manual	Red / Green light	6:6h, 2:2d, 14:14d	5	140 days	~8h45min	~380	Strategies of adaptation

Table INT12: Studies measuring organism fitness during/after growth in binary periodic fluctuations.

C My project: yeast genomics of fitness non-additivity

As we saw above, subjecting micro-organisms to Periodically Fluctuating Conditions (PFC) can teach us a lot about adaptation and evolution. This includes the study of: the impact of lag times and memory effects on fitness in fluctuating environments, the optimal strategies of adaptations, or the ability to predict fitness in changing environments from fitness values in steady conditions. These issues have already been approached by studies performing environmental fluctuations manually or via microfluidics tools, as shown in the previous section. However, throughput is always limited with these methods, preventing large scale screens.

The aim of this thesis was to explore these evolutionary aspects, by adopting a systematic approach where we measured the impact of thousands of yeast genes deletions on fitness in periodically fluctuating conditions. Specifically, we grew the homozygous Yeast Deletion Library in 4 different medium fluctuations (salt, no methionine, sodium metabisulfite, high glucose), with steady controls and 10 different periods each (from 6 to 60 hours). This approach has the advantage to directly points to genes and potential mechanisms that are important in fluctuating conditions. Moreover, screening a large pool of mutants with different phenotypes in different conditions may allow to drive general principles about fitness in fluctuating conditions.

We focused our efforts on the salt fluctuations since phenotypes were the strongest in this condition.

Results showed that predictability of fitness in PFC from fitness in steady conditions was a gene-by-gene question. However, we found that most genes were non-predictable at fast fluctuations, while most genes were predictable at slow fluctuations. Many genes with extremely high fitness in fluctuating conditions were identified at short periods of fluctuations, and two of them were validated by complementation assays. This study opens the door to mechanistic analysis for determining why these genes protect the cell from hyper proliferation in fast PFC.

2 Methods

In this section are methods not described in the submitted publication.

Fluctuation experiment, run 2 and spike-in.
The Methionine and Glucose fluctuation experiments followed the same protocol than the Salt and Sodium Metabisulfite experiments, at the exception that a spike-in control was included. The PCR reaction described in the submitted publication was used to amplify the DNA of two strains (Pcl9 and Ipt1) from single colonies on streaked YPD plates. Then aliquots of these two strains, and a population of amplified barcodes from a Yeast Deletion Library, were mixed at various concentrations. After PCR amplification in the BarSeq experiment, 8 samples had low concentrations: the N condition at day 0 of the Glucose and Methionine experiments. These samples were not included when pooling PCR amplicons and thus were not sequenced. Instead, the 8 spike-in populations were added to the pool of amplicons. After sequencing, spike-in data were analyzed (before normalization) and were then removed from all analysis. The populations that had been removed (N condition at day 0) were computationally replaced. At day 0, all populations experienced only the N medium. Therefore, the counts of mutants in each missing population were replaced by their median counts in 11 other populations. A different set of 11 populations was used to replace each missing population (the other wells in the same row on a 96 well plates), in order to create an artificial variance between replicates. Similarly, 6 populations had very low replicability at day 0. Those were replaced by taking the median count per mutant in 7 wells (in the same column on a 96 well plates).

Calculation of Fold Changes and Principal Components.
Analysis were performed using R version 3.4.0 (2017-04-21). Fold changes were computed using the function DESeq (argument fitType = “local”) and the function “results” (default paramaters) from the package DESeq2 (version 1.14.1). Principal Components (PC) were computed using the “prcomp” function (default parameters).

Estimation of cell number and doubling time
Computing absolute fitness values from FACS data was possible since the FACS mixed several times vigorously each well, before aspiration at a constant speed. The rate of sampling was 1µl per second. Since the time needed to acquire 10,000 cells was recorded in the data files, it is possible to estimate the concentration of cells in the population. I estimated the concentration of cells in each sample/well/population (c) by this formula:
$c = ((N/t) / (dil1 . Dil2))$.
N is the number of events (cells) acquired for a sample on the FACS (10,000 for most samples). t is the time of acquisition of the sample on the FACS. dil1 is the dilution of the sample before fixation, equal to 0.4. dil2 is a second dilution of the sample before acquisition, that is variable. It was set between 0.055 and 0.2 for flow-cytometry run 1 to 4. For run 5 (data of Fig IV-B.7), dil2 was set at 0.02, to keep low speed of samples acquisition on the FACS. The unit of c is s^{-1}. It is equivalent to l^{-1} because the speed of aspiration of the FACS was constant, with unit $l.s^{-1}$.

I estimated the theoretical volume of cultures that would have been produced if there was no dilutions,

by this formula:
$tv[i] = tv[i - 1] + v0 . (v1/v2)^{(i-1)}$, with $tv[1] = v1$
i is the number of fluctuations that occurred (from 1 to 25). v1 is the volume of cultures, equal to 220 µl. v2 is the volume of culture that is kept at each fluctuation, equal to 130 µl. v0 is the volume that is discarded at each fluctuation, equal to 90 µl.

Cell number (cn) in each population was estimated by this formula:
$cn = tv[i] . c . p_t$.
p_t is the percentage of cells of type t. t can be either mutant, wild-type or all cells. Classification of cells was based on a threshold of fluorescence levels, as described in the submitted publication.

A linear model was fit on: log2(cn) ~ time, with time in hours. The doubling time was then computed as the inverse of the slope coefficient of the model.

Growth media
BarSeq assay. For the Sodium metabisulfite and Glucose experiments, N medium was made as described in the submitted publication. Medium S_{Met} (Sodium Metabisulfite) was made by adding 10 ml/L Sodium Metabisulfite 78.37 mM to medium N (final concentration of 0.7837 mM). Medium S_{Glu} (Glucose) was prepared as the N medium, except that 50 g of glucose were added per liter (final concentration of 5%) instead of 20g for the N medium (final concentration of 2%). Medium S_{Met} (Methionine) was prepared as the N medium, except that a mix of amino-acids without methionine was used. N medium for the methionine experiment was made by using the same mix of amino-acids lacking methionine, and supplementing the medium with 149 mg of methionine per liter of medium.
Flow cytometry pleiotropy experiments. The NaCl 0.4 M medium was made by adding 80 ml of NaCl 5M per liter of N medium. The Sorbitol 0.4 M medium was made by adding 73 g of sorbitol per liter of N medium. The raffinose 1.9% medium was made by making a N medium with only 1g of glucose per liter, and 19g of raffinose per liter. The ethanol 5% medium was made by adding 50 ml of ethanol 100% per liter of N medium. The KCl 0.2 M medium was made by adding 14.9 g of KCl per liter of N medium. The LiCl 0.2 M medium was made by adding 8.5 g of LiCl per liter of N medium.

3 Results

I Submitted publication

Genomics of cellular proliferation under periodic stress

Jérôme Salignon * , Magali Richard * , Etienne Fulcrand, Hélène Duplus-Bottin and Gaël Yvert
*) these authors contributed equally to this work

Laboratory of Biology and Modeling of the Cell, Ecole Normale Supérieure de Lyon, CNRS, Université Claude Bernard de Lyon, Université de Lyon, 69007 Lyon; France.

In revision at Molecular Systems Biology and posted on Biorxiv on April 20, 2017.
doi: https://doi.org/10.1101/129163

II Summary of the project

Figure II.1 illustrates the 2-steps screening approach that we adopted for this study: a genomic screen followed by an individual mutant assay to confirm results. For both experiments, cultures of yeast in binary periodic fluctuations were automated on a liquid handling platform. The genomic screen was a high-throughput assay allowing to assess simultaneously the fitness of thousands of mutants from hundreds of different populations. This was achieved by the BarSeq technique (Fig II.2), in which mutants' barcodes are amplified by PCR and sequenced. Magali Richard designed a library of 384 indexing primers that allowed multiplexing by amplifying mutants from different populations with different primers. This way, amplicons from PCR could be pooled and sent to sequencing together. After parsing of reads, the abundance of each mutant in each population could be quantified. The secondary screen experiment was similar to the primary screen, but precision was much higher (Fig II.1). Indeed, in this experiment there were only two strains in the pool instead of ~3500. This prevented potential interactions; such as exchange of metabolites. Moreover, cells were directly counted by a flow-cytometer, instead of gDNA extraction, PCR and sequencing that all introduce biases. The purpose of the secondary screen was to both validate the effect individual mutants of interests and to determine the success rate of validation, in order to infer how reliable the genomics analysis of our BarSeq experiment was.

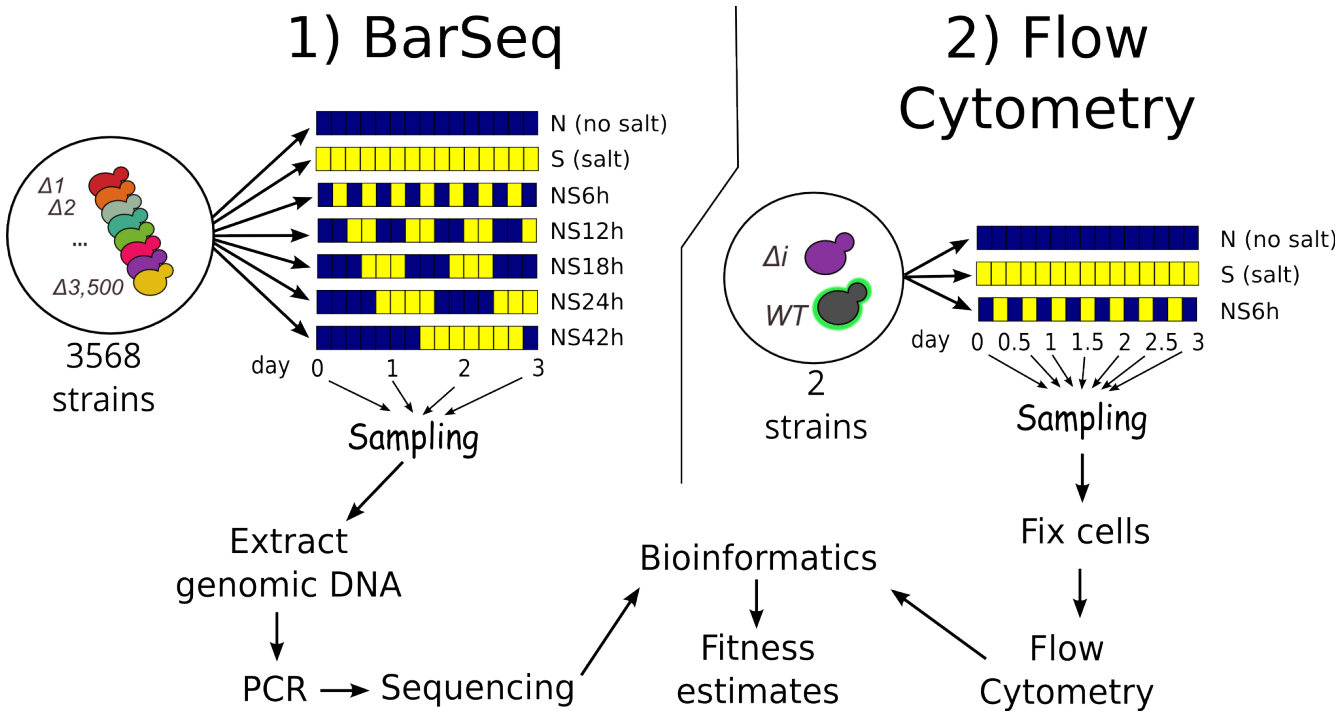


Fig II.1. Experimental design of the project

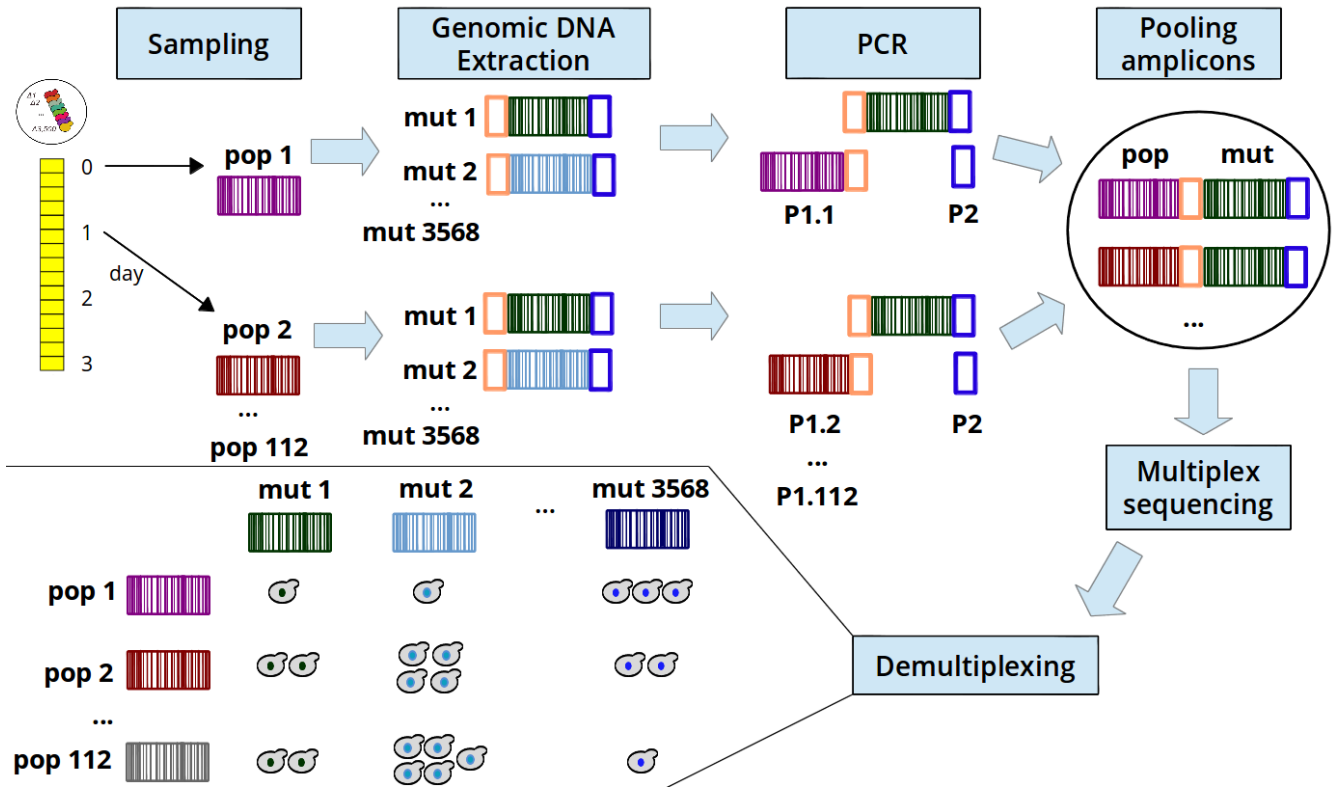


Fig II.2. BarSeq: multiplexing of Barcode Sequencing

Finally, we can make a parallel between the issue of determining if a system is linear and if fitness is predictable in fluctuating conditions from fitness in steady conditions. Indeed, as shown in Fig II.3, we can see that if the fitness of a mutant in periodic stress is equal to the (weighted geometric) average of its fitness in steady conditions, then it probably satisfies the conditions of homogeneity and additivity. Therefore, mutants that have (un)predictable fitness will be referred to as (non)linear mutants. Importantly, there is a slight discrepancy in the terminology used in the submitted publication and in this thesis. In the submitted publication, we used the term homogeneity, instead of linearity, to describe the level of predictability of fitness in fluctuating conditions from fitness in steady conditions.

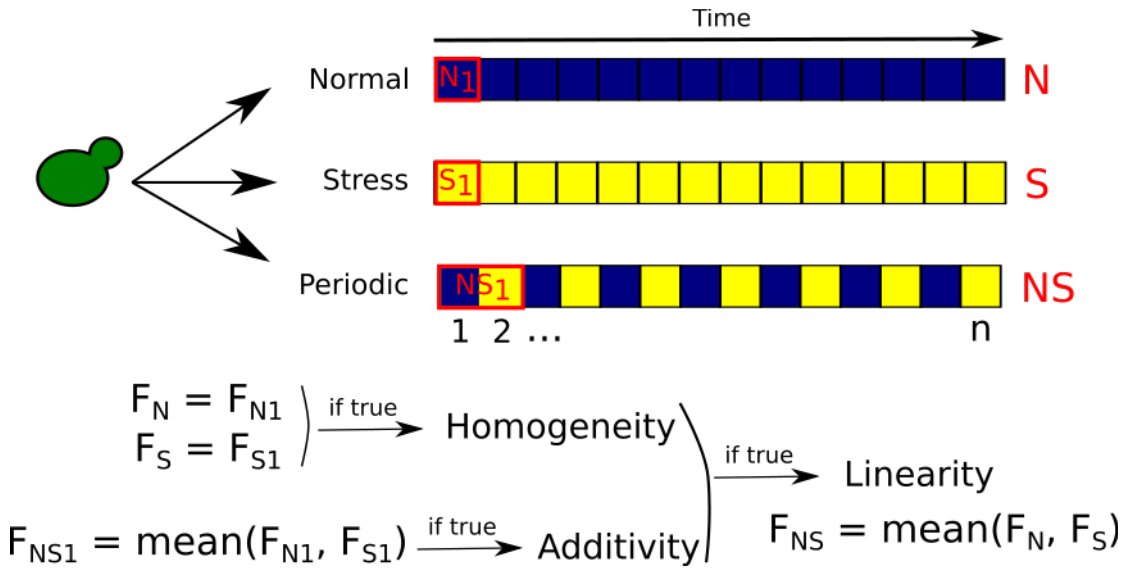


Fig II.3. Testing for linearity of fitness in fluctuating conditions. F stands for Fitness. Non-additivity indicates an environment x environment interaction. Inhomogeneity indicates a time x environment interaction.

III Four genomic screens

A Experiment

i Beginning of the project

Magali Richard arrived at the laboratory 8 months before me, and started this project, together with another project based on wild yeast isolates. After my arrival, I took over the study of deletion mutants, while she continued on the study of wild strains. The experimental designs of the projects were very similar. In her case, she screened a library of ~25 wild isolates while I screened a pool of ~4000 yeast null mutants. During her first year in the laboratory she set up different protocols that I used afterwards. She designed a library of 384 primers for the PCR step of the BarSeq experiment (see part 3-III-B-ii), and optimized conditions for the PCR. She also set up an automated high-throughput robotic protocol for extracting DNA genomic. Finally, she wrote a first version of the script for automating cultures in fluctuating conditions. She performed a BarSeq-based analysis of the wild strains and then tested pairwise strains comparisons by flow-cytometry. Unfortunately, strain-to-strain differences were not validated by this secondary assay. If she had validated fitness differences in fluctuations between two wild isolates strains, she would have then mapped corresponding QTL, through the technique of Bulk Segregant Analysis. Magali Richard then worked on another project.

ii Description of the experiment

The experimental design of plates for this experiment is illustrated in Figure III-A.1. Briefly, each well contained the full Yeast Deletion Library. 96-well plates were split in two, horizontally, with each part being used for a different type of medium fluctuation. Each column contained replicates of the same condition. During three days, binary periodic fluctuations were made between a non-stressful (N medium) standard medium (SD all), and a stressful medium (S medium) in which wild-type yeast growth is suboptimal. For each type of stress, there was a total of 12 conditions: two steady conditions (N and S) and 10 different periods of fluctuations (NS6-60). Finally, samples were frozen once a day, for later genomic DNA extractions. Tested stressful media were NaCl 0.2 M (S_{Salt}) and Sodium Metabisulfite (Na2S2O5) 0.8 mM (S_{Sul}) in the first run (run 1), and glucose 5% (S_{Glu}) and methionine 0 mM (S_{Met}) in the second run (run 2). These concentrations were chosen to be low-enough to allow yeast cells to grow well. Indeed, it was important for us not to lose populations growing in the stressful condition for 3 days, since they were later necessary for computing the degree of fitness linearity at all periods. However, stress should be tough enough in order to observe a phenotype for most mutants. For all experiments, we decided that I would use the same concentration as Magali Richard used in her experiment. The reasoning was that we might be able to compare our results later on. However, she used a concentration of 0.8 M for her salt experiment, which revealed to be too strong and eradicated all populations in the salt steady condition. Thus, for the salt experiment, I tested growth of a wild-type yeast in different salt concentrations. I chose the concentration of 0.2 M, in which cells showed a small but detectable increase in doubling time after 5 hours of exponential growth (Fig III-A.2).

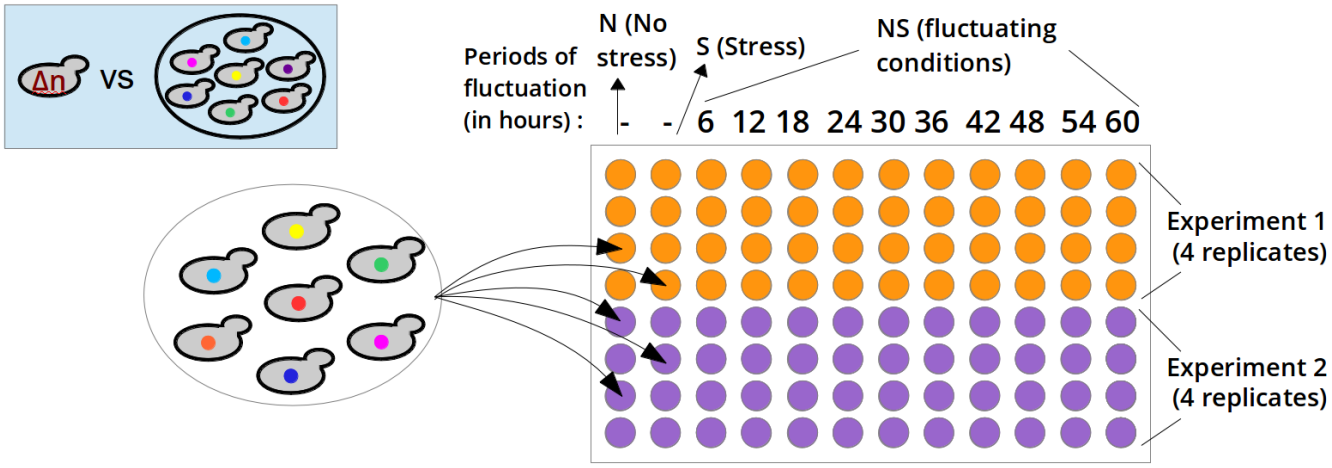


Fig III-A.1. Design of one 96-well plate for the primary screen fluctuation experiment

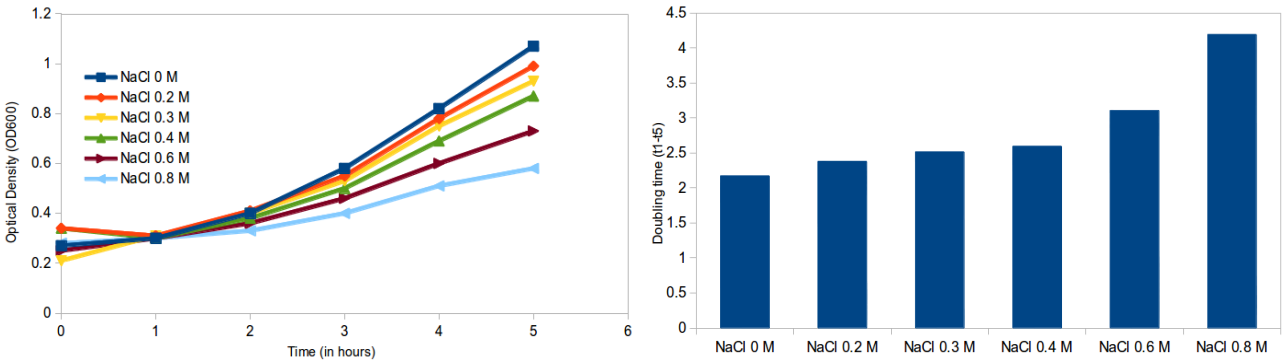


Fig III-A.2. Optical density (left) and doubling time (right) of a wild-type yeast strain at various salt concentrations

iii Growth of populations during fluctuation experiments

The fluctuation experiment lasted only 3 days, however, I needed 4 months to complete it correctly twice. This delay was due to several aborted experiments after crashes of the robot. Indeed, the protocol required a lot of robustness to perform precisely interventions multiple times during few days and nights. Improvements to the protocol included re-writing the code in a loop fashion; preparing 96-well plates with media for the future fluctuations before the night (and not during the fluctuations); avoiding trashing tips or plates; and switching from 4 to 5 autonomous fluctuations per night.

During the experiment, Optical Density (OD) was measured every second fluctuation by a sunrise plate reader (Tecan) on the robotic platform. This allowed to observe how populations grew over time in the different conditions. Figure III-A.3 to III-A.6 show OD measurements in the different media fluctuations. We can see on those Figures that populations sizes drop every day. This occurred when I collected and froze samples for later gDNA extraction. As expected, most of the fluctuating conditions had population sizes that were in-between the two associated steady conditions (data not shown). The

no-stress steady condition (N) should have similar population sizes in these Figures. However, we observed that in the Methionine experiment, growth rate in N was lower than in the other experiments. There were almost no growth differences between the N and S_{Met} populations. This result suggested that we should not expect to observe many strains with an important growth rate difference between the N and S_{Met} conditions after data analysis, which could limit potential genomics analysis. On the contrary, the biggest difference in fitness between the two steady conditions was observed in the NaCl experiment. Thus, we could expect that many mutants would have a detectable phenotype. Finally, the Glucose experiment had a higher variability in N than other experiments, which could potentially decrease the power to detect mutants with interesting phenotypes.

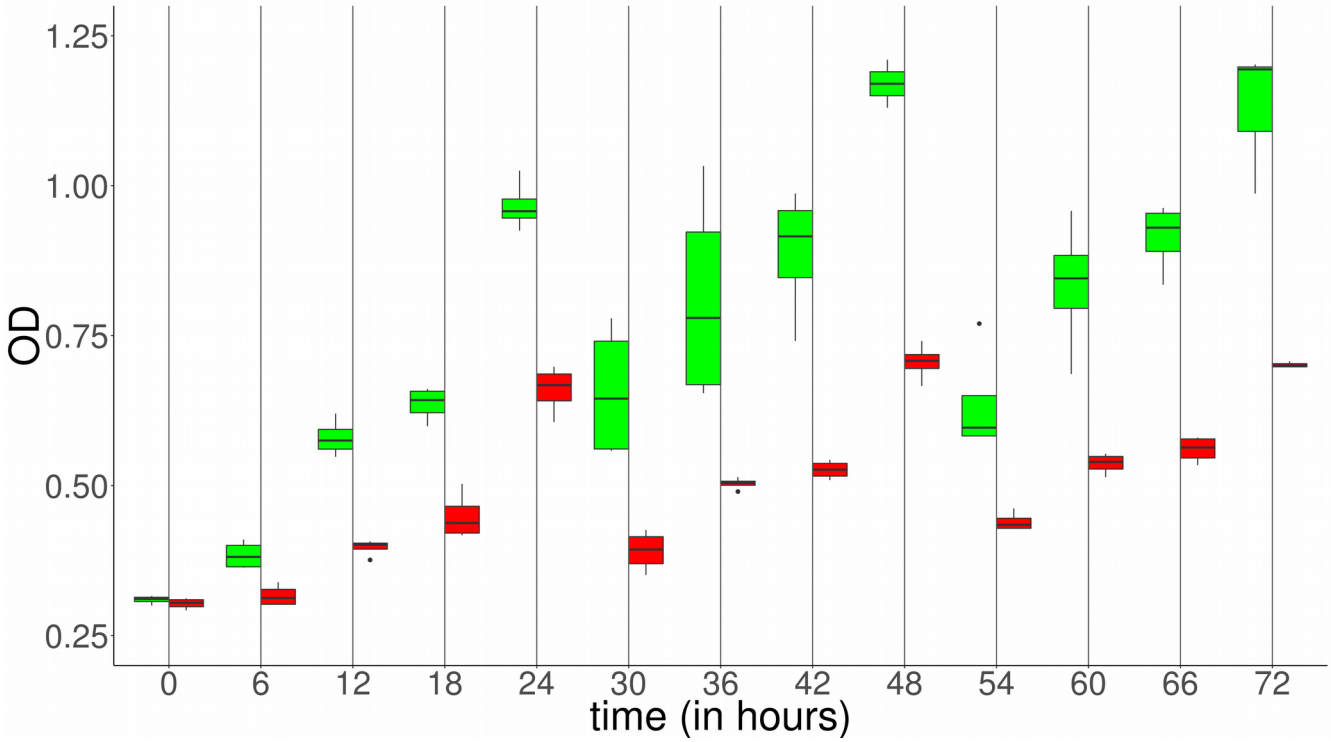


Fig III-A.3. Optical density measurements of populations in the salt experiment. Green: N. Red: S_{salt}.

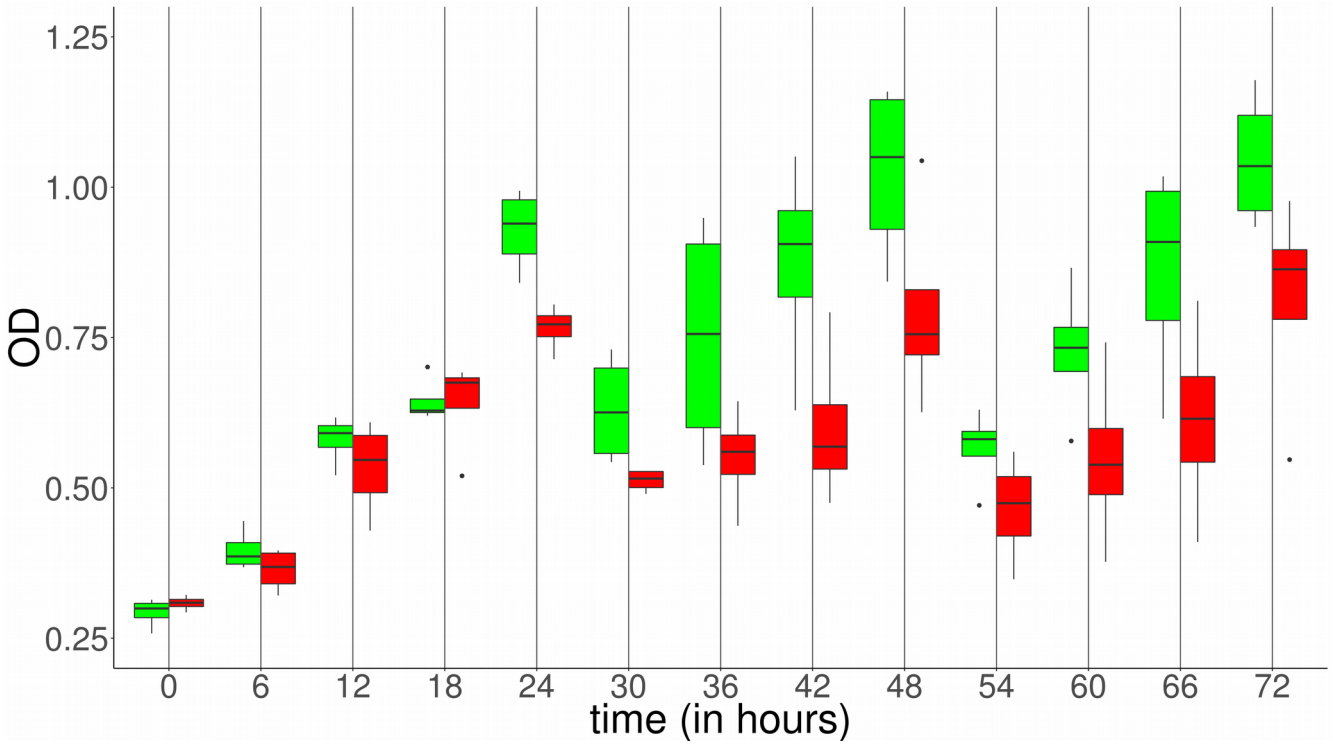


Fig III-A.4. Optical density measurements of populations in the Sodium Metabisulfite experiment. Green: N. Red: S_{sul}.

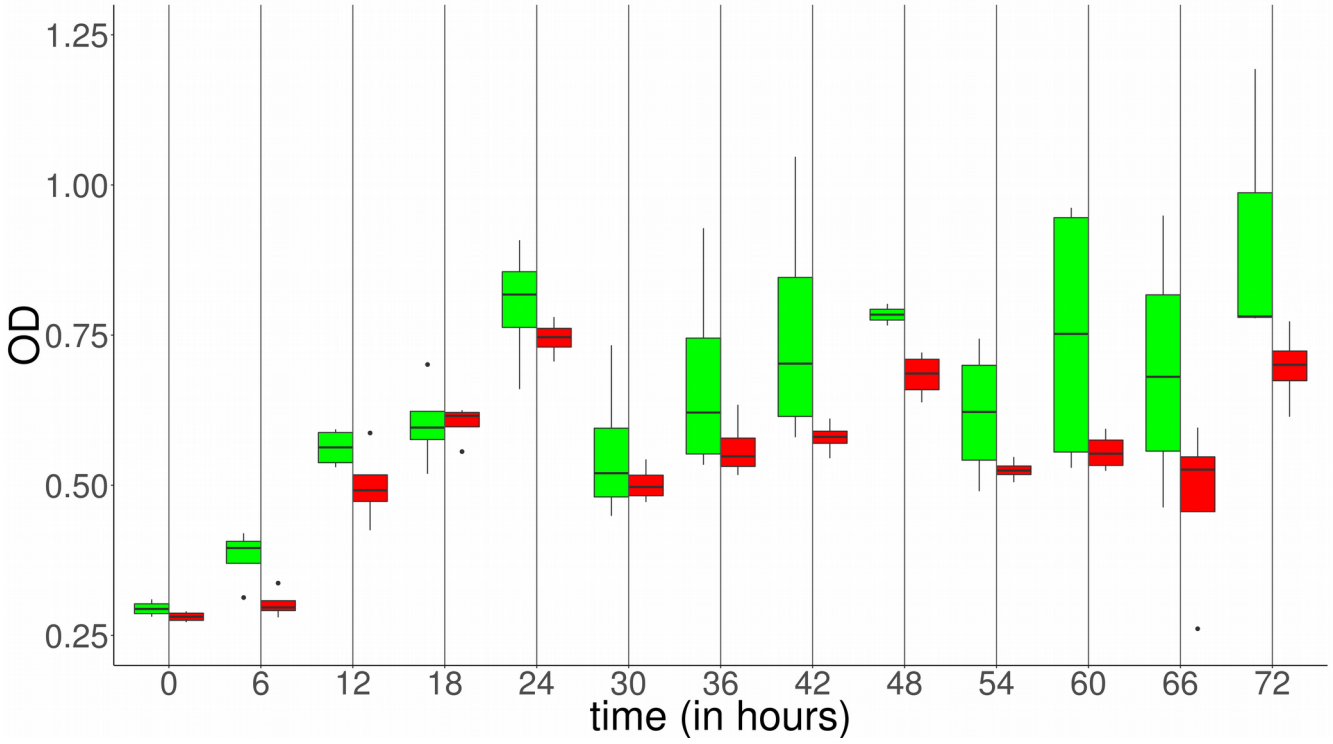


Fig III-A.5. Optical density measurements of populations in the Glucose experiment. Green: N. Red:

S_{glu}.

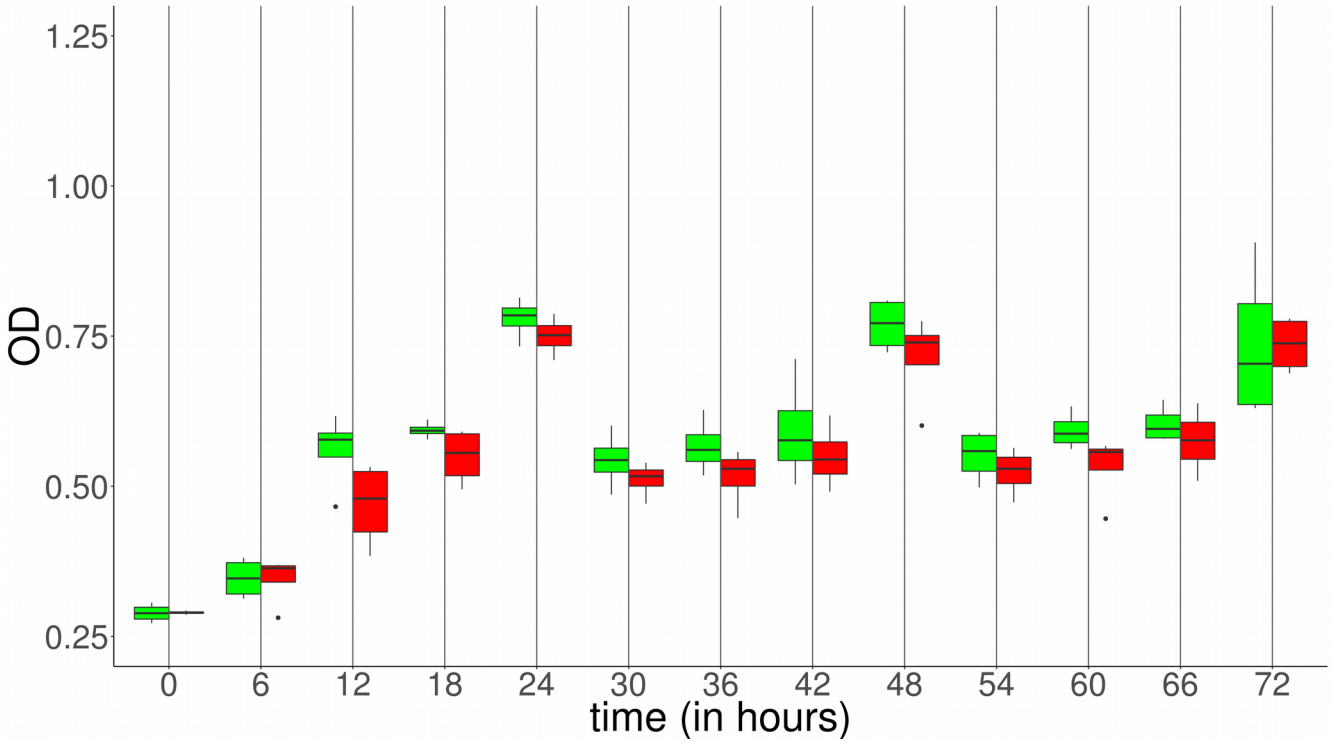


Fig III-A.6. Optical density measurements of populations in the Methionine experiment. Green: N. Red: S_{met}.

iv Solving an apparent issue of PCR amplification of the barcodes

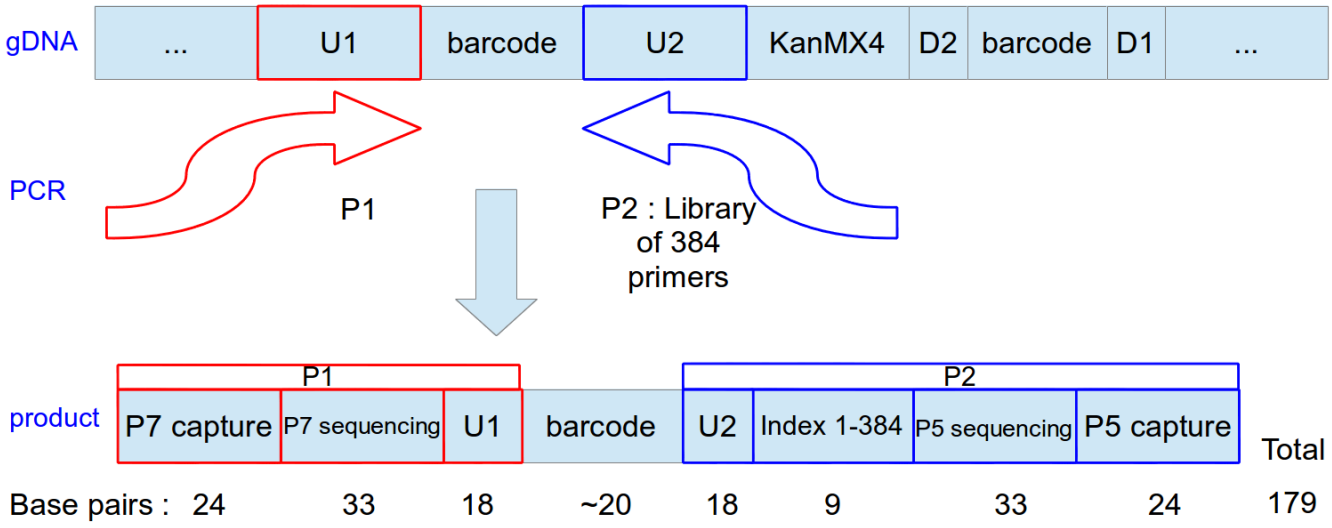


Fig III-A.7. Design of the PCR used to amplify barcodes during the BarSeq experiment.

The design of the PCR used to amplify barcodes in the different populations is shown in Fig III-A.7. Our amplicons contained two barcodes: the mutant barcode and the population barcode. For simplicity, in the following text we will refer to mutants barcodes as “barcodes” and to populations barcodes as “indexes”. Capture sequences allows to hybridize single stranded (denatured) DNA sequences to oligos of a Illumina flow cell. Bridge amplification is a technique were capture sequence primers are used to amplify each sequence. After that all sequences are bound by P7 capture sites, and P5 capture sequences are used for sequencing. In this PCR design, the size of the index, U2 and barcode is about 50 base pairs. Thus, I sequenced my PCR amplicons in single reads 60 base pairs.

I expected that my PCR would yield a single band at ~180 bp. However, I unexpectedly observed two bands: one at ~200 bp, and one at ~180 bp (Fig III-A.8). Before purifying the low band, I wanted to assess the nature of the high band, in order to determine if I should include or exclude it, to ensure high-quality of samples. I have performed this PCR repeatedly for 3 months, during which I tested different conditions to get rid of the second band. However, after changing reagents and materials, two bands were still present. Samples sent to sanger sequencing revealed that both bands contained barcodes. This also allowed to confirm that barcodes of different strains could be identified. I finally found in a publication that it is normal to observe a second band when amplifying barcodes of the Yeast Deletion Library (Pierce et al., 2007). This second band originates from non-specific hybridizations between amplicons from different barcodes during the annealing step of the PCR (Fig III-A.9). These structures migrate slower due to their less compact form, and thus form another higher band. According to the authors, this theory is supported by the fact that boiling amplicons briefly removes the higher band. Indeed, boiling results in only non-specific hybrids and thus enrich for the higher band. Fortunately, I observed the same effect after boiling an amplicon, which confirms that my problem is indeed due to non-specific hybridizations. Since non-specific hybrids also contain relevant barcodes, both bands were cut, purified and sent to sequencing.

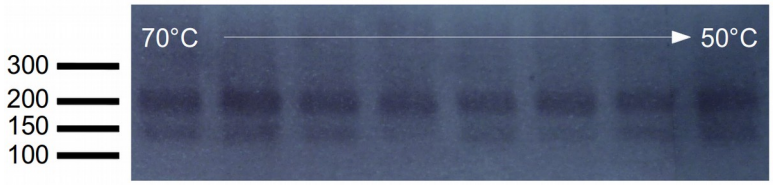


Fig III-A.8. PCR with a gradient of temperatures from 70°C to 50°C and 2.5 mM of MgCl2. Amplified DNA is a genomic DNA sample from the fluctuation experiment

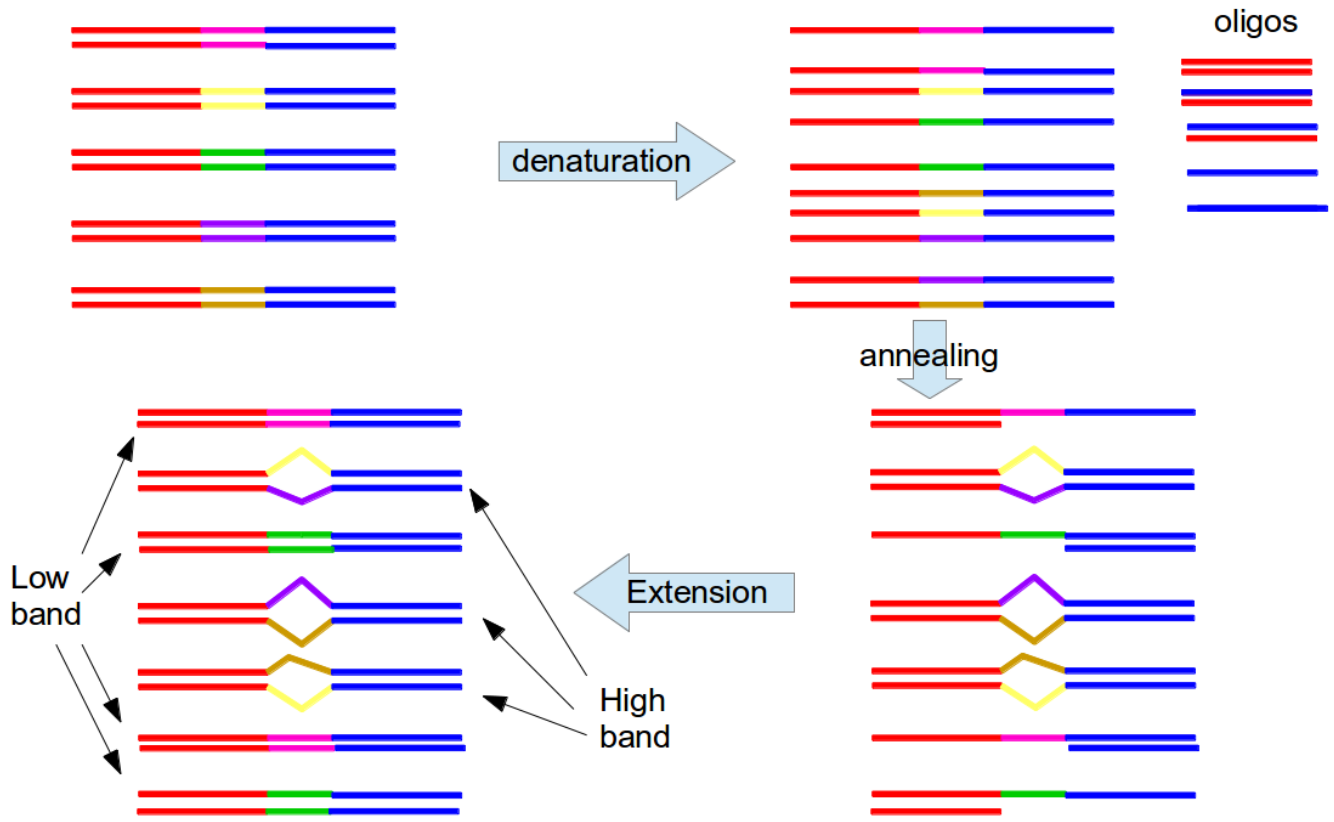


Fig III-A.9. Two bands in PCR with pooled barcodes are expected because of heteroduplexes and homoduplexes.

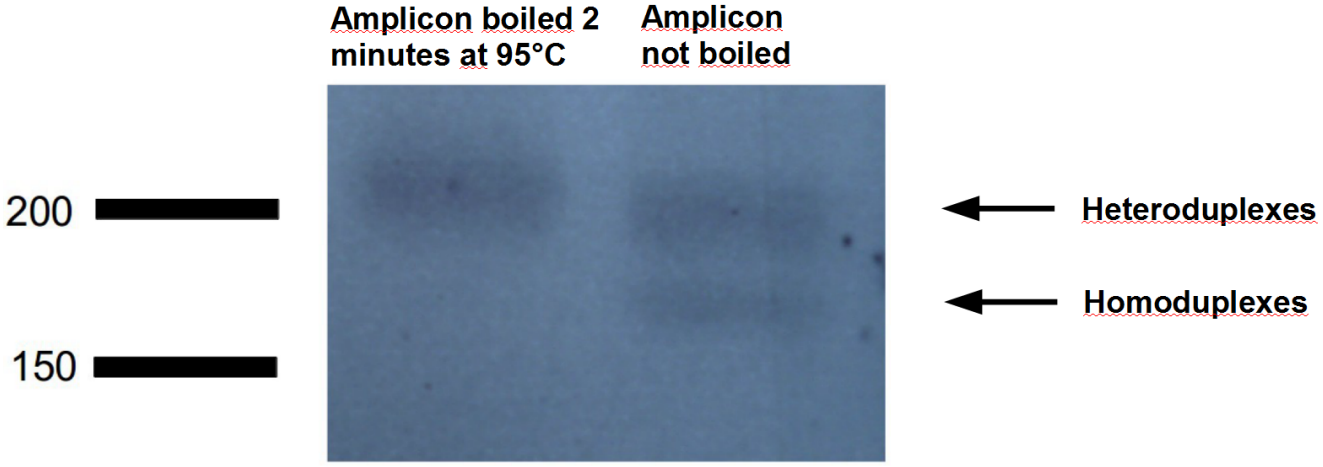


Fig III-A.10. Boiling amplicons before running the gel removes the low band of homoduplexes.

Since my library size was 384 primers, and there were 192 conditions per medium fluctuation (4 replicates, 4 days, 12 frequencies), I pooled samples from two media fluctuations, concentrated DNA, and sent pooled amplicons to sequencing together on a single lane of a Illumina sequencing HiSeq 2500. I decided to send to sequencing samples that originated from the same run; that is NaCl and

Sodium Metabisulfite (run 1 on lane 1). I applied the same design on another sequencing lane, pooling the amplicons of the Glucose and of the Methionine experiment (run 2 on lane 2).

B First, sequencing results

i Quality of sequencing

There were about 145 million reads per Lane. The Phred quality score (orQ-scoree) is calculated with this formula: -10log10(probability that the base is wrong). For instance, a quality score of 20 (resp. 30, 40) means that there is one error in 100 (resp. 1.000, 10.000) base calls. Importantly, in my data, more than 97% of reads had a quality score higher than 30. Thus, sequencing was of high quality.

I confirmed the quality of sequencing data using the software fastQC (Andrews, 2010). In the following Figures is shown a fastQC analysis of one out of 26 fasta files. Other files were qualitatively similar. In Figure III-B.1, we can see the distribution of the Q-score of all sequences in this fasta file: with a mean quality score of ~39. As we can see on Figure III-B.2, the quality score was high for all bases. It was a bit lower for the bases of the index, however, it is usual that the first bases have a lower quality than others (Krueger et al., 2011). On Figure III-B.3, we can see that all sequences contained 9 base pairs varying indexes, followed by the conserved U2 sequence, the ~20 bp varying barcodes and finally half of the conserved U1 sequence.

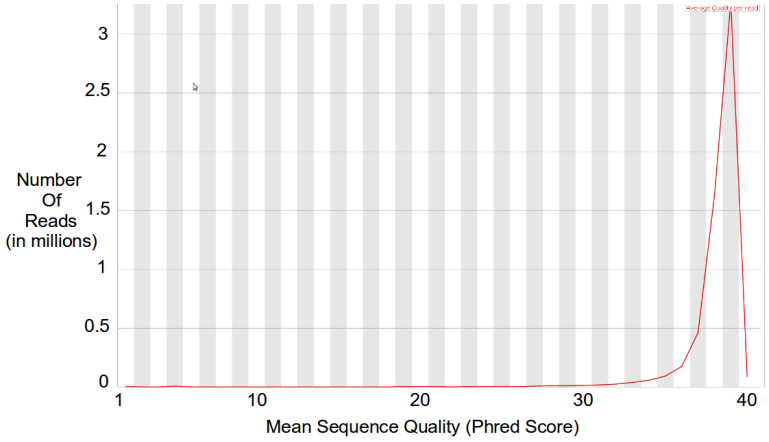


Fig III-B.1. FastQC – Phred quality score distribution over all sequences.

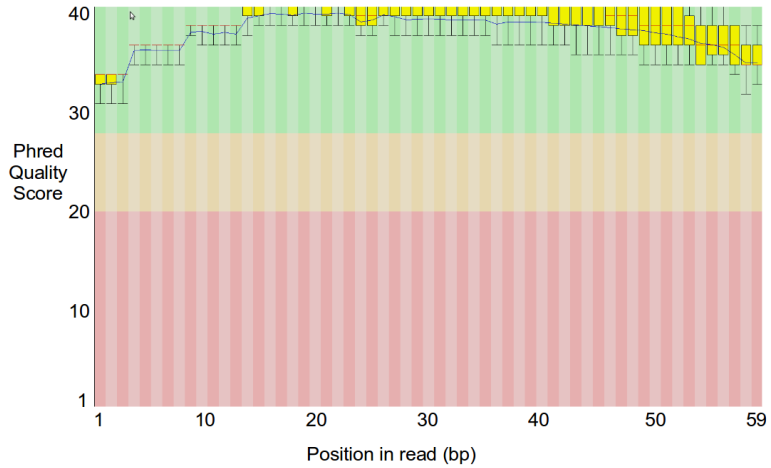


Fig III-B.2. FastQC – Phred quality score distribution over all bases.

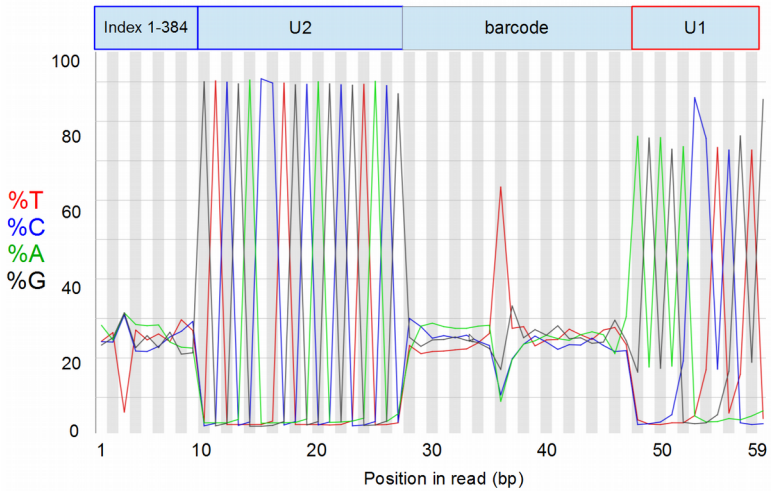


Fig III-B.3. FastQC – Per base sequence content.

ii Reads parsing

Parsing of reads consisted in determining both the index and the barcode in each read. Parsing on raw reads resulted in identification of 57% (resp. 86%) of barcodes (resp. indexes) (Table III-B.4). It is known that PCR and sequencing can introduce errors in DNA sequences. I tried two strategies to correct for these errors: using the distance of Levenshtein (Levenshtein, 1966), or using a code of Hamming (Hamming, 1950).

The distance of Levenshtein indicates how many changes are needed to transform one sequence in another one. One change can be an insertion, a deletion or a substitution, of one letter. The Leveinshtein Distance (LD) can be used to identify the barcode/index that is the closest (that match) to an erroneous reads. A maximal LD can be set, that indicates how many changes are allowed. If a read do not match any other barcodes/indexes at the given maximal LD, then it is discarded. Otherwise its sequence will be corrected for the matched barcode/index. One potential issue when using the LD to correct errors in

reads (Leveinshtein correction) is to match error-containing reads to a wrong barcode. For instance, let's consider two barcodes with a LD of 3 to each other. If a read for one of those two barcodes have two errors, then it might erroneously match to the other barcode. For this reason, to use Leveinshtein correction, it is preferable if barcodes in the library have high LD one from the others.

Therefore, in order to determine if I could use the correction of Leveinshtein, I computed the LD of all barcodes to all other barcodes in each library. And for each mutant I kept the minimal value (minimal LD). We can see on Fig III-B.4 that minimal LD values are higher for the barcode library (Fig III-B.4A) than for the indexes library (Fig III-B.4B). This difference is mainly explained by the shorter length of DNA sequences in the indexes library (9 base pairs) than in the barcode library (~ 20 base pairs). Using the Leveinshtein correction for correcting errors in the index library is not possible since minimal LD are too low (either 2 or 3).

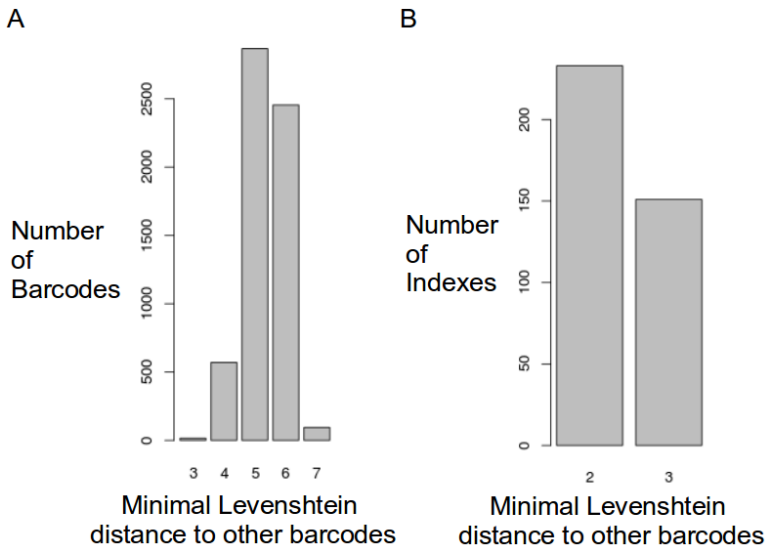


Fig III-B.4. Distribution of minimal Levenshtein distances for **A** the library of mutants' barcodes and **B** the library of indexes.

In contrast, it is possible to use the Leveinshtein correction for the barcode library. The authors in Robinson *et al.* chose to a use a maximal LD of 2 (Robinson et al., 2014). In this case, if one read has two errors, it may erroneously match to another barcode, since few barcodes have a minimal LD of 3. However, in most cases it would not be an issue. I found that the number of corrected reads was modest when using a maximal LD of 1 or 2 (5%, Table III-B.5). Therefore, in order to be conservative, I chose to use a maximal LD of 1. This allowed to rescue 14% of the reads.

Since it was not possible to use the Leveinshtein distance to correct for errors in the indexes library, I tried to use a Hamming correction. Each primer from the indexes library, has been specifically designed by Magali Richard to contain a code of Hamming. A hamming code is a type of error correcting code, were parity bits are inserted within sequences, every 2^n letter (Fig III-B.6). The specific pattern that those parity bits cover in the sequence make it possible to detect errors, but also sometimes to correct them. For instance, if there is only one parity bit that is false, then the error is on the parity bit itself. A recent paper has shown how to efficiently implement quaternary Hamming codes

when designing libraries of oligos (Bystrykh, 2012). Specifically, Magali Richard designed a quaternary Hamming(9,5) code. This means that there are 5 data bits and 4 parity bits. This design allows to create up to 1024 sequences that can be error corrected. She implemented a R script code to correct reads using a Hamming decoding algorithm. Applying her code to my data revealed only a modest rescue of reads (2% as compared with no corrections) (Table III-B.5B). I kept this correction since the error correction system is robust and is not likely to introduce supplementary biases.

A

	Robinson 2014 *	Parsed barcodes °
No correction	62.6%	56.75 %
Maximal Levenshtein Distance of 1	84.6%	71.15 %
Maximal Levenshtein Distance of 2	93.3%	76.38 %

B

	Parsed indexes
No correction	85.52 %
H4(9,5)	87.48 %
Maximal Levenshtein Distance of 1	87.91 %
Maximal Levenshtein Distance of 2	89.54 %

Table III-B.5. Percentage of barcodes parsed using different approaches. The red circle indicates the final strategy that was adopted. **A** mutants' barcodes. * mean parsing on quality filtered reads. ° mean parsing on raw reads. **B** populations indexes.

Bit position		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Encoded data bits		**p1**	**p2**	**d1**	**p4**	**d2**	**d3**	**d4**	**p8**	**d5**	**d6**	**d7**	**d8**	**d9**	**d10**	**d11**	**p16**	**d12**	**d13**	**d14**	**d15**	
Parity bit coverage	**p1**	X		X		X		X		X		X		X		X		X		X		
	p2		X	X			X	X			X	X			X	X			X	X		...
	p4				X	X	X	X					X	X	X	X					X	
	p8								X	X	X	X	X	X	X	X						
	p16																X	X	X	X	X	

Fig III-B.6. Parity bits in a Hamming code. Source: "Hamming code." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. 17 May 2017. Web. 23 May 2017., en.wikipedia.org/wiki/Hamming_code

Finally, after parsing, I obtained a matrix of 6004 mutants by 768 populations. Many of those mutants have zero counts in all conditions and were not considered further (see below).

iii Quality controls

A Spike-in: reliable quantification of mutants in a large range of counts

In order to determine the accuracy of the BarSeq assay for quantifying changes in DNA

concentrations, I designed and used spike-in controls. It consisted in doing a PCR on specific strains at known concentrations, and to include those samples with the other samples sent to sequencing. Then, after sequencing and parsing of reads, I compared known relative concentrations of those controls, with the number of reads I observed in my matrix of counts.

Specifically, I performed a PCR in 8 wells with increasing concentrations of 2 known strains plus a constant concentration of the pool of strains. The genomic DNA concentration was low in the 8 wells of the N condition of the Sodium Metabisulfite and the Glucose experiment (run 2) at day 0. Thus, before pooling all samples for sequencing I replaced those wells by the spike-in controls. On Fig III-B.7, we can see that there is a good precision to detect subtle changes in DNA concentration for individual strains over a wide range of concentrations (in a log2 scale). This comforted us on the validity of BarSeq to quantify relative mutant abundances.

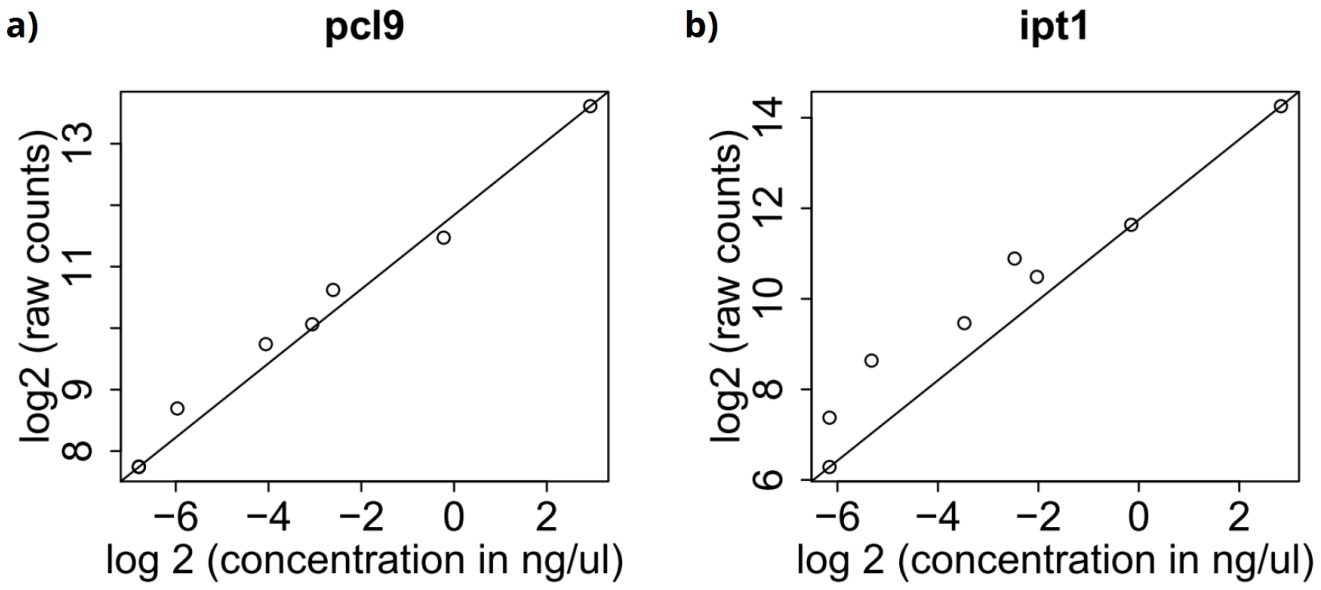


Fig III-B.7. Results of the spike-in control experiment.

B Filtering data

I first removed populations and mutants with very low number of counts. Specifically, I removed 2057 mutants that had less than 100 counts across all conditions. And I removed 3 populations that had less than 500 counts in total. Then, I individually checked distributions of counts in every population, as shown in Fig III-B.8. I removed 97 populations (out of 765) were distributions of counts were looking bad as compared to their replicates.

Then, I removed the wells that contained the spike-in data (Lane 2, day 0, N condition) and replaced

them by the median of 11 other wells at day zero (see methods). Indeed, at the beginning of the experiment, all populations have experienced the same conditions (pre-growth overnight, followed by 6 hours of growth in no stress medium). However, as a result, those wells' variance is artificial at day zero, and the samples are no more paired between day 0 and day 3.

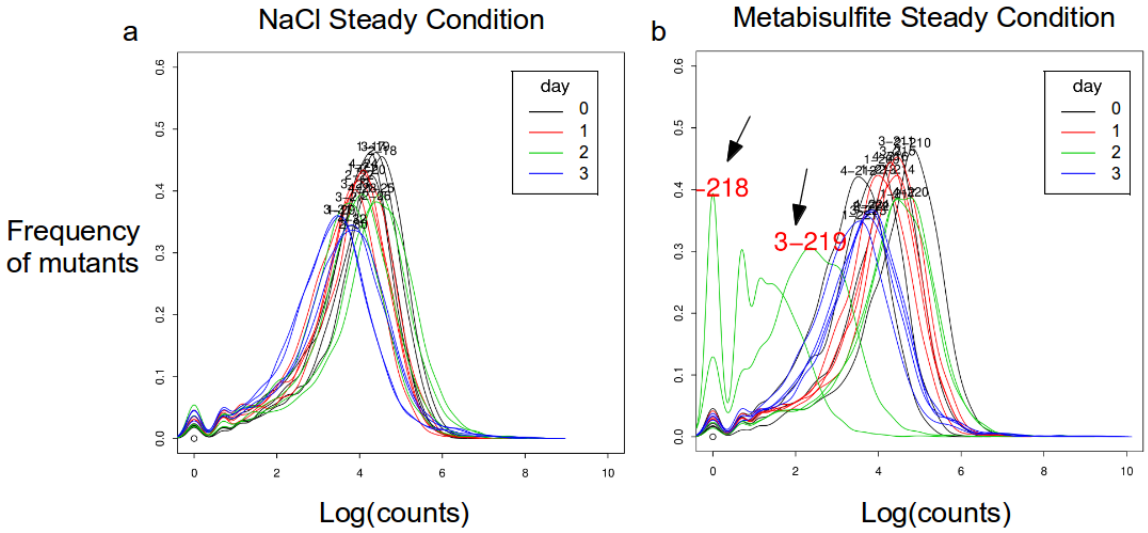


Fig III-B.8. Distributions of raw counts in individual populations. Populations of the same environmental regime were plotted together. The colors indicate the different time points of the experiment. Numbers above populations indicate population replicate number and id in the count matrix. Two populations have weird distributions (arrows), as compared to their replicates, and were thus removed from further analysis. **a)** populations in the NaCl steady stress condition. We observe a broadening of the distribution over time, with many mutants being depleted from populations. This pattern was seen in most conditions (mutants in the pool are unfit). **b)** populations in the Sodium Metabisulfite steady stress condition. Replicates two and three at day 2 (arrows) showed unexpected distributions for most conditions in Sodium Metabisulfite. There was probably a pipetting issue for those two rows in the 96-well plates.

C Data normalisation

Normalisation of count data is essential for most types of differential analysis in genomics. Indeed, the matrix of counts have populations of different sizes, with different distributions of mutants. It is assumed that most genes are not differentially represented between two conditions, and that distributions of replicates should look similar. Differences between distributions can originate from variations during library preparation or sequencing.

For unknown reasons, a small number of genes make up most of the counts in a library. This means that naïve normalisation approaches where all libraries are adjusted to the exact same size do not work: they only normalize for few highly abundant genes. Optimal normalisation methods thus try to adjust the number of counts for most of the genes so that they have comparable total counts across all conditions. We used the R package DESeq2 for normalisation of counts data and computation of Fold changes

(Anders and Huber, 2010; Love et al., 2014a). A recent survey has shown that this tool is among the most effective and robust for normalizing count data with different library sizes or library composition (Dillies et al., 2013). In DESeq2, normalisation is achieved by computing a normalisation factor, called the Size Factor, in each population, and then by multiplying all counts by this scale factor. Scale Factors are computed by dividing each counts of a mutant by its geometric mean across all conditions. Then, a population size factor is the median of those values within the population. As shown in Fig III-B.9, normalisation allows to adjust the distributions of counts, so that we can compare the different populations.

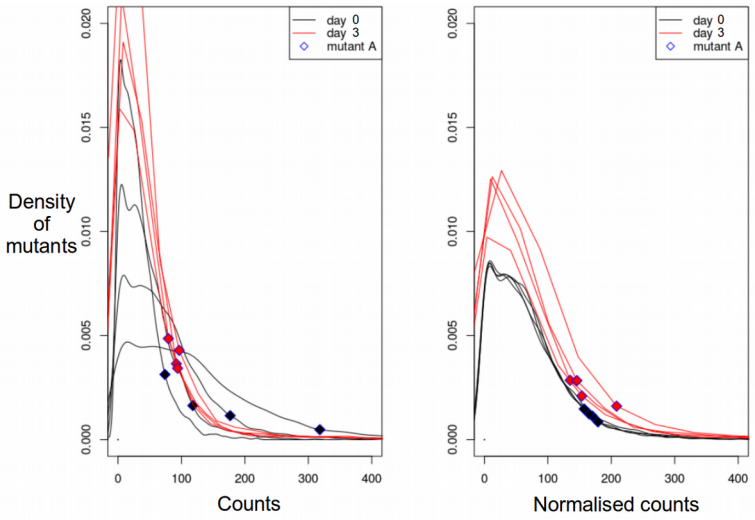


Fig III-B.9. Illustration of the importance of normalisation. After normalisation, we observe that there is no differential abundance of the mutant A between day 0 and day 3. The difference before normalisation was due to non-comparable distributions of counts between populations.

For my purposes I wanted to compare the abundance of mutants in the different periodic regimes for one type of medium fluctuation. Thus, I adopted the following design in DESeq2: the size factors were estimated together for all populations in a given medium. As shown in Fig III-B.10a, populations among the same fluctuation regime have more comparable distributions of count values after normalisation. I obtained a broad distribution of Size Factors, confirming the importance of normalisation (Fig III-B.10b).

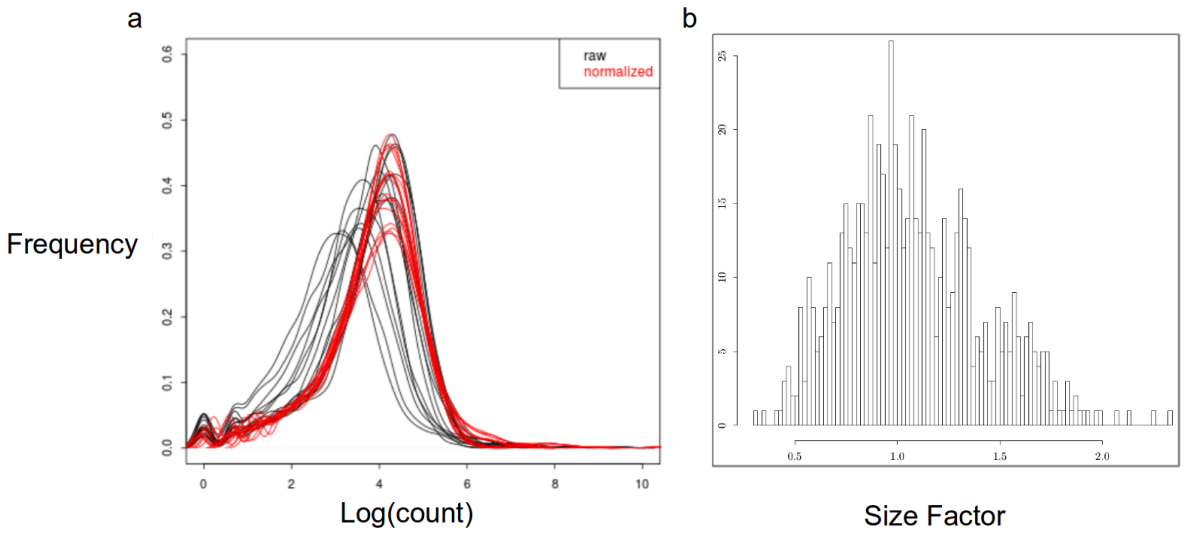


Fig III-B.10. Normalisation in DESeq2. a) Distributions of counts for the populations in the 6-hours periodic fluctuation regime in the salt fluctuation, before (black) and after (red) normalization by DESeq2. **b)** Distribution of calculated size factors in all populations.

D Computation of Fold Changes

Differential analysis is key to genomics studies. To determine if a gene is differentially abundant between conditions, a Negative Binomal (NB) model is fitted to the gene in each of the two conditions, using data points from the available biological replicates. NB models are well-fitted to genomics counts data due to the over-dispersion of counts: genes with high mean counts have higher than expected variance (between replicates). Then, a Wald test is performed to determine if there is statistically-significant difference between the two conditions, according to their NB models. Two parameters are estimated for fitting a NB model: the mean and the dispersion. The dispersion is the most difficult parameter to estimate that is key for the success of a differential analysis method.
In DESeq2, dispersion is estimated by a 3 steps approach: first a NB model is fitted to each gene, second a regression of the mean versus the variance for all genes is made, third the dispersions of genes are "shrunk" toward the regression line, at the exception of outlier genes that are far above the regression line. This approach allows to consider the underestimation of gene dispersion observed in genomics counts data (Love et al., 2014a).

In my case, data are over-dispersed, as expected from count matrix sequencing data (Fig III-B.11). I computed fold changes of normalized counts for all possible combinations of days (e.g. day 0 vs 1, or 1 vs 2 …), and stored the results in a matrix. There are 6 possible combinations of days, 12 frequencies and 4 media, thus the matrix had 288 fold changes values per mutant.

We will later see that I also used a Generalised Linear Model (GLM) for data analysis. Fold changes were used mostly for preliminary analysis, to detect interesting mutants for secondary screens.

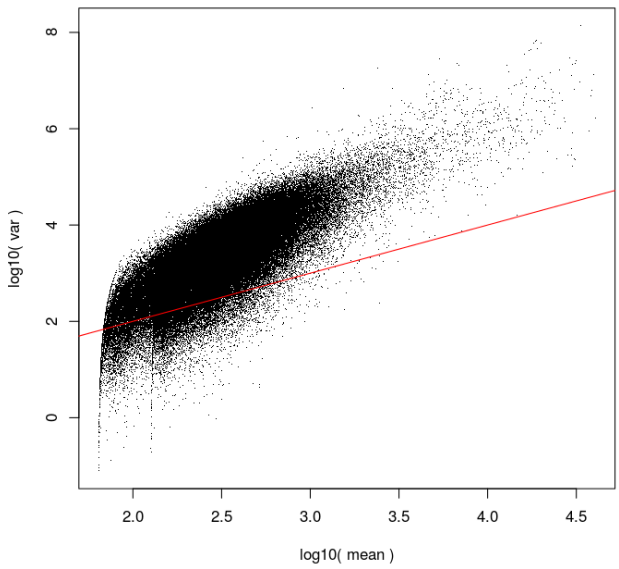


Fig III-B.11. Over-dispersion of normalized counts. Each dot indicates the mean and variance of normalized counts for one mutant in one condition (medium, fluctuation, day), across its biological replicates. The red line is the identity and corresponds to what is expected from a Poisson distribution (mean is equal to variance).

E Expected observations for well-known mutants

On Table III-B.12, we can see that mutants that are important for response to salt stress such as the MAPK and the MAPKK *hog1* and *pbs2*, or *gpd1*, the key enzyme for glycerol synthesis, are lost in populations growing in 0.2 M salt. Similarly, mutants important for the synthesis of methionine, such as *met5* and *met10* (sulfite reductases), or *met8* (siroheme synthesis, cofactor for sulfite reductase) have a marked growth defect in the steady condition without methionine.

a)

Mutant	Fold Change N0 → N3	Fold Change $S_{Salt}0 \rightarrow S_{Salt}3$
pbs2	1.43	-4.53
gpd1	-2.44	-5.51
hog1	2.48	-2.35

b)

Mutant	Fold Change N0 → N3	Fold Change $S_{Met}0 \rightarrow S_{Met}3$
met10	0.782	-5.98
met5	-0.646	-6.9
met8	1.85	-6.82

Table III-B.12. Fold changes between day 0 and day 3 of different mutants important in salt (a) and methionine (b), in the corresponding steady conditions.

F Correlations between replicates: highest reproducibility for the Salt and Sodium Metabisulfite experiment

Correlations between replicates at day 3 were higher than at day 1 or 2, and thus we decided to use data at day 3 for estimating fitness of mutants (data not shown). On Figure III-B.13 we can see the Pearson correlations between replicates in the steady conditions for days 0 and 3 (data used for computing fitness).

First, we observed that data from the run 1 (Salt and Sodium Metabisulfite) seemed to be of higher quality than data from the run 2 (Glucose and Methionine). At day 0, all conditions experimented only the N medium, therefore they should be correlated. Correlations between replicates of the N condition were higher at day 0 for the run 2. However, this is due to the method used to replace spike-in (see methods). Correlations between replicates of the S condition, or between the S and N condition were higher for the experiments from the run 1. This indicates a better quality of experiments in the run 1.

Second, we observed that the effect of the day is very strong. Indeed, for all media, correlations between replicates were very high after 3 days of growth, within the same condition (0.97 or above), but also between the steady N and S conditions (0.77 or higher).

Third, we observed that in Sodium Metabisulfite and Methionine, correlations between stressful and non-stressful media at day 3 are extremely high (~ 0.95), close to correlations within the same media. This indicates that few mutants were affected by the stress, which could be due to either to either the nature of the stress (specific instead of pleiotropic effect), or its concentration (see discussions). For glucose, correlations between conditions were a bit lower but still at ~ 0.9. Only salt seemed to have an effect on a substantial number of mutants, as the correlation between N and S diminished to ~ 0.8.

This analysis suggests that data from run 1 have the highest quality, and that there is more potential for observing pronounced genomic effects in the salt dataset.

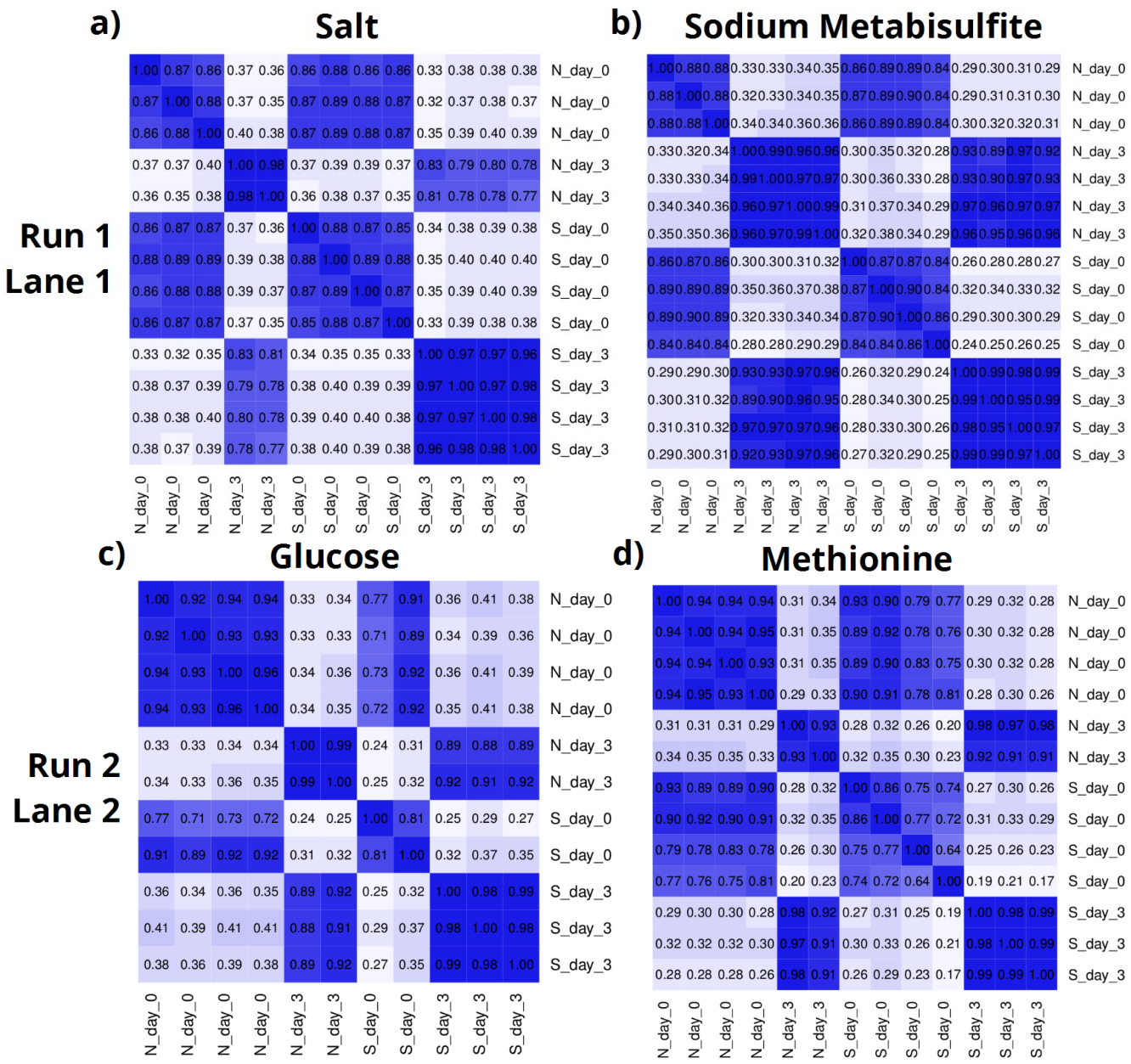


Fig III-B.13. Pearson correlations of normalized counts between replicate populations in steady conditions at day 0 and 3

iv Principal Component Analysis (PCA): days and conditions discriminate populations of the salt experiment

On Fig III-B.14 is shown the first and second principal components of a Principal Component Analysis on the normalized counts. We can see that for all media, those two components explained most of the

variations (~80% for PC1 and ~16% for PC2), and were capturing mainly the effect of the day. In salt, we observed that the steady conditions stood apart at day 2 and 3 from the fluctuating conditions, with the highest periods being further apart. This result suggested that most mutants have important fitness differences in the different conditions, which is desired for genomic analysis. In addition, the data seemed to follow an interesting pattern in the fluctuating conditions: a gradual separation from the longest to the shortest periods.

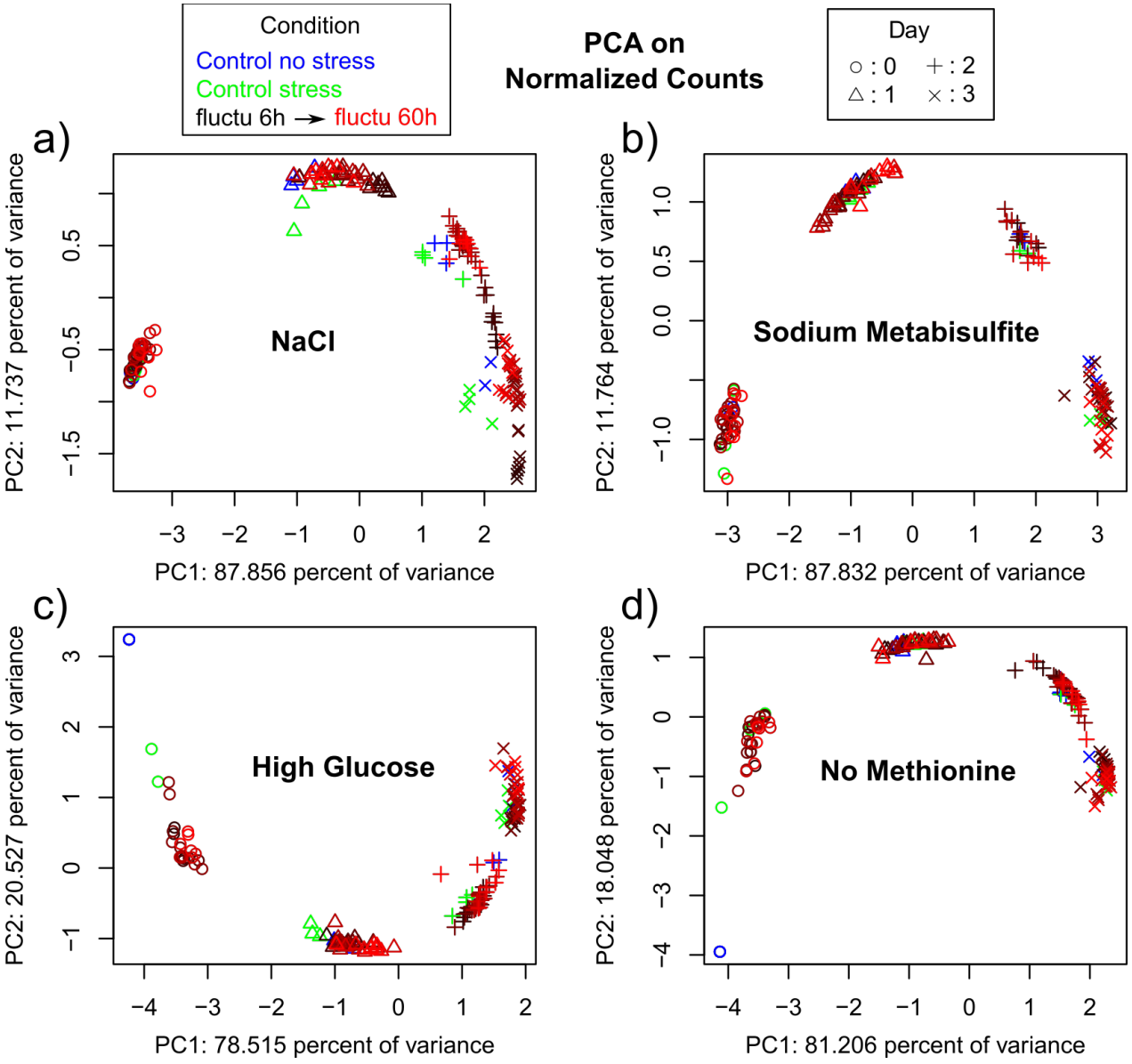


Fig III-B.14. PCA on normalized counts.

Overall, results of the analysis of correlation of replicates and of PCA on normalized counts data indicated that the salt experiment showed high data quality and affected the fitness of numerous mutants. In addition, this stress is highly studied in yeast which allows to compare our results and to infer the function of different mutants in salt stress. Thus, I decided to focus my efforts on the salt

experiment for the rest of my PhD.

C Resequencing

i Resequencing increased the power to detect mutants with subtle effects

In fall 2015, I had validated the nonlinear transgressive effect in salt of various mutants (described below). However, those validated mutants all had higher than expected fitness. None of them had lower than expected fitness. This could be due to the fact that it is easier to detect mutants with high fitness, since their count numbers increase, than mutants with low fitness, since they have very few counts, with large variabilities between replicates. We therefore considered to re-sequence our libraries in order to increase our power to detect small effects.

Robinson *et al.* tested different experimental designs of BarSeq experiments with a prototrophic haploid Yeast Deletion Library (Robinson et al., 2014) for determining the number of differentially abundant mutants between growth in glucose or galactose. They mapped almost 60 million reads per condition. They found that by sequencing only 6 million reads per condition with 4 biological replicates, the statistical power was sufficient to detect most (>70%) mutants with a fold change difference detected in their full dataset. This corresponds to ~1400 reads per mutant per condition, or 350 reads per mutant per population. Interestingly, they also showed that increasing the number of biological replicates drastically decreases the number of reads needed in order to obtain the same statistical power of detection. However, increasing the number of technical replicates did not increases the power of detection. These observations guided us for planning the re-sequencing.

Our first sequencing provided a mean of ~70 reads per mutant per population (Table III-C.1). We thought that we could obtain the level of coverage advised in Robinson *et al.*, in order to detect small effects in our pool of mutants. Thus, we sent to resequencing only part of our frozen stocks. We resequenced the Salt and Sodium Metabisulfite experiments, at 7 conditions: the 2 steady conditions and 5 periods (6h, 12h, 18h, 24h, 42h).

This second round of sequencing was also of very high quality, with 96% of reads having a quality score above or equal to 30 (table III-C.1). The quality of the data was overall similar to the quality of the first sequencing. After resequencing, we obtained a mean of 363 mutants per population. Thus, we achieved the high level of precision that we were targeting (Table III-C.1).

	# of raw reads (in M)	# of parsed and filtered reads	# of conditions	# of reads per condition	# of parsed mutants	# of reads per mutant per rep	Median # of reads per mutant per rep
First round of sequencing	296	190	48*4 = 192	0.98	*3568*	69	52
Second round of sequencing	416	234 + 57 = 289	14*4 = 56	5.16	*3568*	363	258
			Threshold :	6		349	

Table III-C.1. Number of reads sequenced per mutant per condition after the first and second sequencing. 57 million reads from the first round of sequencing were added to data from the second round of sequencing.

ii Data filtering

In order to analyse only high data quality, I looked at the distributions of counts of mutants and populations, and I set up arbitrary thresholds (Fig III-C.2). I ended up with a matrix of 3568 mutants (rows) and 208 populations (columns). As for the first sequencing, I also looked at the shape of the distributions of the replicates in each condition. No population displayed the unexpected distributions we observed in the first sequencing.

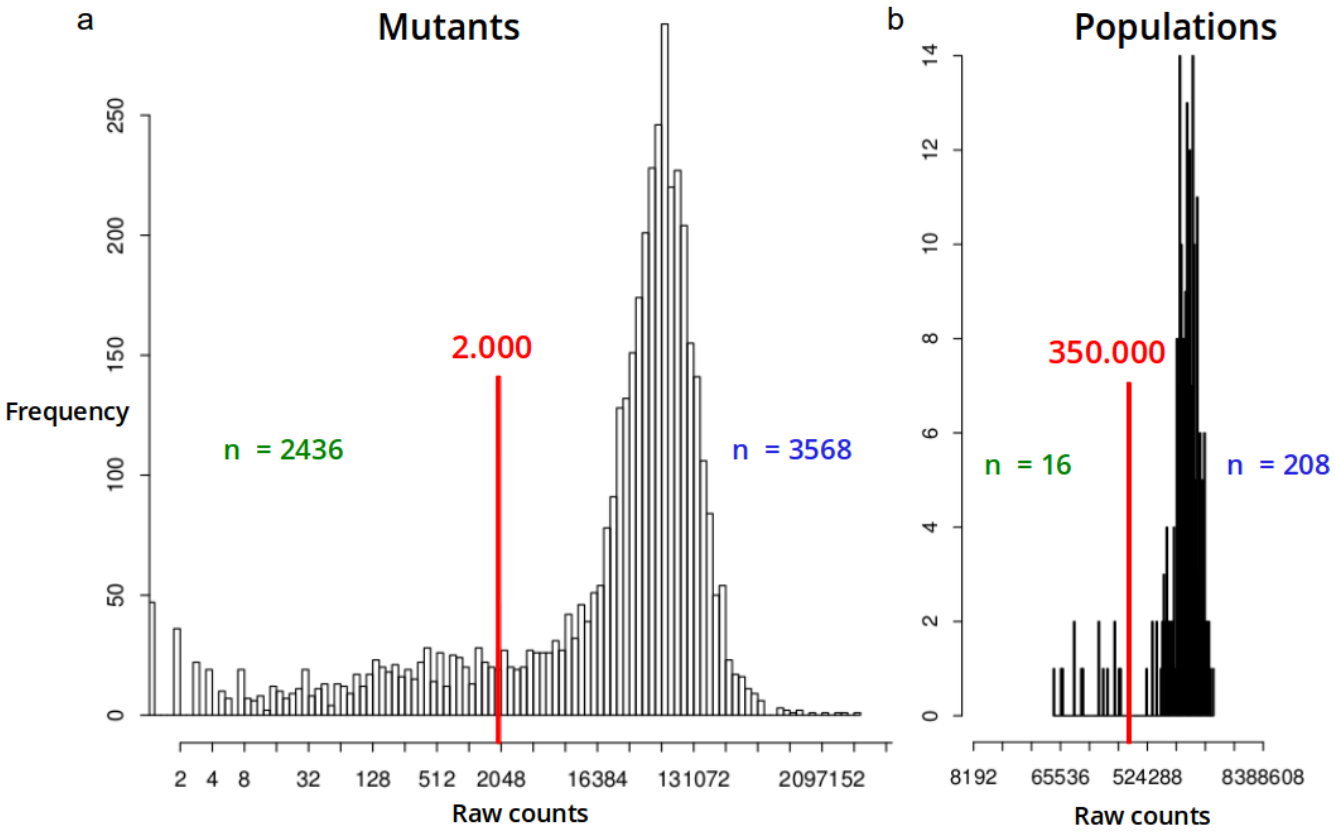


Fig III-C.2. Distribution of counts for mutants (a) and populations (b). In red is the threshold of minimal counts for keeping mutants/populations for further analysis. In green (resp. blue) is the number of mutants that are discarded (resp. conserved).

iii Correlations between replicates

Correlations between normalized counts of replicate population did not significantly change after re-sequencing (Fig III-C.3). It increased by 0.01 between some replicate populations, and decreased by 0.01 between some non-replicate populations.

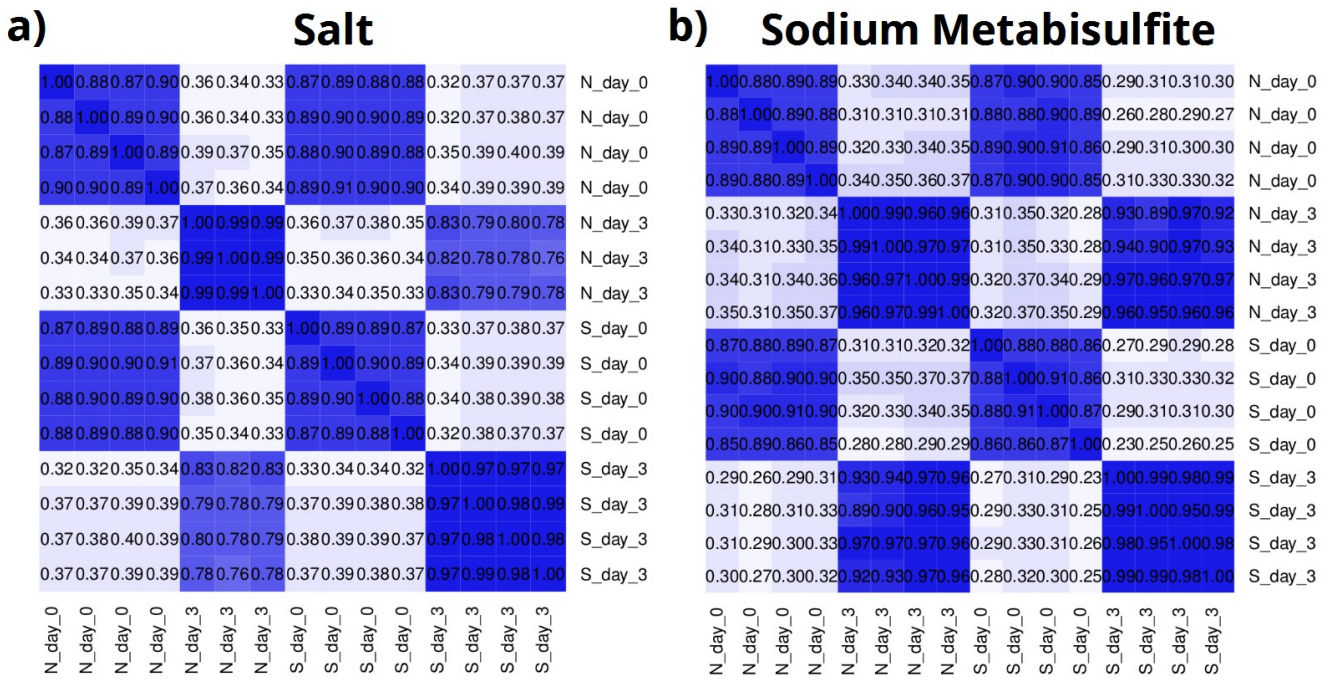


Fig III-C.3. Correlations of normalized counts of replicate populations after resequencing.

iv Number of paired replicates of steady conditions used to compute fitness

Fitness was computed using the formula of (Qian et al., 2012). It was necessary to consider the presence of replicates in my dataset. Thus, I computed fitness on paired data for a given replicate between day 0 and day 3. Then, as explained in my submitted publication, I computed the nonlinearity of fitness (wdev) by this formula:

$wdev = \frac{w_{observed}}{w_{expected}}$, where $w_{observed}$ was the fitness of the mutant strain experimentally measured in the periodic environment and $w_{expected}$ was the fitness expected given the fitness of the mutant strain in the two steady environments (N and S), calculated as:

$w_{expected} = w_N^{f_N} . w_S^{f_S}$, where f_N and f_S were the fraction of time spent in N and S media, respectively, during the course of the fluctuation experiment. In most cases $f_N = f_S = 0.5$.

However, my quality control analysis of the distribution of counts in the different populations led me to remove several populations that had unexpected distributions (Fig III-B.8). This was necessary only for data from the first round of sequencing, as the coverage was insufficient. Missing populations is an issue for analysis of paired data. Indeed, if data are available for one population/replicate at day zero

but not at day 3, then fitness cannot be computed for this population. Fitness values in the steady conditions are critical as they are used to compute the nonlinearity of all other conditions. Having more replicates in the steady conditions means more robust wdev values in all associated fluctuating conditions.

We can see in table III-C.4 the number of replicates that were used to compute fitness in the steady conditions of the four experiments. We can see that experiments that were sequenced only one time had between 2 and 3 paired replicates per steady condition. In contrast, experiments that were sent to re-sequencing had between 3 and 4 paired replicates per steady condition. Overall those results mean that the experiments that were not resequenced have fewer replicates to compute fitness in steady conditions and have thus less robust wdev values in all conditions. Thus, care should be taken for interpreting genomic results of those experiments.

Round of sequencing	Periods sequenced	Medium	N day 0	N day 3	# of paired replicates	S day 0	S day 3	# of paired replicates
First	6h → 60h	Salt	3	2	2	4	4	4
First	6h → 60h	Na2S2O5	3	4	3	4	4	4
Second	6h, 12h, 18h, 24h, 42h	Salt	4	3	3	4	4	4
Second	6h, 12h, 18h, 24h, 42h	Na2S2O5	4	4	4	4	4	4
First	6h → 60h	Glucose	4	2	2	2	3	2
First	6h → 60h	Methionine	4	2	2	4	3	3

Table III-C.4. Number of paired replicates in the steady conditions in all experiments after read parsing and data filtering.

D Generalised Linear Model (GLM)

We wanted to set up a model to determine which mutants are statistically significantly nonlinear in periodic conditions. Poisson models are well-fitted to data of counts of events within a specific time period. The Poisson law can be applied to data that respect the property of homoscedasticity: meaning that variability is constant with the mean, and does not increase or decrease with higher count values. However, count data from sequencing are usually over-dispersed: higher counts have higher variabilities. A Negative Binomial model is more fitted to this type of data (Love et al., 2014b). GLM use a link function to predict model parameters from linear combinations of predictors.

As described in the submitted publication, we fitted a GLM to our count data. We assumed that the normalized counts of mutant i in condition c (N, S or periodic) at day d in replicate population r originated from a negative binomial distribution NB(λ_i, α), with :

$$\log(\lambda_i) = offset_{i,c} + \beta_{i,1}.t^{N}_{c,d} + \beta_{i,2}.t^{S}_{c,d} + \beta_{i,3}.N^{changes}_{c,d} + \varepsilon_{i,c,d,r}$$

and $offset_{i,c}$ being the median of normalized counts for condition *c* at day 0, $t^{N}_{c,d}$ and $t^{S}_{c,d}$ being the amount of time spent in medium N and medium S at day *d*, respectively, $N^{changes}_{c,d}$ being the number of changes between the two media that took place between days 0 and *d*, and ε being the residual error.

If fitness is linear in a fluctuating environment, then it is insensitive to the number of changes and $\beta_{i,3} = 0$. Nonlinearity can therefore be inferred from the statistical significance of the term $N^{changes}_{c,d}$ of the model. The corresponding *p*-values were converted to *q*-values, using package *qvalue* version 2.0.0 in order to control the False Discovery Rate.

I wanted to determine if mutants with high nonlinearity of fitness at some frequencies also had high nonlinearity of fitness at other frequencies. For this purpose, I made correlations on the q-values of the $\beta_{i,3}$ term of the GLM (Fig III-D.1). The only experiment for which I observed such correlation was the salt experiment: the three shortest periods were more highly correlated to each other than other conditions. This result is consistent with the PCA described above.

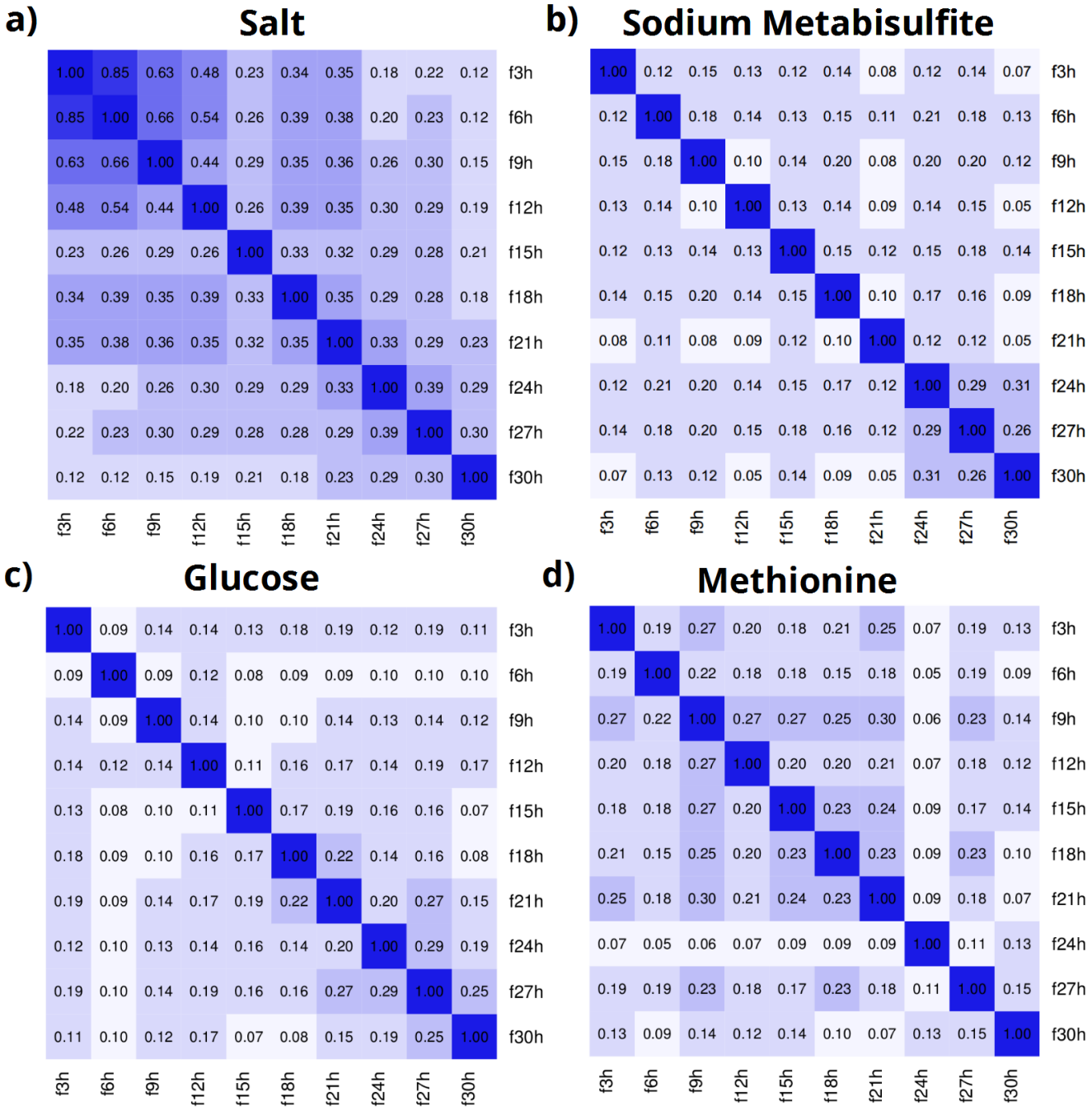


Fig III-D.1. Pearson correlations of the -log10 of the qvalues of the GLM in the different conditions after the first sequencing.

E Genomic analysis

i No general relationship between expected and measured fitness at the 6 hours period

Comparing expected fitness versus measured fitness is indicative of the predictability of fitness in fluctuating environments. We observed in the Salt experiment that all nonlinear mutants with expected fitness higher than ~1.07 have higher-than-expected measured fitness (Fig III-E.1a, submitted publication). One could think that it may be a general effect observed in fluctuating conditions. That is, having a high fitness in two conditions induces an even higher fitness in binary periodic fluctuations between those two conditions. However, this is not a general rule as we did not observe this phenomenon in other types of media fluctuations (Fig III-E.1b-d). Indeed, in the Methionine experiment, most mutants with high expected fitness had lower than expected fitness (Fig III-E.1d). In sodium metabisulfite, the *thi2* mutant stood out (Fig III-E.1b). It had the highest fitness nonlinearity (see section 3-III-F-iii). Interestingly, the *thi2* gene plays an important role in resistance to sodium metabisulfite stress (see discussions).

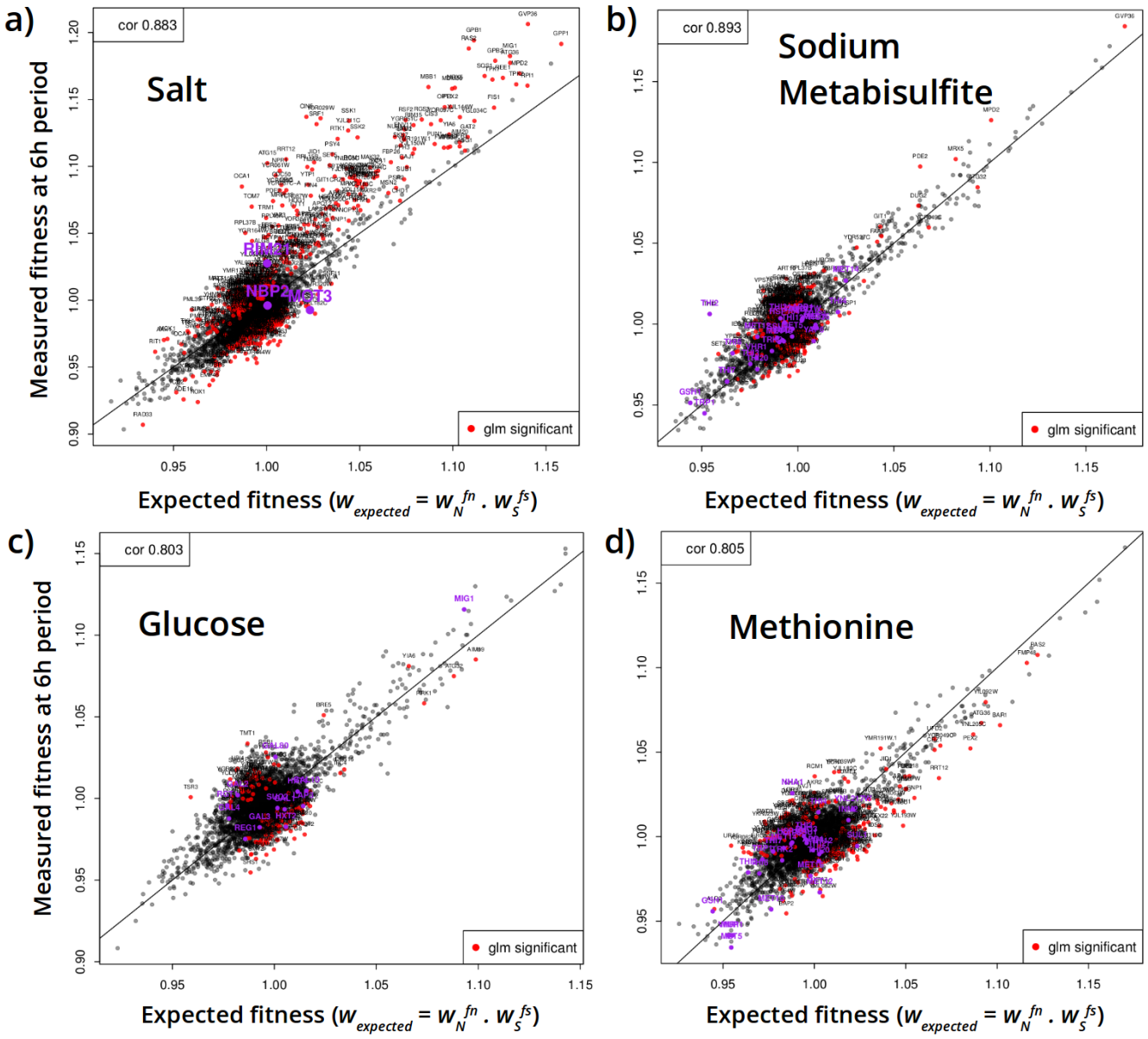


Fig III-E.1. Expected vs measured fitness at the 6h period in the different media fluctuations. Presented data are from the first sequencing for the Glucose **(c)** and Methionine **(d)** media, and from the re-sequencing for the Salt **(a)** and Sodium Metabisulfite **(b)** media. On the top left of each plot is shown the Pearson correlation coefficient between expected and measured fitness. Genes highlighted in red have a qvalue below 0.0001, and genes highlighted in purple are known in the literature to play a role in the corresponding stressful condition. w_N (resp. w_S) is fitness at day 3 in medium N (resp. S). *fn* (resp. *fs*) is the fraction of time spent in the N (resp. S) medium.

ii No general relationship between Antagonistic Pleiotropy and fitness nonlinearity

In the submitted publication, I defined a mutant as Antagonistic Pleiotropic (AP) if it had w_N (fitness in N) and w_S (fitness in S) values of different sign, and a large absolute difference between w_N and w_S. In the salt fluctuation, we observed that mutants with the highest nonlinearity were mostly Antagonistic Pleiotropic (AP) mutants with significantly higher fitness in N than in S_{Salt} (Fig III-E.2a). We didn't observe this effect in the other media fluctuations. Thus, the buffering effect observed for mutants with much higher growth in N than S was specific to the salt fluctuations.

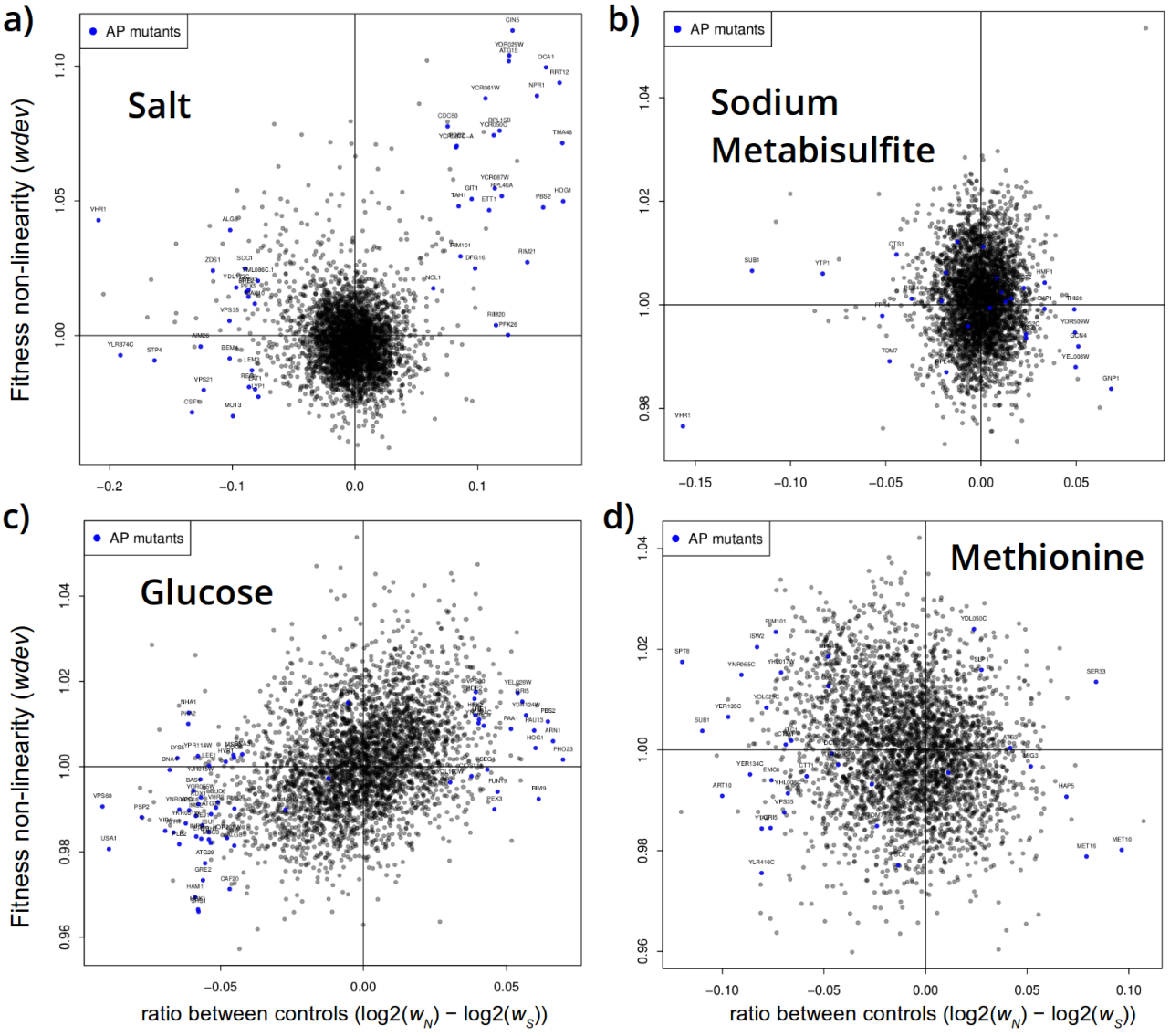


Fig III-E.2. Distance of fitness between steady conditions vs fitness nonlinearity. Presented data are from the first sequencing for the Glucose **(c)** and Methionine **(d)** media, and from the re-sequencing for the Salt **(a)** and Sodium Metabisulfite **(b)** media.

iii Fitness nonlinearity as a function of fluctuating period

Looking at the distributions of fitness nonlinearity values (wdev) in the different media revealed different patterns than in the salt experiment (Fig III-E.3). The bulk of mutants (the extremities of the estimated kernel densities, before outliers) have similar wdev values in all experiments: with upper bound at around 1.025 and lower bound at around 0.975. In all experiments, the median wdev was centered around one. In the sodium metabisulfite experiment, there were very few outliers that had wdev values above 1.05 or below 0.95 (Fig III-E.3b,e). And almost all of them are at the longest periods of fluctuations (48h, 54h and 60h). Finally, in the methionine experiment, the pattern was the opposite than in the salt experiment; with broader wdev values for longer periods (Fig III-E.3f). However, both the high and the low tail were increasing, indicating that mutants with higher or lower than expected fitness appeared in the population. Overall, it seemed that no experiment showed a pattern as clear as in salt: where shorter periods have more outliers with high wdev values. It is possible that the lower level of stress in the sodium metabisulfite experiment, and the lower quality of data in the glucose and methionine experiment, hinder the apparition of a clear shaded effect, as seen in the salt experiment. Alternatively, periodic salt stress may have induced a more pleiotropic widespread genomic response.

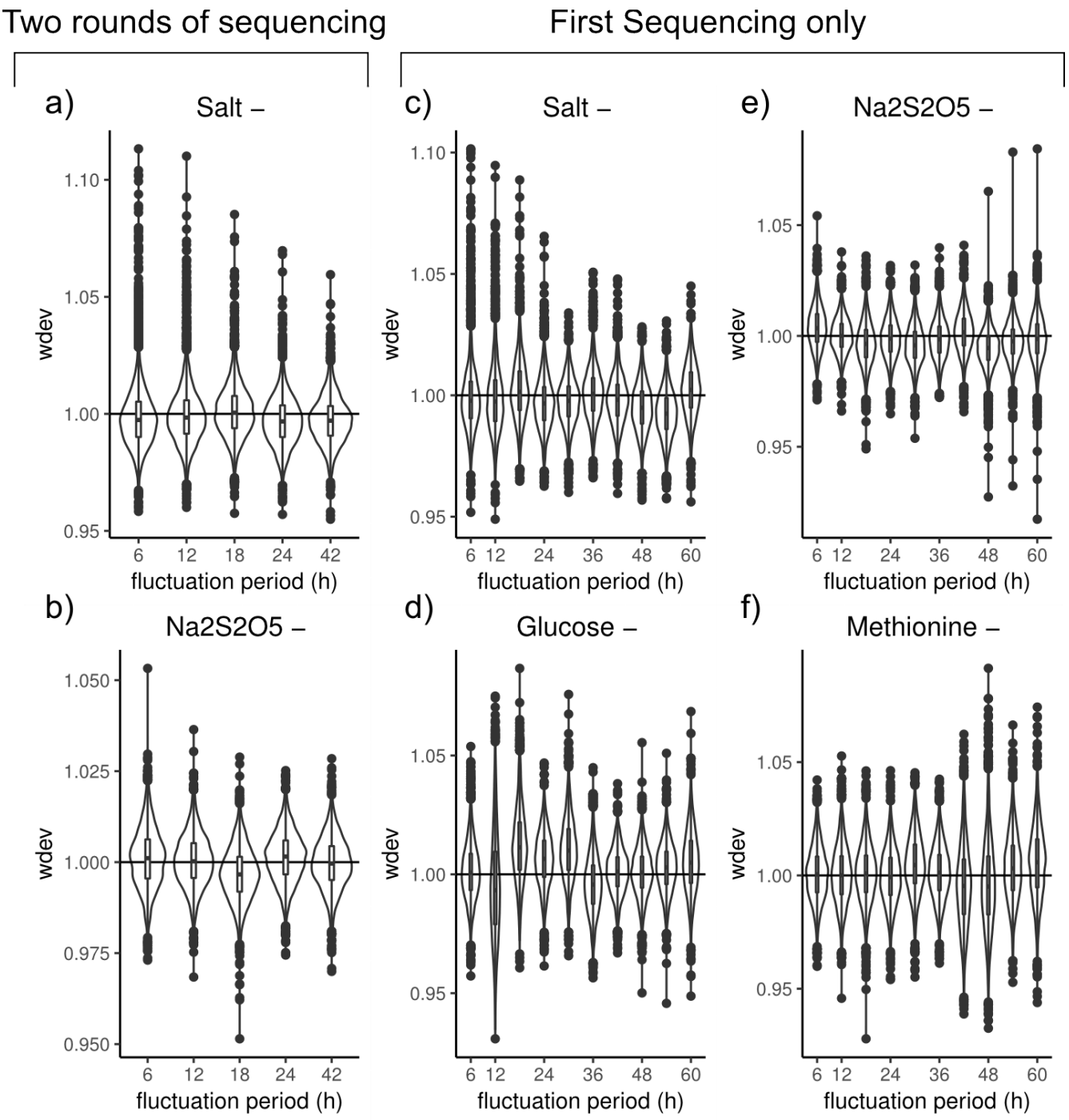


Fig III-E.3. wdev values distributions in the four experiments after the resequencing (a-b) or the first sequencing only (c-f)

F Detailed plots of some interesting nonlinear mutants

A key aim of my BarSeq experiments was to discover mutants with strong nonlinear fitness. I tried to find such mutants using different approaches. In this section, I provide details for some mutants with nonlinear fitness in fluctuating conditions in the 3 experiments that are not described in my submitted publication.

For all the text below; nonlinear mutants that have higher (resp. lower) than expected fitness will be called winners (resp. losers).

i Biases that can lead to false positives

A Mutants with aberrant counts in the steady conditions at day 0

In many cases, I observed that one of the two steady conditions had zero counts at day 0 or 3. For instance, we can see that there was probably an issue with the raw counts of the *ayr1* and the *mrx12* mutants in the S condition at day zero (Fig III-F.1). Indeed, these data point seems completely aberrant and result in high fold changes of counts in the S condition (up to 23 for the *ayr1* mutant) that are probably artificial. The normalisation step didn't correct for that, since this issue is probably specific to these genes and not to an error in the whole population.

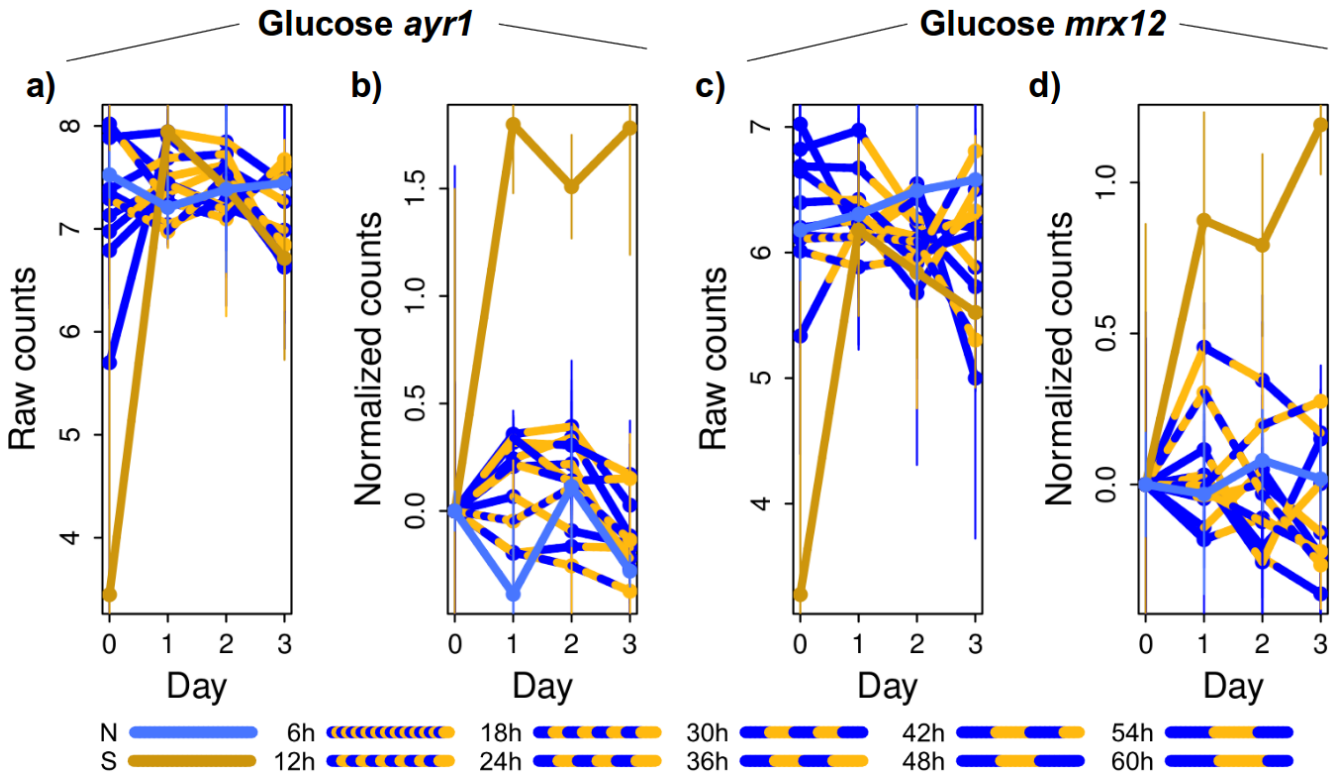


Fig III-F.1. Counts of the *ayr1* and the *mrx12* mutants over time in the Glucose experiment. Value: log2(counts) (± s.d.) (in order to easily visualize fold changes).

I observed that data from the Glucose and Methionine experiments are very noisy in general, with many other mutants that had normalisation issues at day zero or day 4, especially in the steady conditions, which can bias computation of fitness and wdev values. This is probably related to the lower number of replicates available in those conditions (Table III-C.5). Another factor that explain the increased noise in those experiments is the lower level of correlations between replicates observed previously (Fig III-B.13). Future experiments should probably include more replicates, especially for the steady conditions as they are critical for computation of wdev at all frequencies.

B Mutants with aberrant counts in the steady conditions after day zero

Even when it is biologically relevant that the mutants reach zero counts after some time, this can bias the analysis. For instance, the *thi3* mutant is undetectable after 1 day in sodium metabisulfite (Fig III-F.2). After this point, its fitness in Na2S2O5 cannot decrease anymore (since counts cannot be negative). It even artificially increases due to the normalisation procedure: more mutants are depleted in the pool over time and thus relative fitness that is computed by using an “artificial” wild-type will increase comparatively.

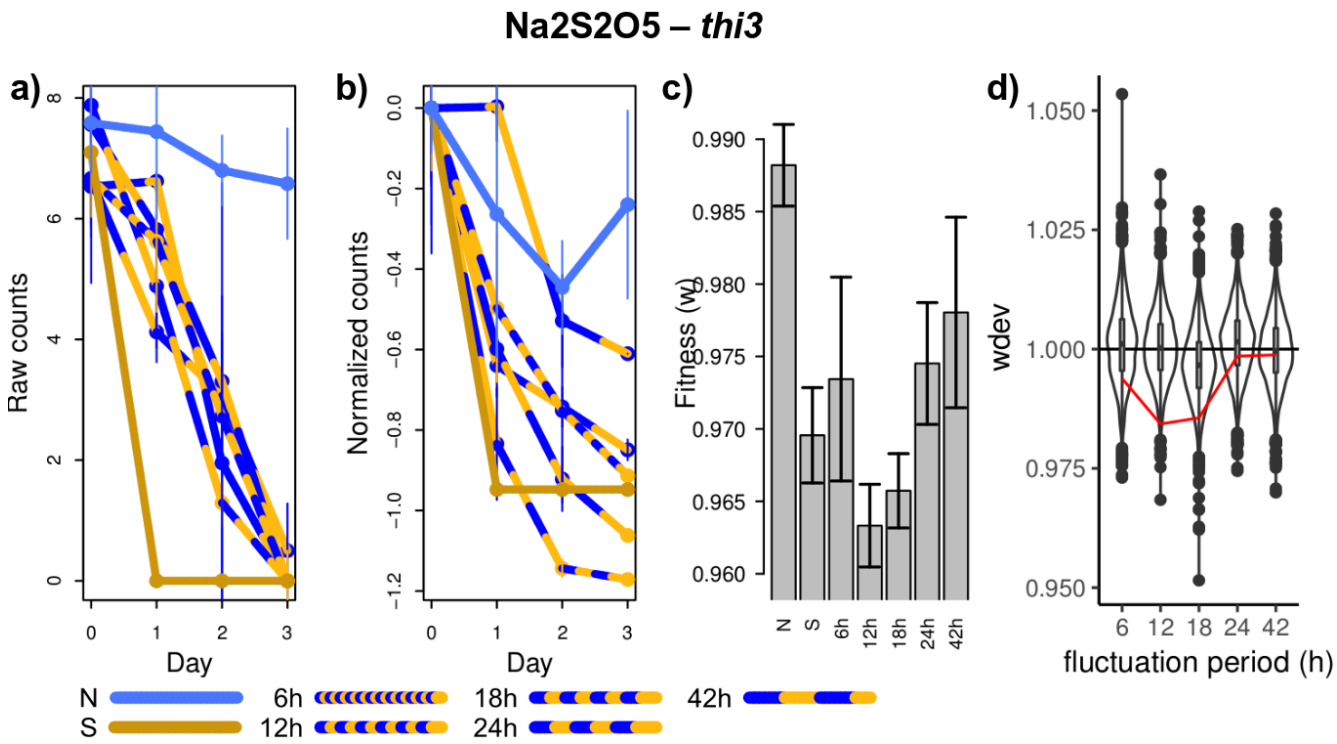


Fig III-F.2. ***thi3*** **in the Sodium Metabisulfite experiment: a-b)** median of log2(counts) (± s.d.), **c)** mean fitness (± s.e.m) at day 3, **d)** distributions of wdev values of all mutants, with those of *thi3* in red.

C Mutants with very low counts or no counts at day 0

Some mutants were present at a very low abundance in the pool at the beginning of the experiment. Those mutants, such as *apq12*, are absent (or almost) of the raw count table at day zero and then they may increase over time. However, the increased variability of very low counts at day zero can bias the analysis and lead to false estimation of differential abundance. For instance, we can see on Fig III-F.3, that *apq12* was a transgressive winner in Salt and a transgressive loser in Glucose when looking at normalized counts. However, raw counts show a different pattern: variation in the initial abundance at day zero fully explained the detected effect. Those mutants have a too low data quality and were not considered for secondary validation assays.

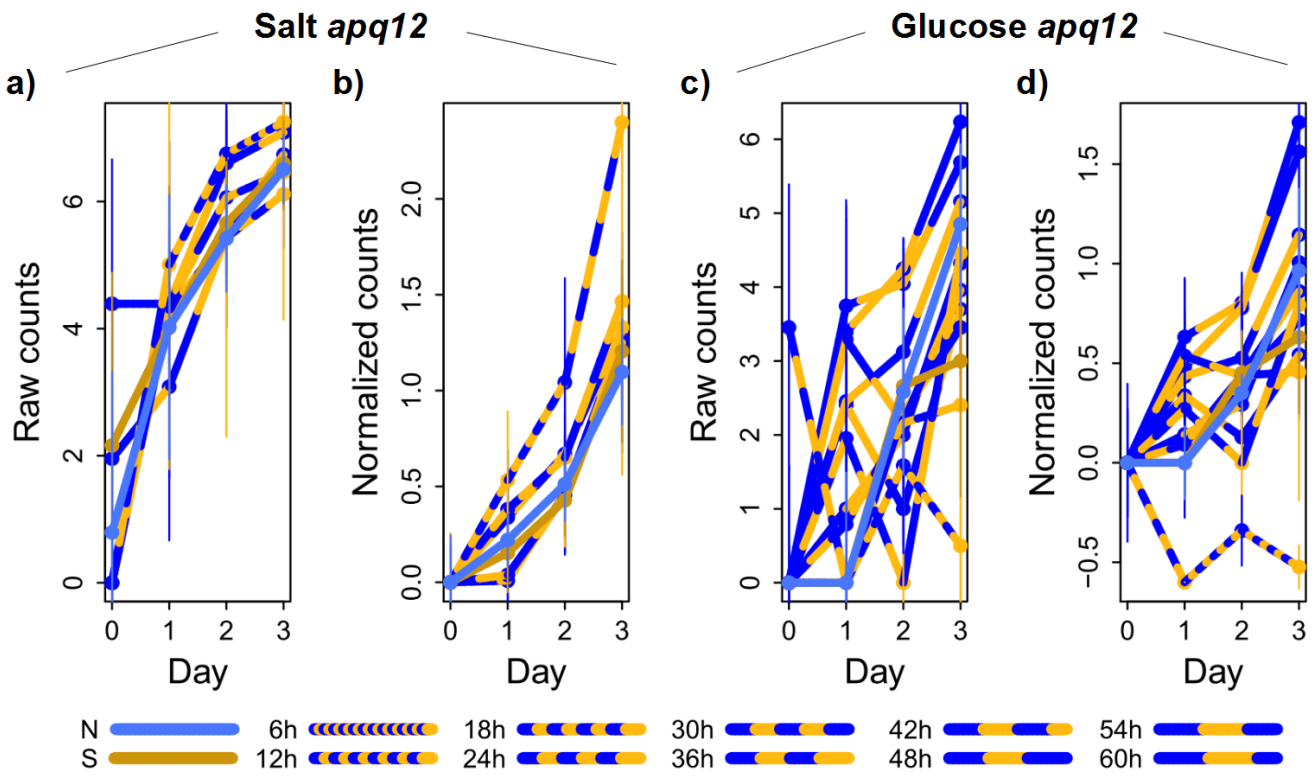


Fig III-F.3. Counts of *apq12* over time. Value: median of log2(counts) (± s.d.).

D Mutants with very high growth in all conditions

Some mutants grew faster than others in all conditions. For example, the *ras2* mutant increases in abundance 3 to 4 fold in most conditions of the methionine experiment (Fig III-F.4). In this situation, very slight differences in normalisation can largely impact the differential abundance of those mutants in various conditions, resulting in important fitness differences, and thus extreme wdev values, while raw counts data did not support such differences. Thus, caution should be taken when considering data from such mutants.

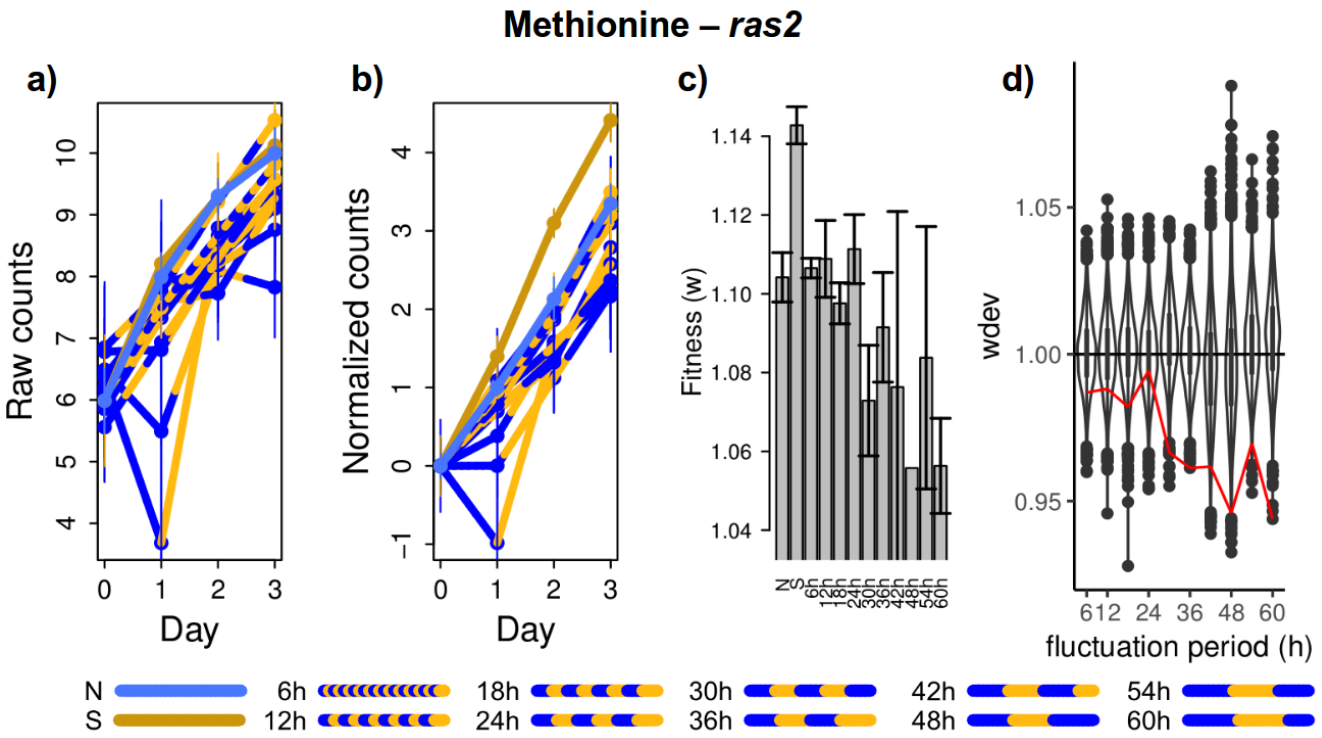


Fig III-F.4. *ras2* in the Methionine experiment: a-b) median of log2(counts) (± s.d.), **c)** mean fitness (± s.e.m) at day 3, **d)** Distributions of wdev values of all mutants, with those of *ras2* in red.

E Mutants with very slow growth in all conditions

At the opposite, some mutants were rapidly depleted in the population, such as *ino4* in Glucose or *gtr1* in Methionine (Fig III-F.5). Those mutants are so unfit that they have a competitive fitness close to zero. During the quality-control step applied to the count table, I removed strains with less than 2000 reads in all conditions. However, strains that have many reads at day zero and none at later days, or strains that had many reads only in one experiment and few in the other, were not removed (there were few such mutants). *gtr1* and *ino4* are depleted immediately in all four media. Across all conditions, they have respectively 2414 and 16517 counts at day 0, and 346 and 12 counts after day zero. Comparing quantitative fitness changes for these mutants is irrelevant and they were not considered for further validation assays.

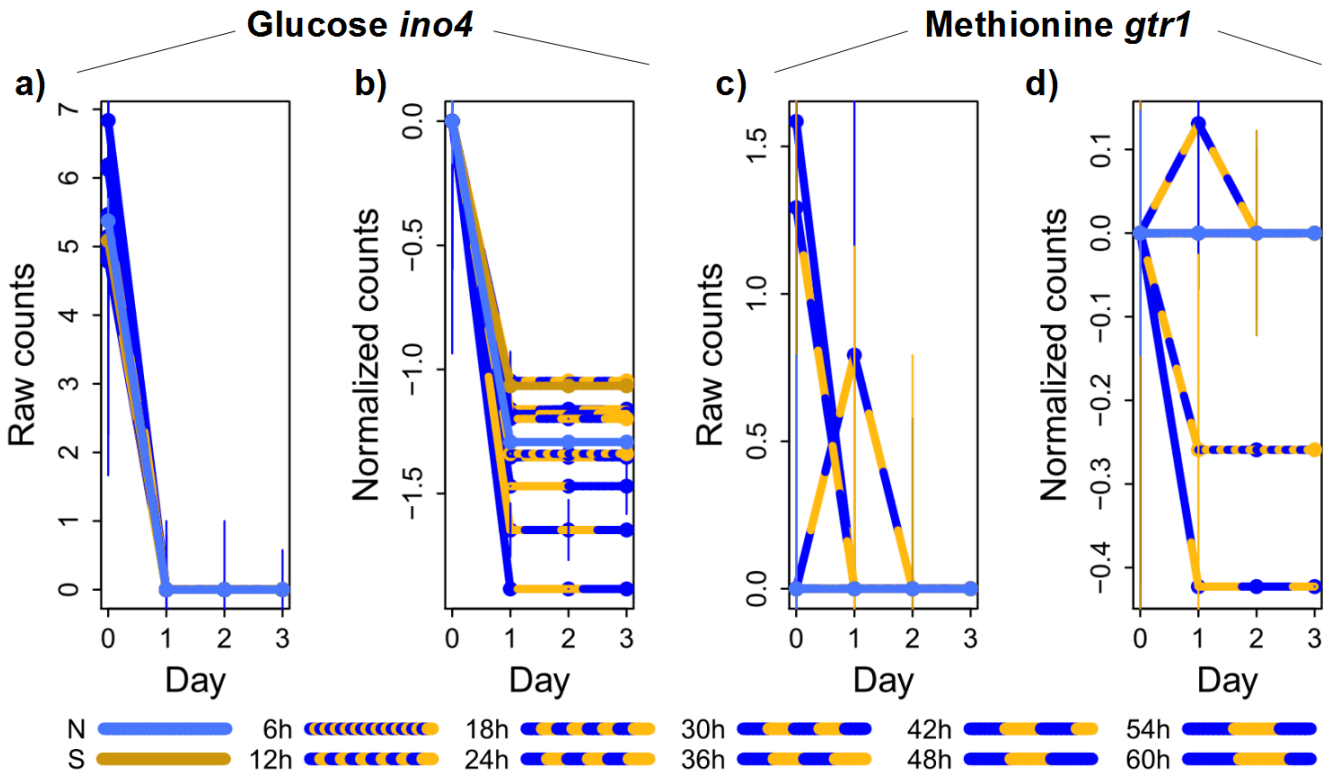


Fig III-F.5. Counts of *ino4* in the Glucose experiment and *gtr1* in the Methionine experiment. Value: median of log2(counts) (± s.d.).

F Conclusions

To conclude, data from BarSeq experiments contain a lot of variability. Careful analysis should be made before concluding on the fitness effect of a given mutant. When looking at individual mutants, I observed higher level of noise in the glucose and the methionine experiment. This is consistent with the lower correlation of replicates described above, and the lower sequencing depth. Thus, genomic results in those conditions should be interpreted with caution.

In order to have a high confidence in the BarSeq data of a given mutant, some criteria can be defined. Trustworthy nonlinear mutants should have: enough raw counts at day zero in all conditions (to be above the noise inherent to small numbers), a monotonic change in fitness of all conditions over time, a growth that is not exceptionally high or low in all conditions, no aberrant points in the steady controls, no condition that reaches zero counts during the experiment, and a strong effect that can resist the noise in the replicates. However, even when all those conditions are met, only secondary experiments can ultimately validate the effect observed in a BarSeq experiment.

Mutants presented below have been selected to have interesting non-linear behaviours, with the fewest possible of the bias described above.

ii vhr1: pleiotropy but condition-dependent response to oscillations

A strong interest was to find mutants with a pleiotropic nonlinear effect in different experiments. If such mutants were found, it could indicate that their deleted gene participates in regulating fitness in unrelated fluctuating conditions. However, I could not find such mutant. This difficulty could be related to the fact that conditions that I tested were probably not stressful enough. Alternatively, it could be related to the lower quality of the glucose and methionine experiments. Finally, it could be that there are no mutant that plays a pleiotropic role in regulating fitness in different types of fluctuating conditions. Nevertheless, one mutant displayed an interesting non-linear behaviour in several unrelated fluctuating conditions.

VHR1 is a transcription factor that activates biotin genes in response to low biotin concentrations. I found that the *vhr1* mutant had a very special behaviour in the different experiments. It had similar normalized counts profiles in the steady conditions of 3 experiments: Salt, Na2S2O5 and Methionine (Fig III-F.6). Normalized counts were relatively constant in S, and sharply decreasing in N (Fig III-F.6). However, the effect in fluctuating conditions varied in the different media. In the salt experiment, fitness was highly non-linear at almost all tested periods (Fig III-F.7a, III-F.8a). In addition, fitness was higher than expected for the short periods, and lower than expected for the long periods (Fig III-F.8a,c). In the sodium metabisulfite experiment fitness was lower than expected at both long periods (Fig III-F.8e) and short periods (Fig III-F.8b,e). By comparing fitness and wdev values for the *vhr1* mutant in the sodium metabisulfite experiment, it could be though that there are apparent discrepancies (Fig III-F.7b and Fig III-F.8b). For instance, *vhr1* has similar fitness in the periodic 6 hours and 12 hours conditions, but different wdev values. However, as mentioned above, plots of fitness represent mean fitness values (± s.e.m.), while wdev values are computed using median fitness values. Thus, those plots are not directly comparable. In the methionine experiment, fitness was linear in most conditions (Fig III-F.8f). It should be noted that variability between replicates was important in the Sodium Metabisulfite and Methionine experiment (Fig III-F.6b,d, Fig III-F.7b,d). Thus, the precise level of linearity of fitness remain to be determined. However, differences are marked with the salt experiment. To conclude, despite having the same growth effect in 3 different media, the *vhr1* mutant had either linear or non-linear effects in fluctuating conditions depending on the specific frequencies and types of media fluctuation.

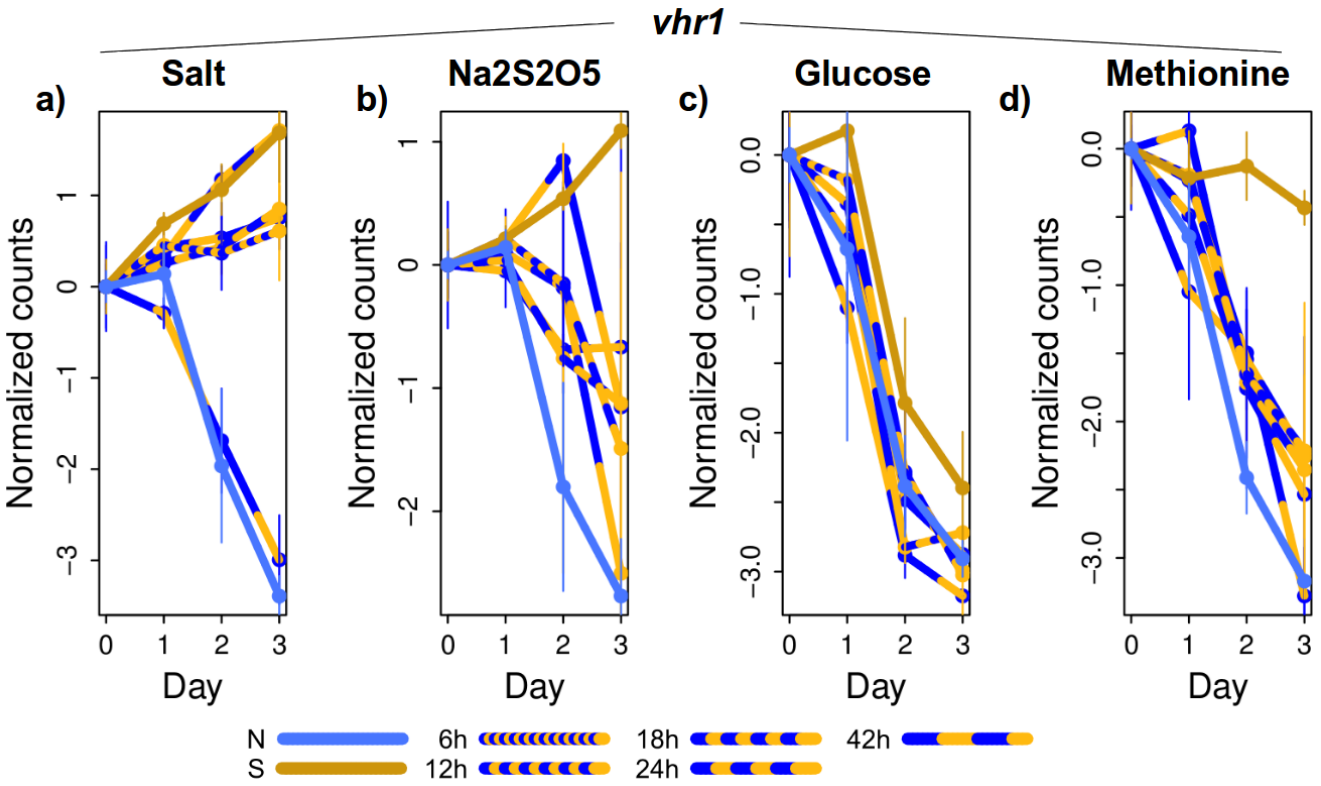


Fig III-F.6. Normalized counts over time (in days) of *vhr1* in the four experiments. Value: median of log2(counts) (± s.d.) adjusted to 0 at day 0.

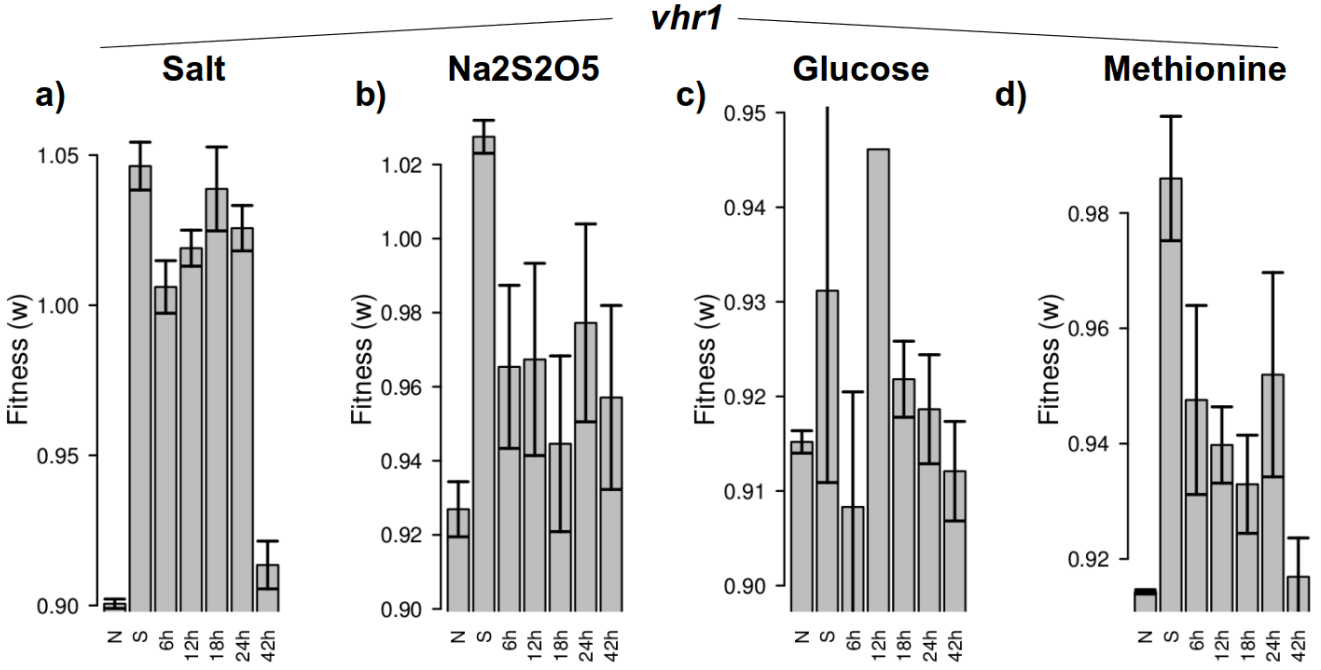


Fig III-F.7. Fitness at day 3 of *vhr1* in the four experiments. Value: mean fitness (± s.e.m) at day 3.

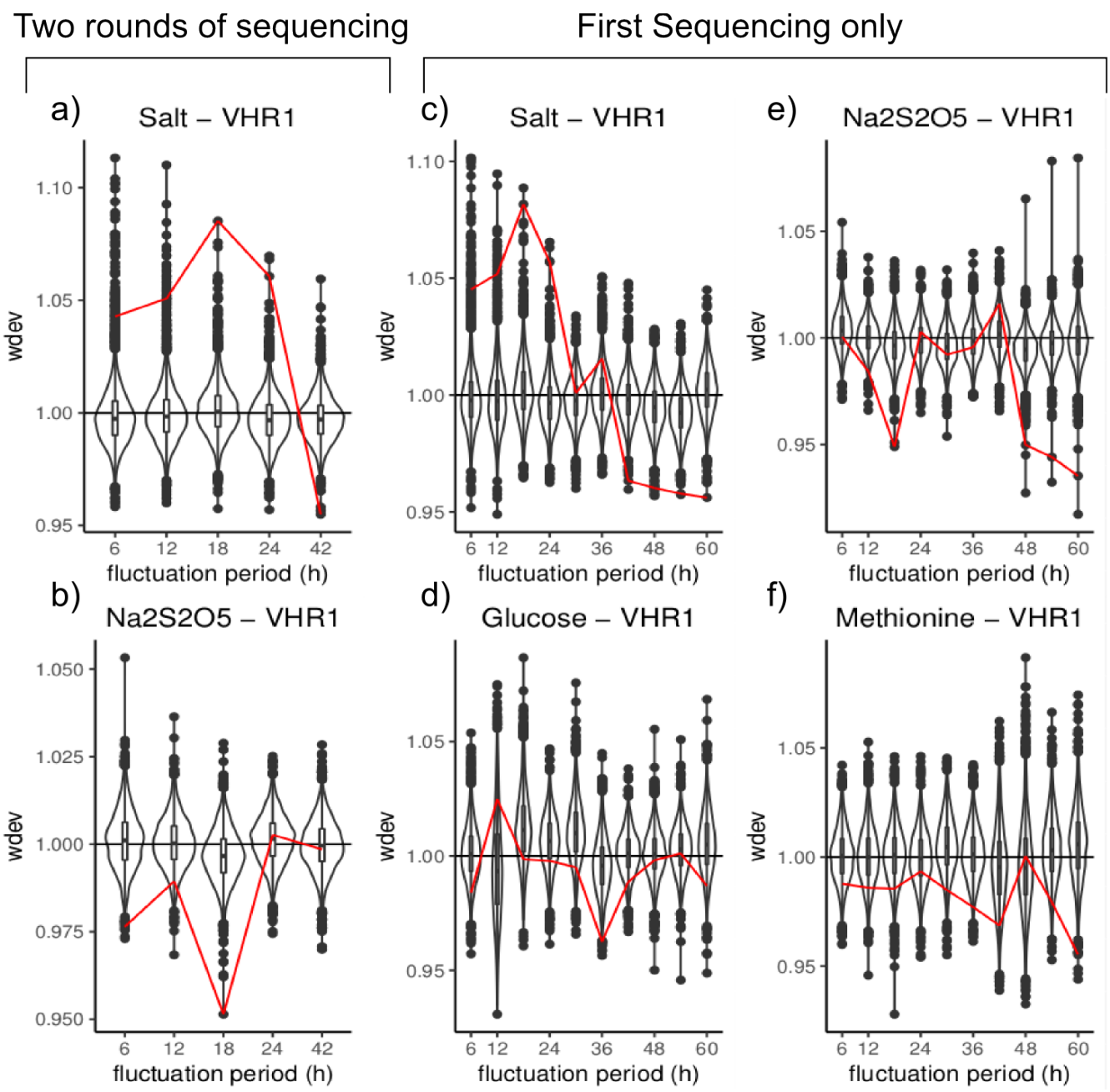


Fig III-F.8. Distributions of wdev values of all mutants, with those of *vhr1* in red: after two rounds of sequencing in the Salt experiment **(a)** and in the Sodium Metabisulfite experiment **(b)**; or after one round of sequencing in the Salt experiment **(c)**, in the Glucose experiment **(d)**, in the Sodium Metabisulfite experiment **(e)**, and in the Methionine experiment **(f)**.

iii thi2 in sodium Metabisulfite: strong antagonistic nonlinear effect across periods

THI2 codes for transcription factor that activates thiamine biosynthesis genes in response to low biotin concentration. As for the *vhr1* mutant, the *thi2* mutant showed a bivalent behaviour: with the highest wdev values (after resequencing) at the 6h and 12h periods, and the fourth lowest wdev value at the 42h period (Fig III-F.9). Data from the first sequencing reveals that this mutant had lower than expected fitness at all long periods (Fig III-F.10c,d). Thus, it seems that activating thiamine genes is useless when the environment fluctuates quickly, while it is critical when it fluctuates slowly (see discussions).

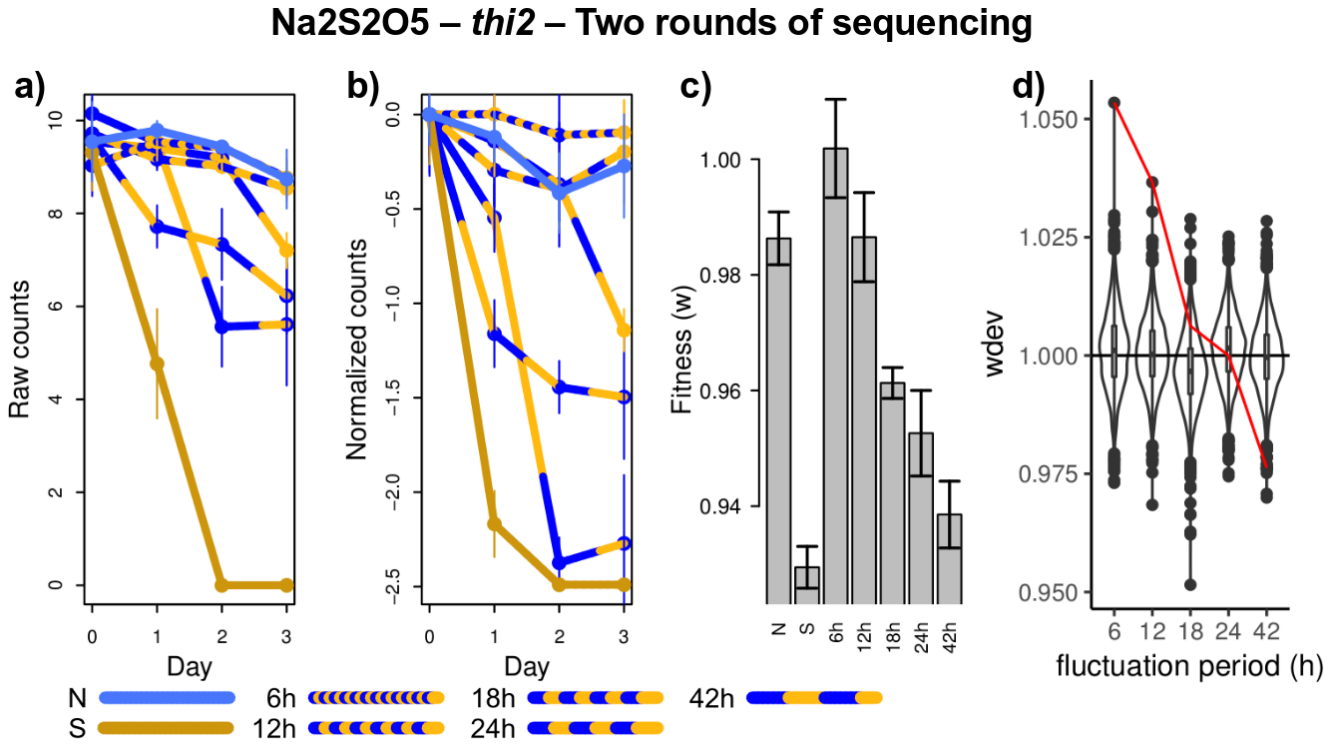


Fig III-F.9. *thi2* in the Sodium Metabisulfite experiment after two rounds of sequencing: a-b) median of log2(counts) (± s.d.), **c)** mean fitness (± s.e.m) at day 3, **d)** distributions of wdev values of all mutants, with those of *thi2* in red.

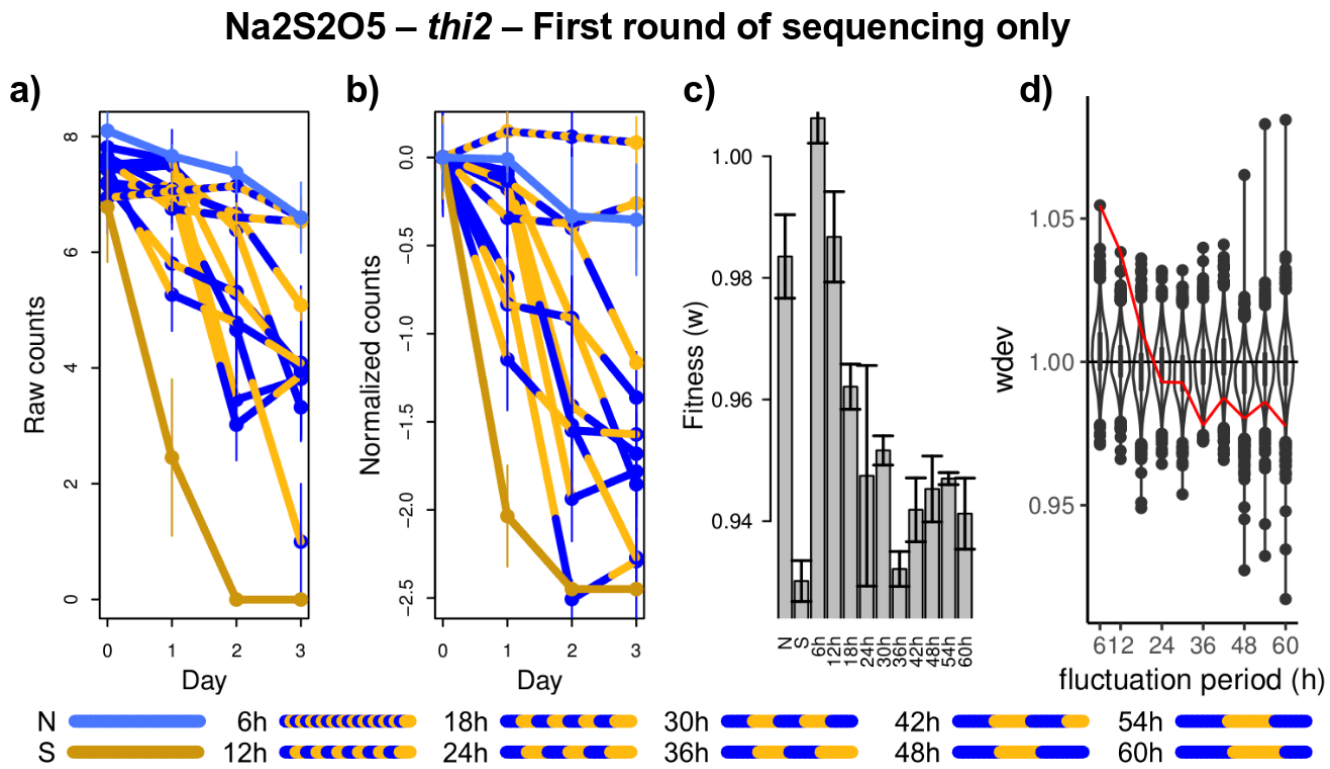


Fig III-F.10. ***thi2*** **in the Sodium Metabisulfite experiment after one round of sequencing: a-b)** median of log2(counts) (± s.d.), **c)** mean fitness (± s.e.m) at day 3, **d)** distributions of wdev values of all mutants, with those of *thi2* in red.

iv met12 in Methionine: nonlinear loser at short periods

The MET12 gene codes for a protein that has a major isozyme of methylenetetrahydrofolate reductase (MTHFR) activity in vitro. Its activity is redundant with MET13 and less pronounced than MET13. The *met12* mutant had transgressive loser phenotypes at most periods (Fig III-F.11) and a good overall data quality. Results for *met12* should be tempered by the fact that at day one, it had consequently decreased in the long fluctuations (> 48 hours periods), while it did not decrease in the N condition. This is not expected since at day 1, mutants in long periods have only experimented the N medium. Still, the results are convincing since we observe that *met12* is depleted very fast at shorter periods; with zero raw counts in the 3 shortest periods at the end of the experiment (Fig III-F.11a). By individually looking at other mutants with low wdev values in the methionine experiment, I observed that none had as high data quality and interesting phenotypes as *met12*.

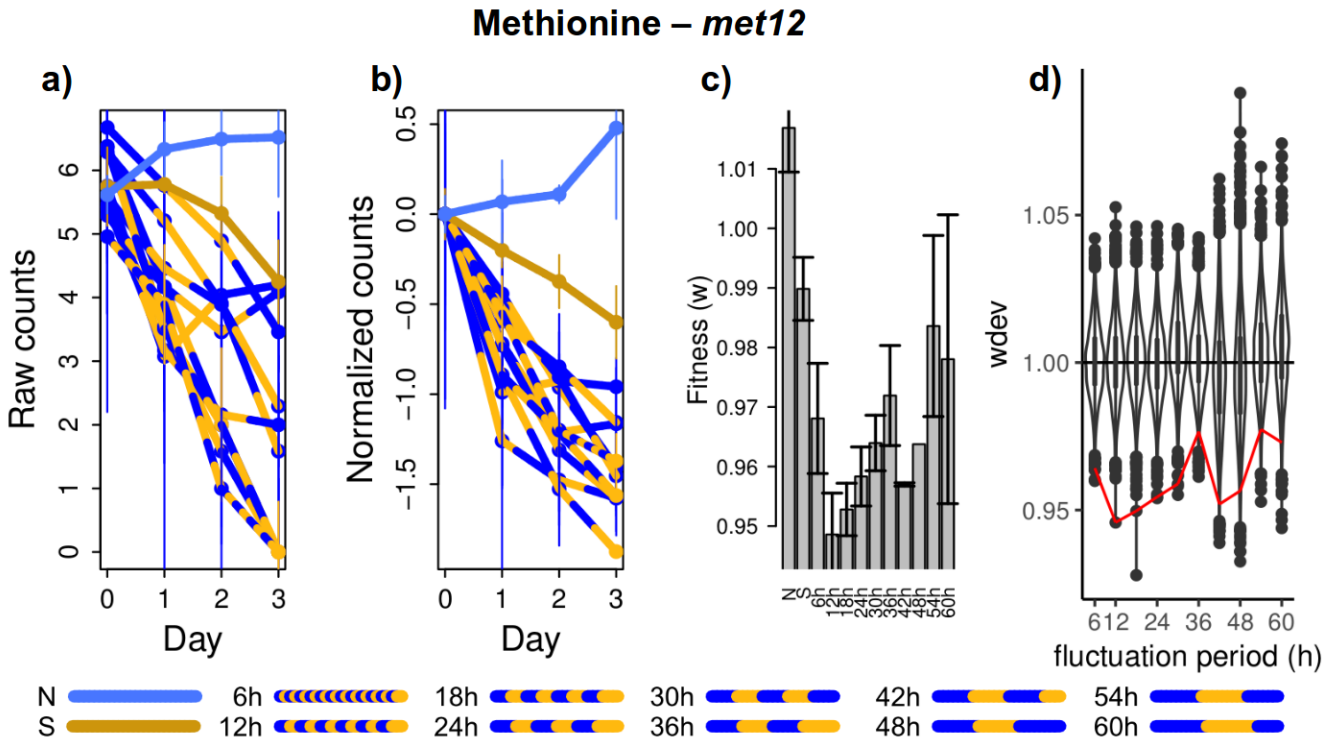


Fig III-F.11. *met12* in the Methionine experiment: a-b) median of log2(counts) (± s.d.), **c)** mean fitness (± s.e.m) at day 3, **d)** distributions of wdev values of all mutants, with those of *met12* in red.

IV Secondary Screen: single-mutant measurements in periodic salt stress

I chose to focus on validating results of the Salt experiment since it had a high data quality and displayed strong nonlinear effects. Secondary screen assays were performed only in the salt fluctuations and at the 6 hours period, since the strongest nonlinear effects were observed at this period.

A Flow-cytometry assays: fitness relative to Wild-Type

In the flow-cytometry or FACS assay wells were inoculated with 50% of wild-type cells and 50% of mutant cells (Fig IV-A.1), OD was not measured and samples were fixed with paraformaldehyde twice a day before flow-cytometry measurements (see the submitted publication for details).

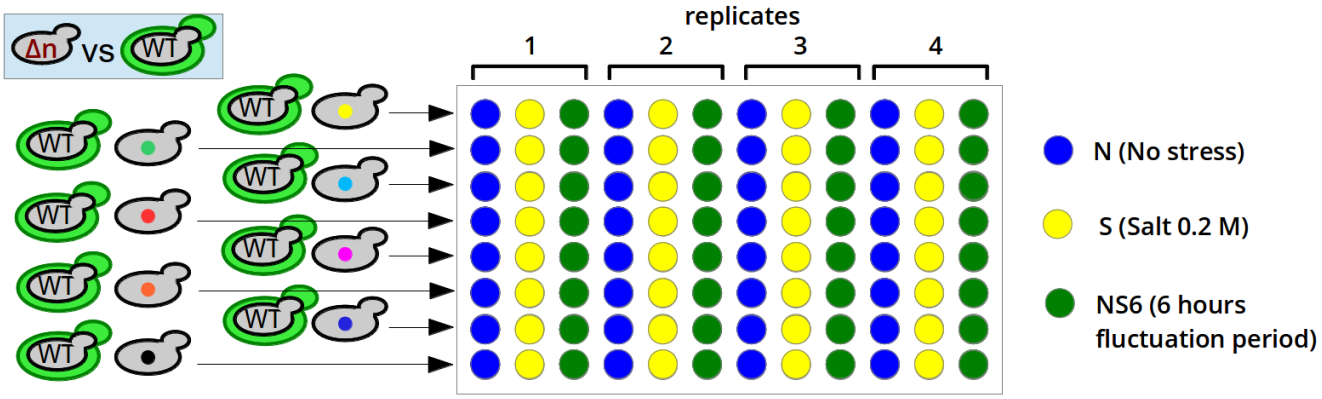


Fig IV-A.1. Design of one 96-well plate for the secondary screen Flow-Cytometry assay.

A correlation of 0.71 was observed between fitness from the BarSeq assay and from the flow-cytometry assay (submitted publication, Fig 1F). This revealed that despite the lower precision of the BarSeq assay, and the interactions between strains that can occur when grown in pool, most results from the BarSeq assay can be trusted. For instance, Figure IV-A.6a-d shows the BarSeq data of 4 mutants with very high wdev values at the 6 hours period of the Salt experiment (3 are in the top 5 highest values, and one is in the top 30). Those mutants had low growth in S, high growth in N and a higher growth at the 6 hours period. The effect of 3 of those mutants was confirmed by flow-cytometry assays: their growth in all conditions was qualitatively similar to the BarSeq assay (IV-A.2a,c,d,e,g,h). *yor029w* had qualitatively similar growth in N and in periodic stress, however, it had a high growth in S in the flow-cytometry assay (IV-A.2b,f).

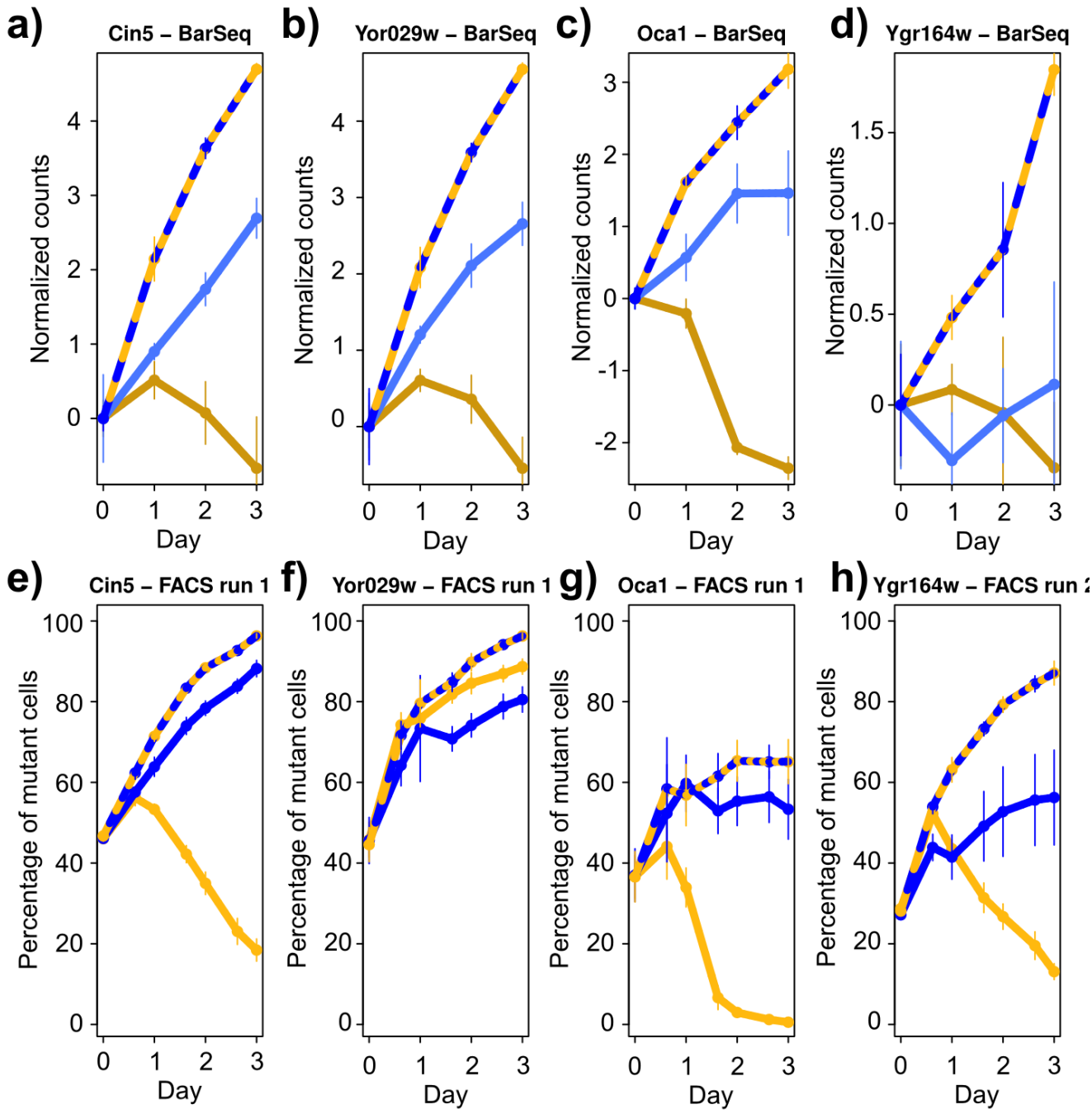


Fig IV-A.2. Data of four transgressive mutants across the 3 days of the BarSeq experiment (a-d) and FACS experiments (e-h). BarSeq log2(normalized counts) were adjusted to zero at day 0. FACS data: frequencies of the mutant strains vs the wild-type strain. Run 1 and 2 indicates two different flow-cytometry experiments. All plots show median values ± *s.d.*.

B Absolute fitness

i Optical density measurements suggest that absolute fitness may not be transgressive

I set up a secondary validation protocol where mutants were grown individually, and their growth was measured by optical density. Because a sunrise plate reader (Tecan) is available on the robotic platform, this protocol was simple and fast.

The design of the experiments was the following: one strain was tested per row of a 96-well plate, with four replicates per condition (Fig IV-B.1). Replicates of the same condition (either N (No stress), S (Stress) or NS6 (fluctuation 6 hours)) were separated across the row in order to avoid biases related to the localisation of wells within the plates. Four wells per run were handled in parallel in the robotic station. Thus, a total of 32 strains were tested in one experiment. I performed two such experiments: the first lasted one day, and the second lasted 3 days. The wild-type strain was not included in the first experiment, while one row of each well was dedicated to the wild-type strain in the second experiment. A total of 45 different mutants were tested, and 16 of them were tested in both experiments.

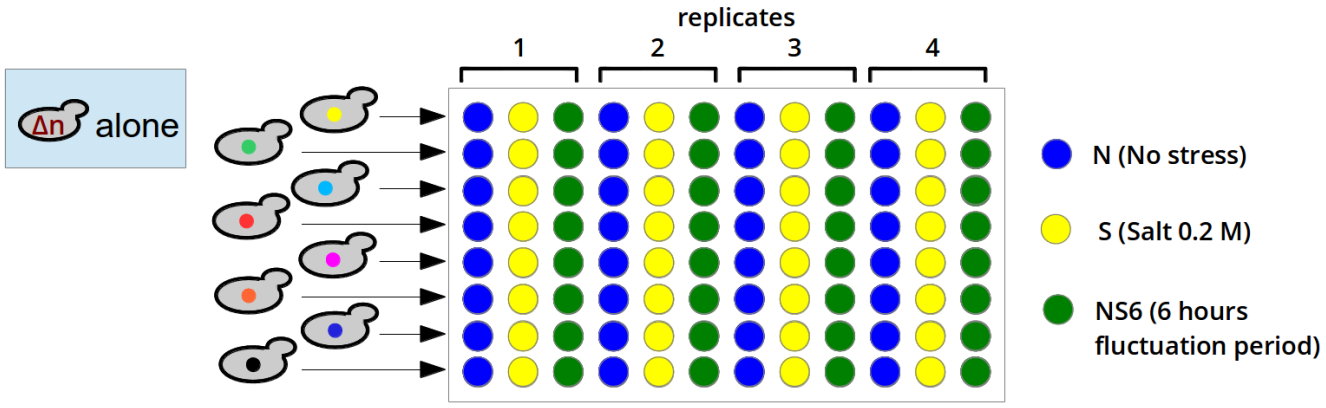


Fig IV-B.1. Design of one 96-well plate for the secondary screen OD assay.

In the first experiment, OD was measured once at every fluctuation event: before dilution (right after taking plates out of the incubator). In the second experiment, OD was measured twice at each fluctuation: once before dilution and once after dilution. The aim was to estimate growth of cultures in the incubator independently of dilution: by comparing OD of cultures entering and leaving the incubator with no bias due to variability in dilutions across wells. In the middle of the second experiment, I decided to make a slight modification to the protocol (just before the fluctuation at 36 hours). This modification consisted in mixing the cultures in all wells before measuring OD in the plate reader. I thought that it could improve the accuracy of the experiment by having more homogeneous solutions. As a result, OD values increased in the second part of the experiment (Fig IV-B.3 to IV-B.5), with the maximum going from ~ 1.1 to ~ 1.3. However, variability between replicates increased, especially for the wild-type in periodic conditions (Fig IV-B.5). Therefore, it is not clear if data quality is higher in the first or second half of the experiment.

Results of those experiments showed no mutant with a strong transgressive phenotype (Fig IV-B.2 to IV-B.5). Only *oca1* had a slight transgressive winner phenotype in the second experiment. I was quite disappointed at the time, and I was worried about the quality of the BarSeq data. However, I later realised that the data quality in both experiments are good (submitted publiction, Fig 1F), but I was comparing two different values: relative fitness and absolute fitness.

Results from the OD experiments revealed that most of the tested strains had nonlinear winner phenotypes; with a growth in the fluctuating condition that was as fast as in the no-stress condition. For instance, Fig IV-B.2 to IV-B.5 show the OD data of 4 mutants with very high wdev values in my BarSeq assay (same mutants as in IV-A.2). This indicates those mutants are non-linear winners but not transgressive in terms of absolute fitness values (at the exception of *oca1* that is slightly transgressive). If we want to see if those absolute data fit with our relative data from the BarSeq experiment, we can compare them to data of the wild-type in the second experiment. Unfortunately, there was noise in the data from the 16 replicates of the wild-type strain in the second experiment. Especially, as mentioned above, there was a great increase in variability of the replicates for the fluctuating condition in the second half of the experiment. In the first half, fitness of the wild-type seemed to be either linear or slightly nonlinear loser. This could explain the pronounced differences between wild-type and mutant strains in fluctuating conditions.

To conclude, those results suggest that interesting informations about absolute growth rate can be measured with a simple assay using a plate reader. Importantly, those experiments revealed that most mutants with high wdev values have a transgressive relative fitness phenotype but not an absolute transgressive fitness phenotype. indeed, out of 45 mutant strains tested for absolute fitness, only one (oca1) had a slight transgressive winner phenotype.

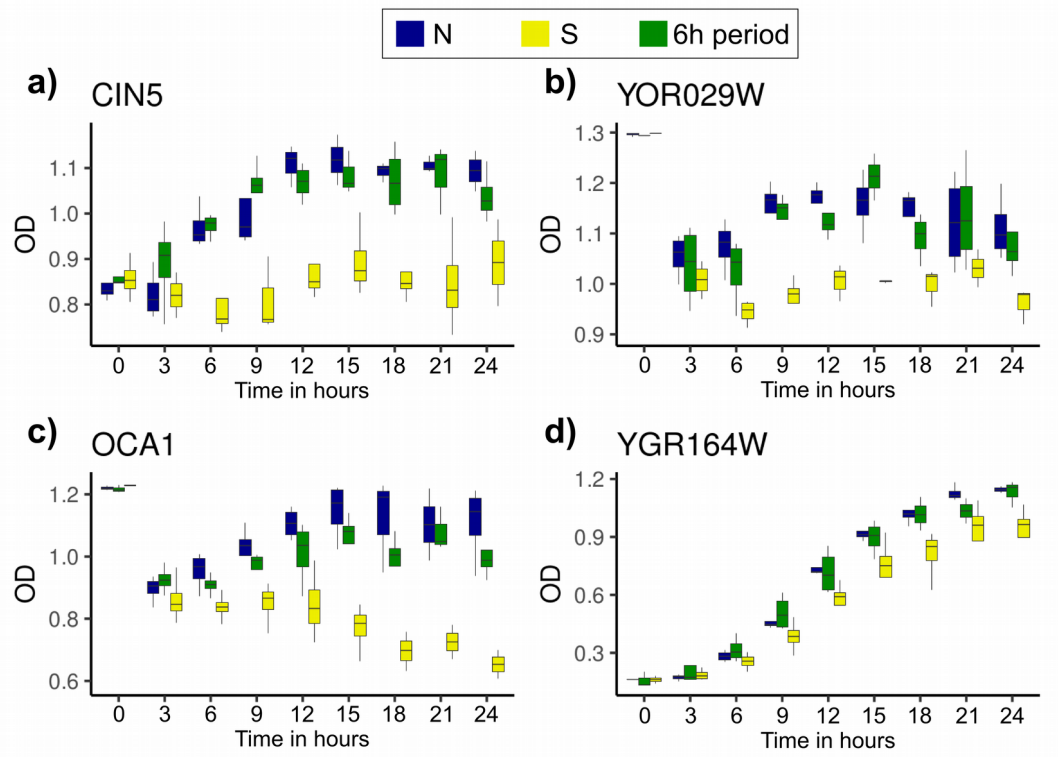


Fig IV-B.2. OD values (four replicates) before dilution of four mutants for the one-day experiment.

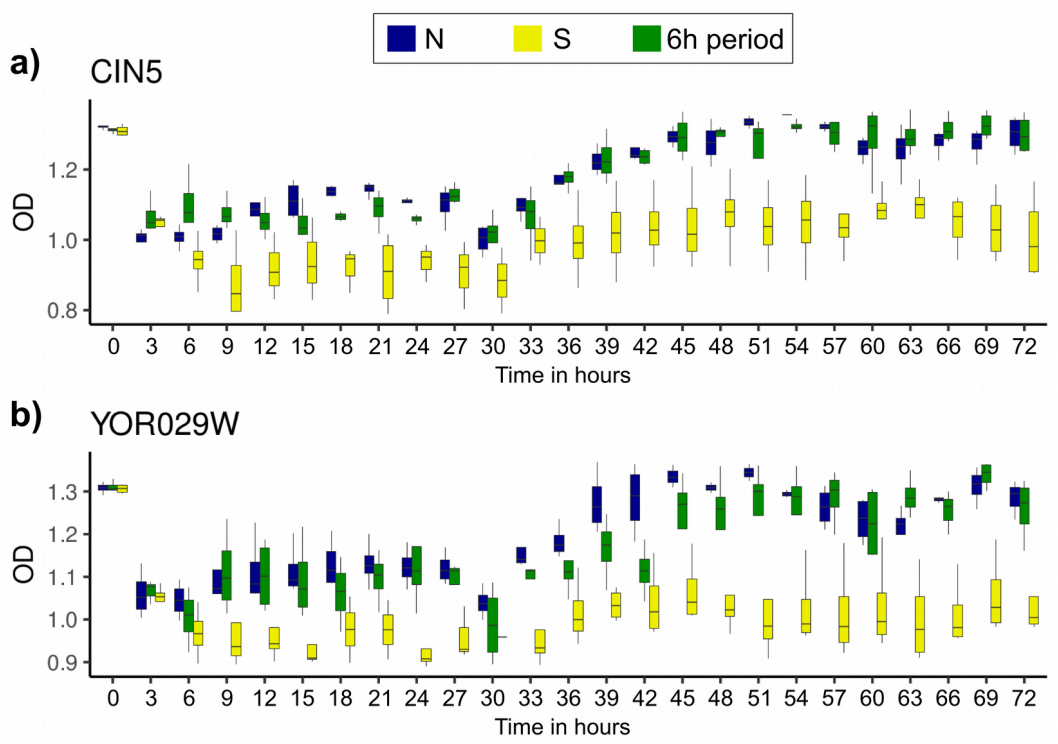


Fig IV-B.3. OD values (four replicates) before dilution of *cin5* and *yor029w* for the three-days experiment.

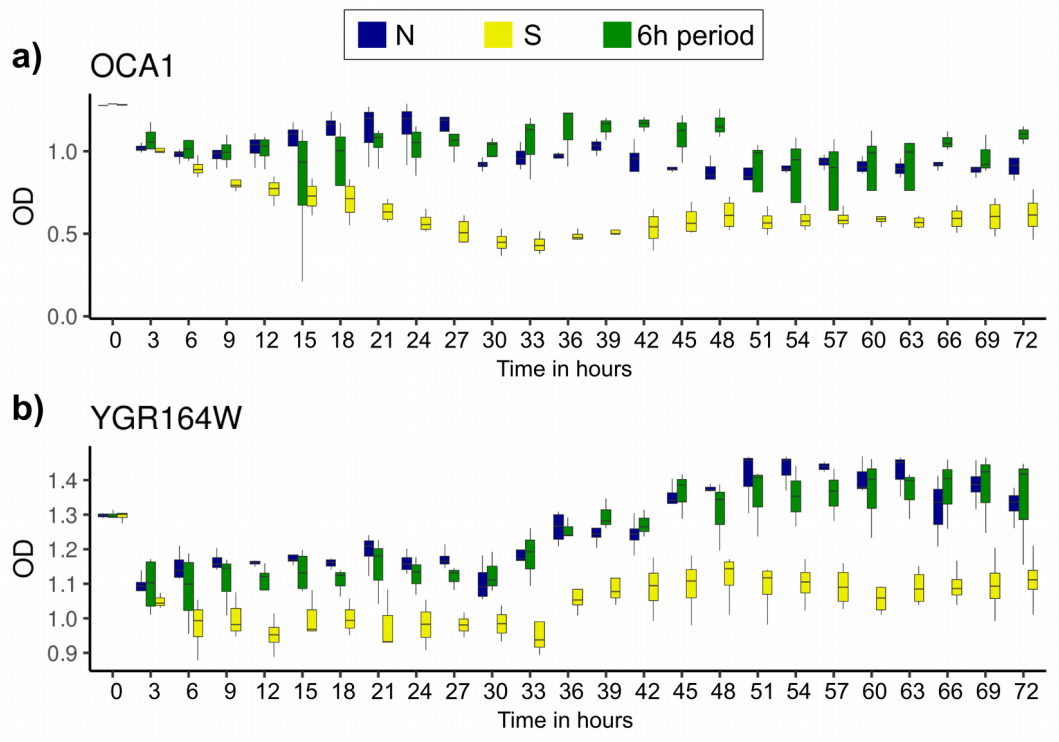


Fig IV-B.4. OD values (four replicates) before dilution of *oca1* and *ygr164w* for the three-days experiment.

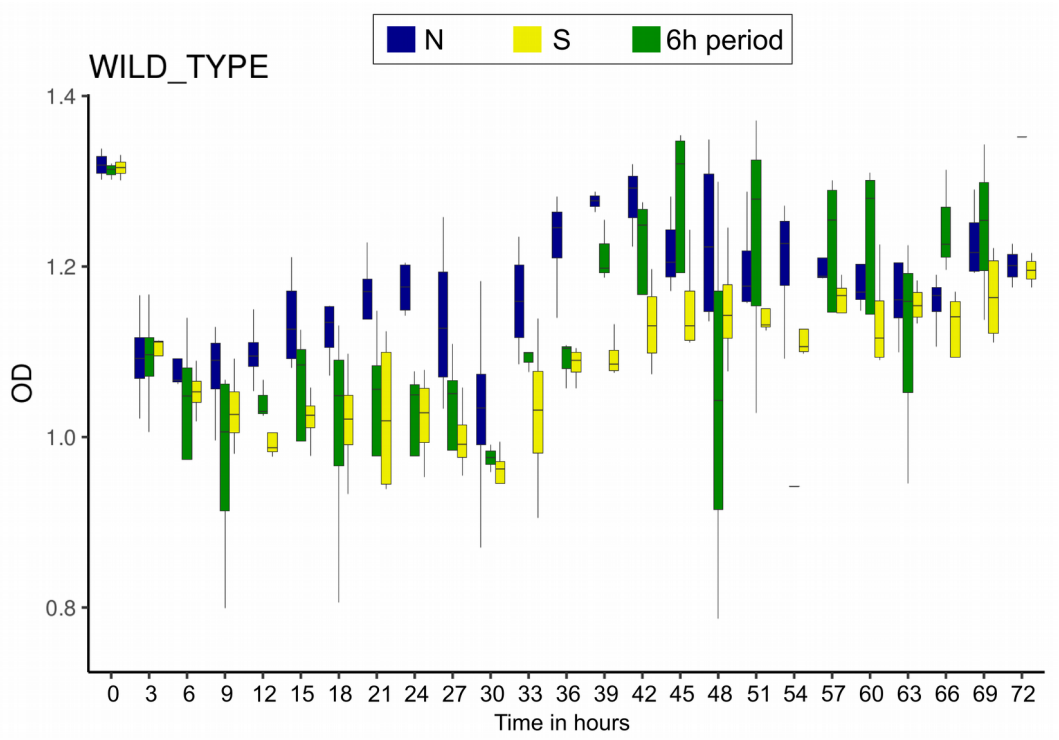


Fig IV-B.5. OD values (16 replicates) before dilution of the wild-type strain for the three days experiment. One row (four replicates per condition) was dedicated to the wild-type strain in each of the four plates.

ii Absolute fitness estimated by flow-cytometry

Analysis of the OD experiment described above suggested that some mutants have a transgressive effect only in terms of relative fitness but not in terms of absolute fitness (with the exception of *oca1)*. I therefore re-analyzed data of flow-cytometry experiments, in order to address this point. Computing absolute fitness values from FACS data was possible since the FACS mixed several times vigorously each well, before aspiration at a constant speed. The rate of sampling was 1µl per second. Thus, I estimated the cellular concentration in each well by dividing the number of cells sampled by the time of sampling. I estimated the theoretical volume of cultures that would have been produced if there was no dilutions (which is exponential, see methods). The total number of cell at each fixation was estimated by multiplying the cellular concentration by the theoretical volume. I could then determine those numbers for both wild-type and mutant cells. Finally, a linear model was fit on: log2(cn) ~ time, with time in hours. The doubling time was then computed as the inverse of the slope coefficient of the model.

In the flow-cytometry experiment, I always included at least 9 wells dedicated to wild-type controls: the wild-type kanamycin strain (k2) alone, the wild-type kanamycin GFP strain (gfp2) alone, and a competition between k2 and gfp2. These controls allowed to check if the cut-off for GFP thresholds were appropriate, and to verify if there was no growth alteration associated with GFP in any condition

(methods).

On Fig IV-B.6 we can see the doubling time (DT) of the wild-type strain gfp2 that was cultivated alone in different experiments/runs, as well as the DT of mutants cultivated in competition with the gfp2 strain. The wild-type strain had small DT differences between conditions (Fig IV-B.6a-c). It has been used for all competition experiments. We can see that it had a higher DT in salt than without salt. The DT in the fluctuating conditions was slightly superior to the salt condition. Comparatively, the mutants had lower DT than the wild-type in the N condition (by ~ 0.15), and significantly higher DT in stressful conditions (Fig IV-B.6d-f). In fluctuating conditions, those mutants had DT slightly lower than their DT in the N condition. Thus, there is a very slight transgressive winner effect in absolute fitness. Those data are consistent with the OD data and could explain why we observe a transgressive effect in fluctuating conditions. Indeed, the differential in DT between the gfp2 strain and the mutant strains is slightly higher in fluctuating conditions than in the N condition.

However, these experiments had two limits to compute absolute growth rates: strains grew in competition, and, more importantly, the rate of sampling was high and cell counting may have been imprecise. 10,000 cells were acquired for each population. About 3,800 cells per well were then gated for further analysis (methods of the submitted publication). Cells were acquired in the flow-cytometer at a rate of ~1,000 cells per second. Thus, the time of acquisition of a well was about 10 seconds. Resolution is limited since the FACS does not record time units shorter than a second.

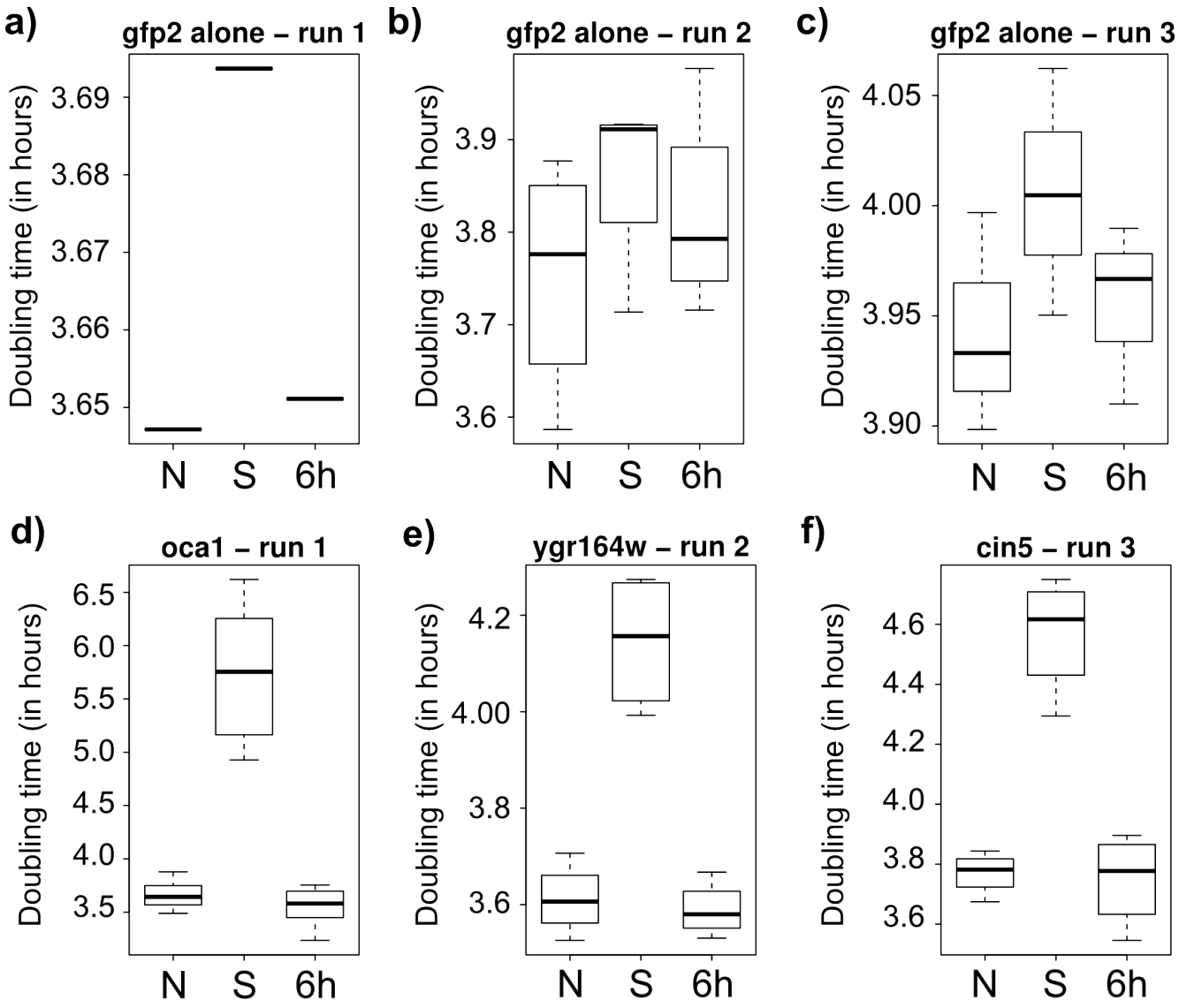


Fig IV-B.6. Doubling time of wild-type **and mutant strains during flow-cytometry assays.** Stress indicates salt 0.2 M and fluctu indicates the 6 hours periodic condition. gfp2 is a wild-type strain bearing a GFP-expressing cassette and a kanamycin-resistance cassete, that was used in all competition experiments. The y axis indicates the doubling time in hours. Number of replicates per condition per plot (from left to right, from top to bottom): 3, 4, 3, 1, 4, 3, 8, 4, 4. The run indicates the flow-cytometry experiment from which data originates.

Therefore, I performed another one-day experiment, in order to take into account those issues and to obtain clearer results on absolute growth rates. In this experiment, different strains were cultivated alone, and the plates were diluted about three times more than in previous experiments. This resulted in a median sampling rate of ~350 cells per second, and a median time of acquisition of 28 seconds per well. Results are shown in Fig IV-B.7 for the *trm1* mutant and a wild-type strain. As expected, the *trm1* mutant had a trangressive winner phenotype when grew in competition with the wild-type (Fig IV-B.7a). Importantly, I found that the wild-type strain alone was linear or slightly nonlinears loser (Fig IV-B.7c). Importantly, the *trm1* mutant alone was slightly transgressive winner in fluctuating conditions

(Fig IV-B.7b). This suggests that some strains can have a higher absolute fitness in binary periodic fluctuating conditions than in both separate steady conditions.

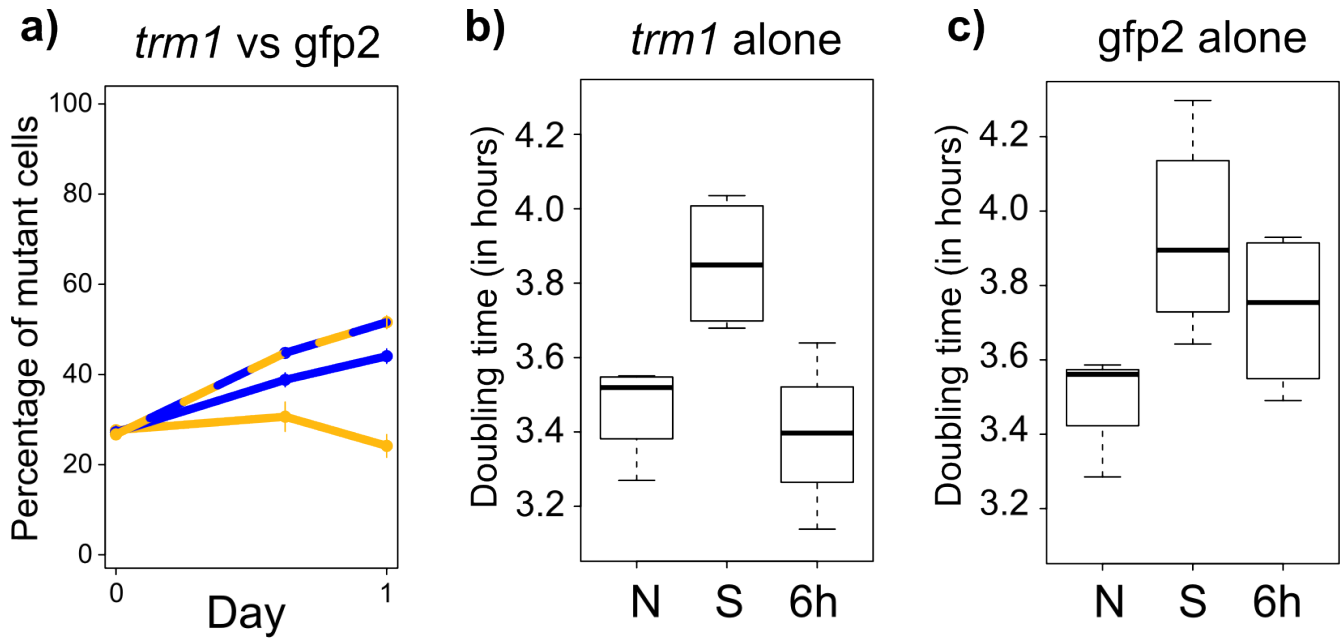


Fig IV-B.7. Absolute fitness of a transgressive winner mutant. a) frequencies of *trm1* vs the *wild-type* strain over time. Values: median ± *s.d.*. Color code: salt 0.2 M in yellow, N in blue and periodic 6 hours in alternating blue and yellow. **b)** doubling times of the *trm1* mutant cultivated alone. **c)** doubling times of the wild-type strain gfp2 cultivated alone.

C Negative and non-reproducible results

i hog1 mutant showed non-reproducible results across experiments

Three strains displayed inconclusive results in the flow-cytometry assays: *hog1*, *ire1* and *hac1*. *ire1* and *hac1* had frequencies very low at the beginning of the experiment (below 10%) and were depleted immediately after (data not shown). *hog1* showed contradictory results (Fig IV-C.1). After my first flow-cytometry experiment, I was excited to observe that I could perfectly validate the nonlinear winner phenotype of the *hog1* mutant at the 6 hours period of fluctuation. However, in the two following experiments I obtained very different results for *hog1*: with a seemingly linear phenotype in fluctuating conditions. After the third experiment, we were not able to conclude on the effect of *hog1* . Importantly, other mutants were tested in different runs (run 1 and 3 or run 2 and 3), and resulted in highly reproducible transgressive phenotypes (cin5, pde2, trm1, tom7, oca1, ygr164w). Additionally, the gfp2 strain showed reproducible results across experiments (Fig IV-B.6). For now, it remains a

mystery why results for *hog1* were not reproducible.

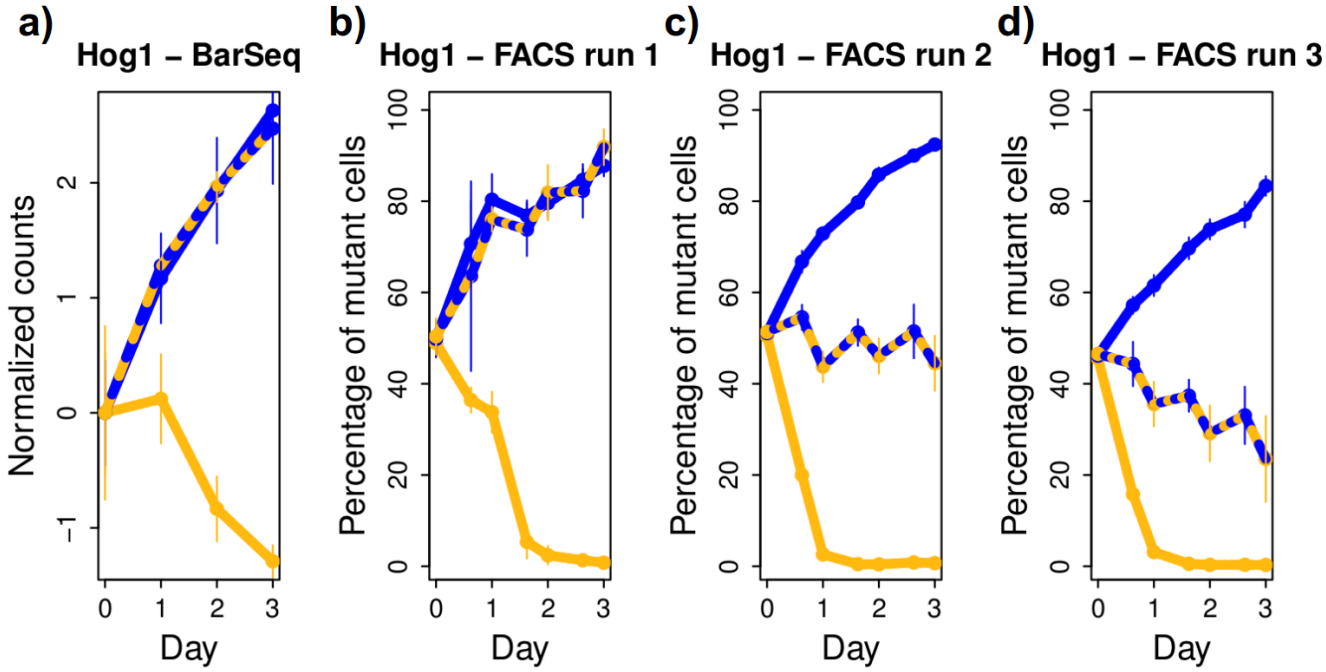


Fig IV-C.1. Plots of the *hog1* mutant: in the BarSeq experiment **(a)** and in different runs of the flow-cytometry experiment **(b-d)**. Values: median ± *s.d.*.

ii Haploid strains were not transgressive

Haploid strains were constructed for the main transgressive winner mutants and analyzed by flow-cytometry. The motivation behind this work was to see how robust my results were to ploidy levels. Moreover, if results were similar to diploids, working on haploid strains would have simplified later molecular biology experiments. Haploid mutants were in competition against a GFP-tagged haploid wild-type strain. However, none of them showed an important transgressive winner phenotype (Fig IV-C.2). And only *pde2* and *cin5* had slight transgressive winner phenotype (Fig IV-C.2c,h). *hog1* had an interesting nonlinear loser phenotype (Fig IV-C.2g). However, it was fully depleted in salt after just one day, which makes it hard to really conclude about its linearity.

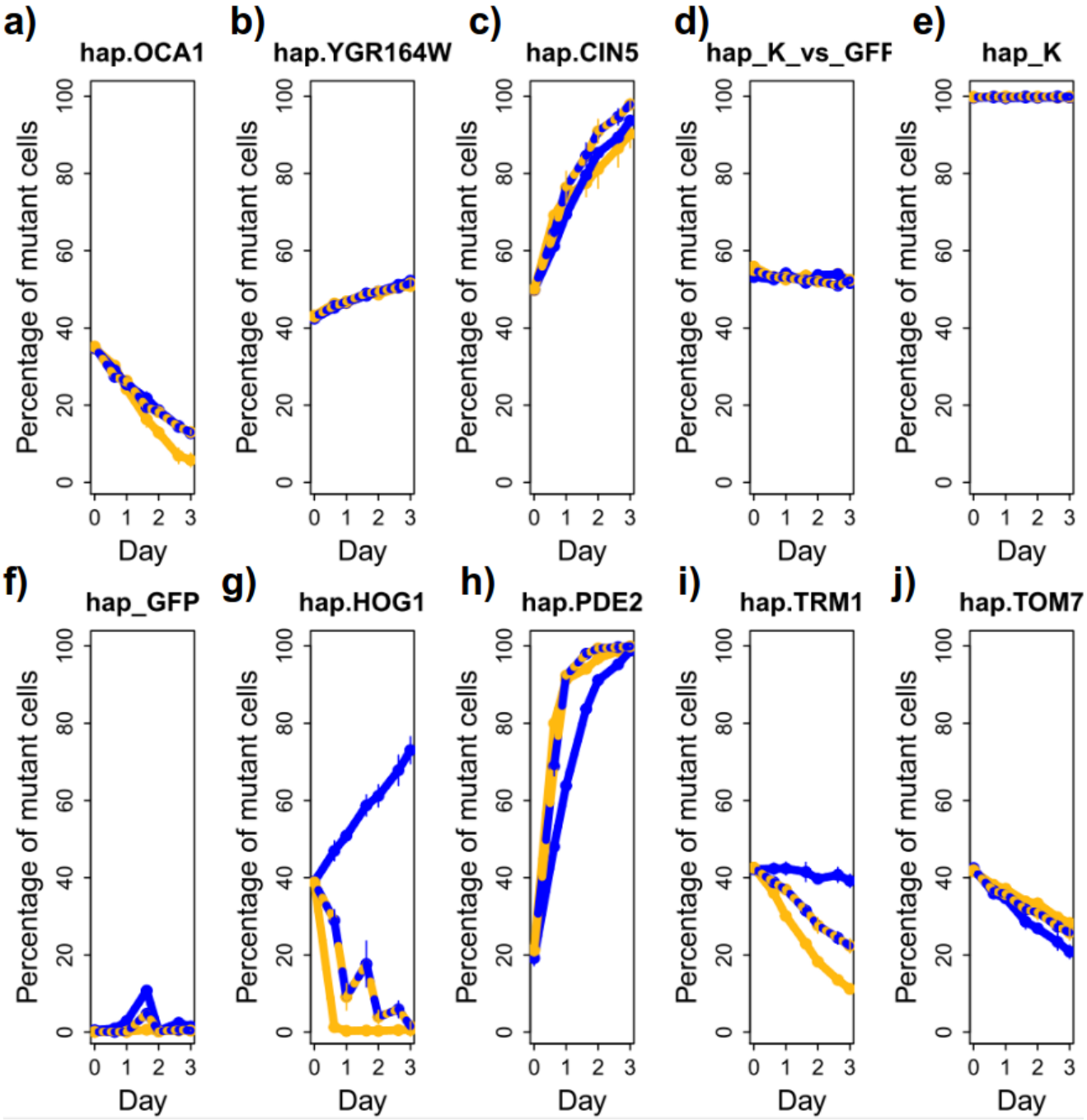


Fig IV-C.2. Null haploid mutants and wild-type haploid controls in a flow-cytometry experiment. K is a haploid wild-type strain with a kanamycin-resistance cassette, and GFP is a haploid wild-type strain with a kanamycin-resistance cassette and a GFP-expressing cassette. **a-d)** and **g-j)** competitions of one strain agains the GFP strain. **e)** K strain alone. **f)** GFP strain alone.

iii Some complementations did not rescue the transgressive phenotype

Complementation assays of null mutants allow to determine if the phenotype observed is caused by the lack of the gene product, or if it is due to other factors such as secondary mutations, perturbed cis-regulations or aneuploidies. Several complemented strains were constructed. So far 2 out of 5 showed successful complementation (*pde2* and *tom7*, see submitted publication).

TRM1 encodes for a tRNA methyltransferase. *ygr164w* codes for a hypothetical protein. Both had strong transgressive winner phenotypes at the 6 hours period (Fig IV-C.3a,c). A tRNA (tR(UCU)G2) is present within the sequence of YGR164W. We thought that regulation of tRNAs might be involved in the strong transgressive winner phenotypes of the *trm1* and the *ygr164w* mutants. Thus, we constructed a *ygr164w* strain that was complemented only for the missing tRNA. However, both the *trm1* complemented strain and the *ygr164w* tRNA-complemented strain still had a transgressive winner phenotype (Fig IV-C.3d,f). Complementation of the *cin5* mutant strain also failed to rescue the transgressive phenotype (Fig IV-C.3e). Importantly, I did not verify that the complementation was functional: that the complemented genes did not contain any error in their DNA sequence after insertion, and that their gene product was express at standard levels. Thus, I cannot conclude that the function of the null gene is not important for the phenotype of the strains that failed to complement.

The complemented strain of the *srf1* and *hal9* mutants were constructed and tested as well. However, the robot experienced a collision, and the experiment was aborted after one day. At this time point, those mutants have no visible transgressive phenotypes, on the contrary of *cin5, trm1 (*Fig IV-C.3b,c), *pde2* and *tom7* (submitted publication Fig 4F,G). Thus, I could not conclude about the involvement of the SRF1 and HAL9 genes in the transgressive winner phenotype of the *srf1* and *hal9* mutants.

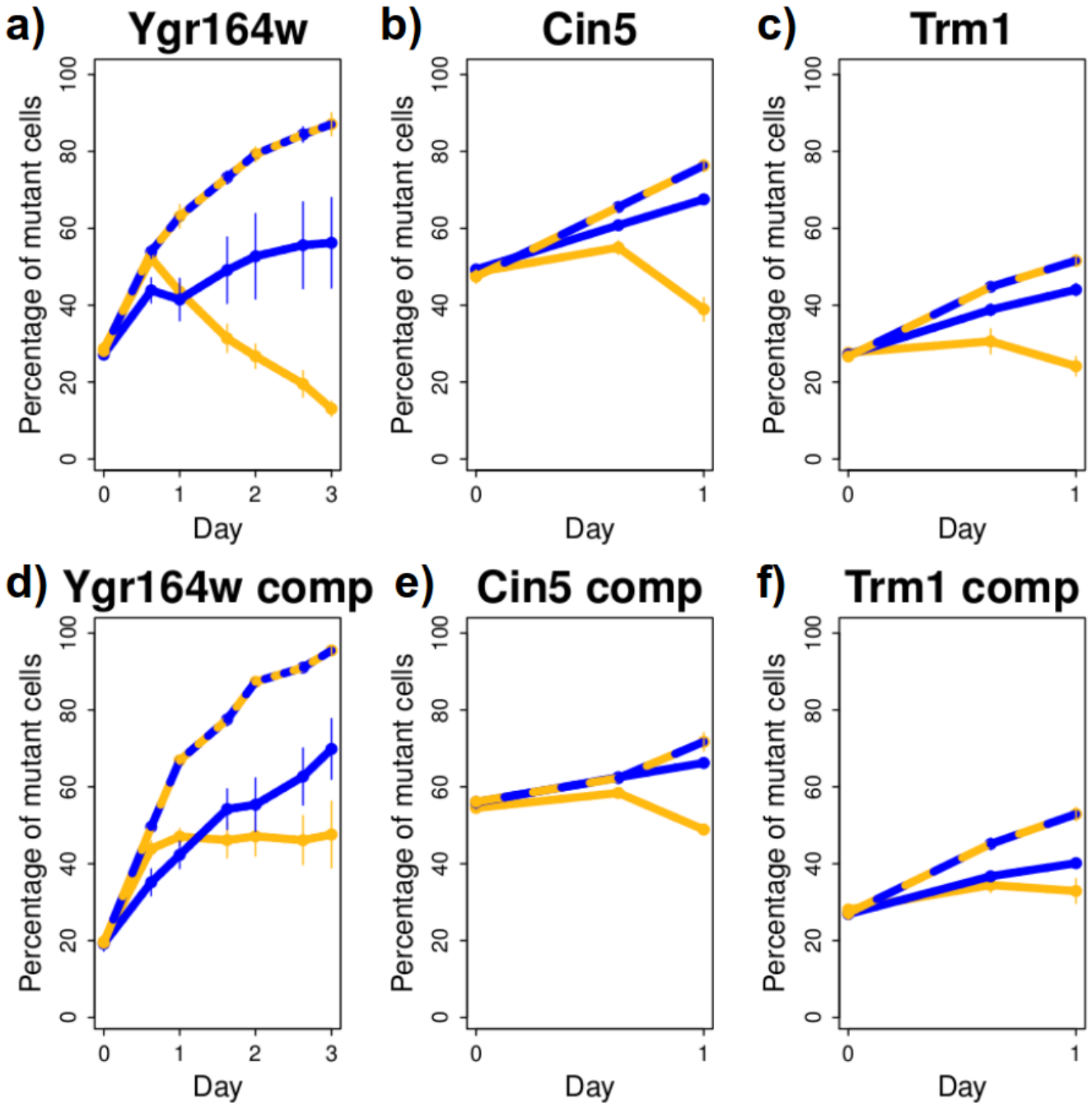


Fig IV-C.3. Flow-cytometry experiment of 3 complemented strains: *ygr164w* **(a)**, *cin5* **(b)**, *trm1* **(c)**, *ygr164w* complemented by the tRNA tR(UCU)G2 **(d)**, *cin5* complemented **(e)**, *trm1* complemented **(f)**.

D Generality of transgressive effect for the *pde2* and *tom7* mutants in other environmental fluctuations

I made two experiments to determine if two validated mutant strains (*pde2* and *tom7*) showed a transgressive behaviour in other conditions as well. However, *pde2* had a nonlinear winner phenotype (run 4, Fig IV-D.1c), or a slightly transgressive phenotype (run 5, Fig IV-D.2b) but no strong transgressive winner phenotype at day 1 as observed in previous experiments (submitted publication Fig 4B). This could be because of a slightly reduced salt concentration, because during the run 4, phenotypes were as expected in the NaCl 0.4 M condition (Fig IV-D.1a,e). The transgressive winner effect of *tom7* was similar in both NaCl 0.2 M and NaCl 0.4 M (Fig IV-D.1e,g).

Interestingly, *tom7* had a transgressive winner phenotype visible at day 1 in raffinose periodic stress (Fig IV-D.1h). Raffinose, contrarily to glucose, favors mitochondrial respiration. Tom7p is involved in mitochondrial protein import (Neupert and Herrmann, 2007). Thus, increased levels of respiratory growth may drive the transgressive winner phenotype of the *tom7* mutant in salt periodic stress. In contrast, *pde2* had a nonlinear loser phenotype in raffinose periodic stress (Fig IV-D.1d). Thus, its growth advantage in salt may be due to increased levels of fermentative growth.

I wanted to determine if fitness is transgressive only in osmotic stress or in ionic stress or in both. I therefore tested the following stressful conditions: KCl 0.2 M, LiCl 0.2 M, NaCl 0.2 M and sorbitol 0.4 M (run 5). Unfortunately, *pde2* had a slight transgressive winner effect in salt 0.2 M (Fig IV-D.2b). It is therefore difficult to know if transgressivity is specific to NaCl. Fitness of *pde2* was nonlinear in both osmotic and ionic stress, but it was not transgressive. The pattern for *tom7* was different: fitness was slightly transgressive at day 1, it was nonlinear but not transgressive in sorbitol, and it was quite linear in ionic stress. Thus, for *tom7*, osmolar stress seemed to induce more nonlinear effects than ionic stress. While for *pde2* both osmolar stress and ionic stress induced nonlinear effects. Finally, only *tom7* was transgressive in another condition (raffinose 1.9%). Thus, the transgressive winner fitness of those two mutants was not generally pleiotropic.

Run 4

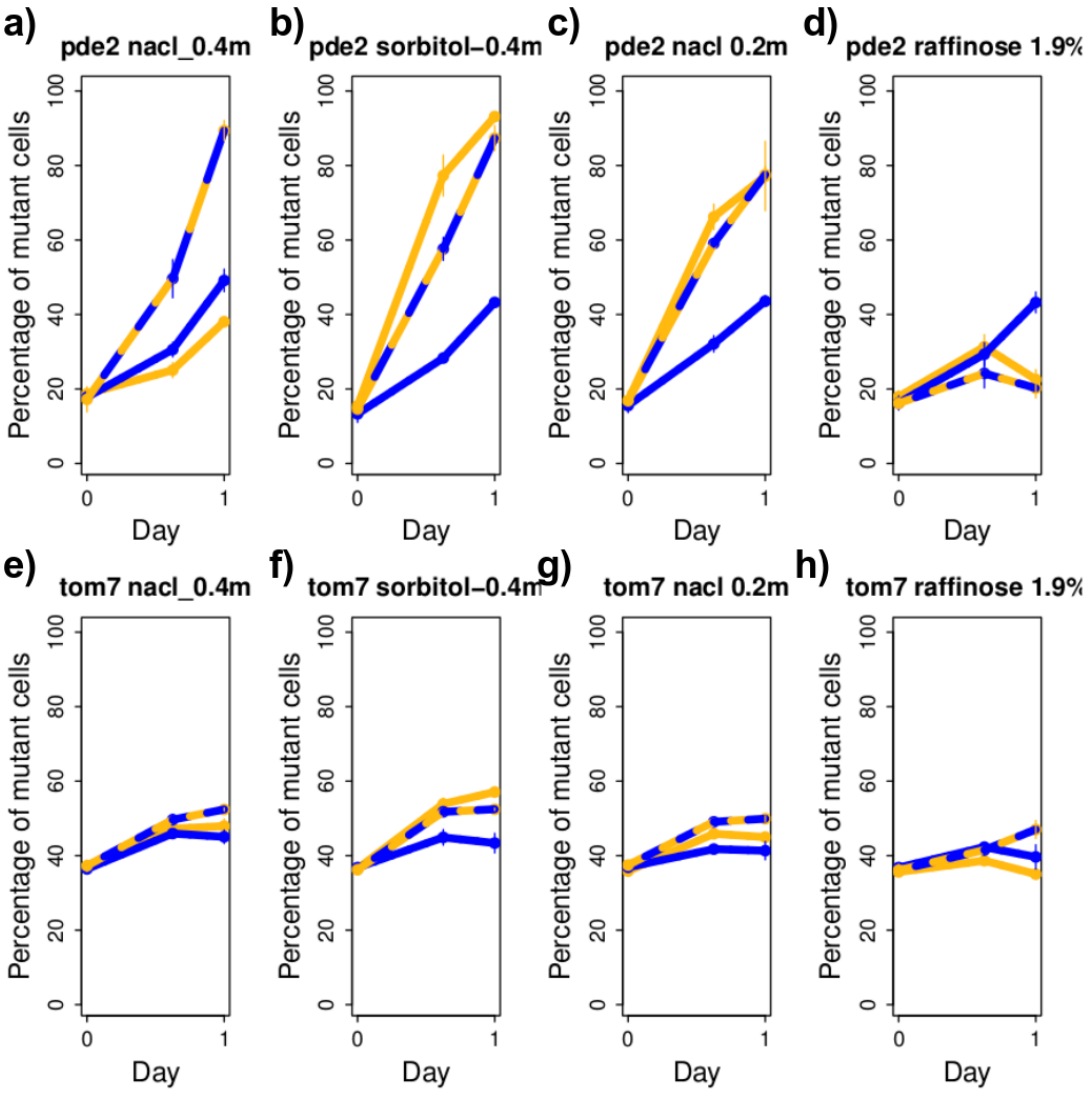


Fig IV-D.1. Flow cytometry experiment (run 4) of the *tom7* and *pde2* mutants in various media fluctuations. Color code: N (blue), S (yellow), 6 hours period (blue-yellow). In the raffinose medium glucose 0.1% was used instead of glucose 2%.

Run 5

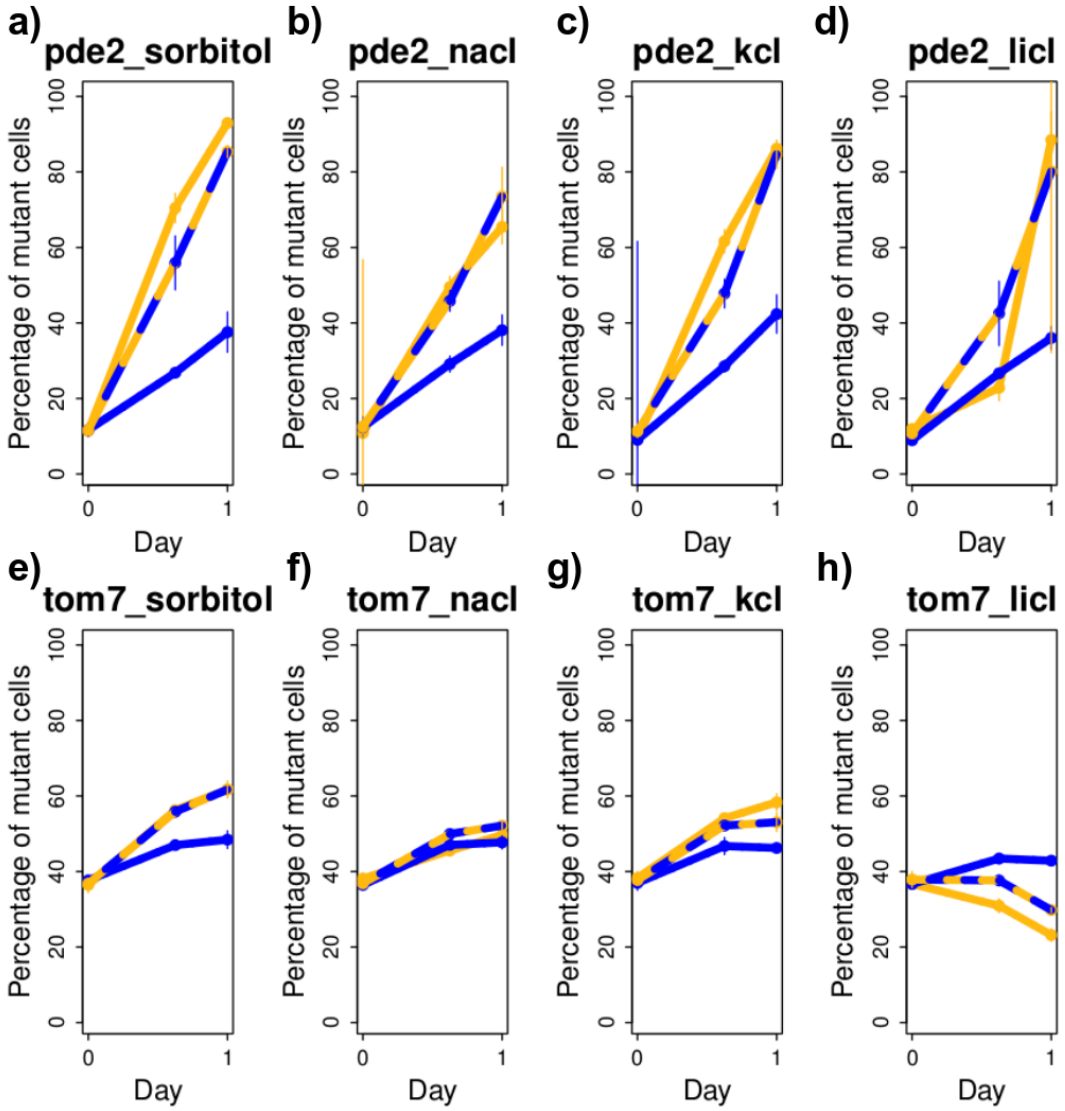


Fig IV-D.2. Flow cytometry experiment (run 5) of the *tom7* and *pde2* mutants in various media fluctuations. Color code: N (blue), S (yellow), 6 hours period (blue-yellow). Concentration tested are: a,e) Sorbitol 0.4 M, b,f) NaCl 0.2M, c,g) KCl 0.2 M and d,h) LiCl 0.2 M.

V Analysing co-fitness data to group mutants with similar phenotypic profiles

I have explored various available datasets and tools in order to try to infer the mechanistic basis of the nonlinear phenotypes. In order to infer the function important for the phenotypes of nonlinear mutants, I was hoping to find datasets where some nonlinear mutants share common characteristics. A type of dataset that I found particularly interesting was datasets of correlation of fitness (co-fitness) of mutants in many conditions (Hillenmeyer et al., 2008). In those datasets, a library of yeast mutants is exposed to hundreds of conditions. It was showed that mutants that have high correlations of fitness across those

conditions share biological functions and Gene Ontology terms (Hillenmeyer et al., 2008). Another study showed that co-fitness data can be successfully used for functional inferences (Hoepfner et al., 2014). My reasoning was that if some nonlinear mutants clustered by their values of co-fitness, then they may be involved in the same pathway/function that would be key for their nonlinear phenotype.

I have looked at 3 databases of co-fitness data where: 1144 (fitDB) (Hillenmeyer et al., 2008), 3258 (HIPHOP) (Lee et al., 2014) and 1800 (FMI) (Hoepfner et al., 2014) chemicals were screened against the whole homozygous YDL. On the FMI database, in general, correlations of 0.8 or above indicate mutants for genes involved in the same molecular complex, or that are directly interacting within the same pathway. For instance, in Fig V.1, we can see that *hog1* is correlated at 0.94 with *pbs2*. Indeed, both are crucial for the activation of the Hog pathway, and are directly interacting together. Upstream regulators *ssk2* and *ssk1* are the next highly co-fit strains, with a correlation of ~0.55. This reduced correlation can be explained by the fact that those genes are involved in only one out of two branches of the Hog pathway (the Sln1 branch). After *ssk1*, co-fitness values drop below 0.5. For the rim101 pathway there are about 10 genes with a correlation above 0.85. Those genes are all key genes of the Rim101 pathway, as shown in the supplementary Figure 1 of my submitted publication. Then, the correlation drops to 0.65 for genes not directly involved in the Rim101 pathway.

a)

Query	Interactor	CoFitHet	CoFitHom
HOG1	PBS2	0.42123	0.93919
HOG1	SSK2	0.1478	0.5655
HOG1	SSK1	0.21462	0.54467
HOG1	DCN1	0.01945	0.49339

b)

Query	Interactor	CoFitHet	CoFitHom
RIM101	RIM8	0.37404	0.89151
RIM101	RIM20	0.24742	0.88418
RIM101	RIM13	0.15517	0.87984
RIM101	RIM21	0.18535	0.87712
RIM101	RIM8_p	0.30591	0.86803
RIM101	RIM9	0.0514	0.86291
RIM101	YGR122W	0.14104	0.85825
RIM101	DFG16	0.17593	0.85766
RIM101	CHS5	0.12169	0.65227
RIM101	VPS28	0	0.64089

Fig V.1. Homozygous mutants most co-fit to *hog1* (a) or *rim101* (b) on fitDB.

I looked at the co-fitness score of mutants from well-known pathways in the three databases. Surprisingly, I found that the fitDB and the FMI databases were more accurate than the HIPHOP database at recovering other mutants of the pathways in the top hit, despite their lower number of tested molecules (data not shown). The quality of the experiments could explain this apparent discrepancy. Thus, I focused on the fitDB and the FMI databases for my analysis.

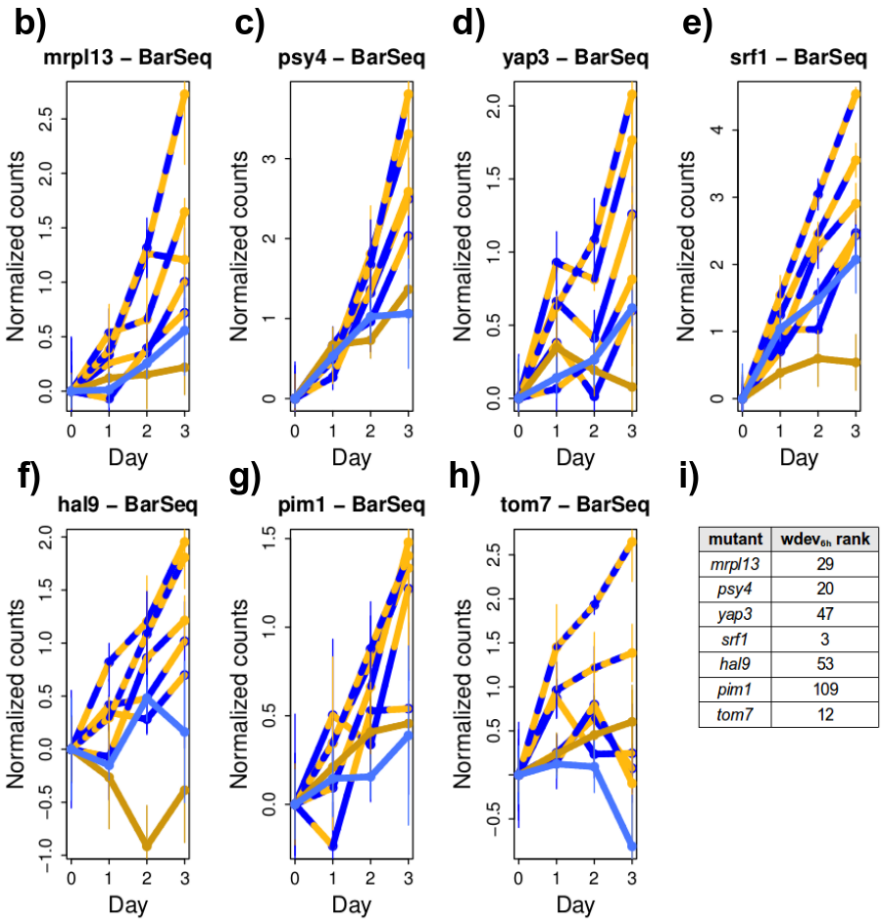
Importantly, data from those databases are very similar to my BarSeq data: they performed the same type of pooling assays with the same homozygous library. Thus, the same possible biases are expected, such as: interactions between strains, exchanges of metabolites within the pools or secondary mutations.

A *tom7* group: a potential common function related to mitochondria

Mrpl13p is a mitochondrial protein of the large subunit. The *mrpl13* mutant had a very strong nonlinear winner effect at the short periods of fluctuations (Fig V-A.1b). I wondered if I could use co-fitness data to infer why this mutant show a transgressive phenotype. We can see on fitDB that there are 6 mutants co-fit to the *mrpl13* mutant, with a correlation higher than 0.6. All mutants in this group have similar phenotypes in fluctuations: with few growth differences in steady conditions, and a strong growth advantage in fluctuating conditions, especially at fast fluctuations (Fig V-A.1). Those 7 mutants have all very high wdev values at the 6 hours periodic condition (Fig V-A.1) that rank among the top 109 highest values. I validated the transgressive effect of 3 (*srf1*, *hal9*, *tom7*) out of 6 of those mutants during the flow-cytometry experiment (Fig V-A.2). This means that the effect of at least 3 mutants (*mrpl13*, *pim1*, *yap3*) is not due to the function of their gene. The complemented strains for *srf1* and *hal9* were constructed, but are not tested yet. So far, only the deletion of the *tom7* gene has been shown to induce a transgressive winner effect. I can hypothesize that mutants in this group might share similar phenotypes due to perturbed mitochondrial functions (see discussion). Mutants that were not validated in the secondary screen assay might mis-regulate mitochondria through secondary mutations, aneuploies or perturbed cis-regulations. Since the *tom7* mutant has been complemented, below, I will refer to this group as the *tom7* group.

a)

Query	Interactor	CoFitHet	CoFitHom
MRPL13	PSY4	0.01946	0.77352
MRPL13	YAP3	-0.10731	0.76109
MRPL13	YDL133W = SRF1		0.72638
MRPL13	HAL9	-0.16201	0.71938
MRPL13	PIM1	0	0.67631
MRPL13	TOM7	-0.29131	0.65005
MRPL13	WSC2	-0.11678	0.58848
MRPL13	SHE1	-0.05033	0.58353
MRPL13	YBL083C	0.08114	0.56174
MRPL13	YKL033W-A	-0.12157	0.53437



i)

mutant	$wdev_{6h}$ rank
mrpl13	29
psy4	20
yap3	47
srf1	3
hal9	53
pim1	109
tom7	12

Fig V-A.1. *tom7* co-fitness group: a) homozygous mutants most co-fit to the *mrpl13* mutant on fitDB, **b-h)** BarSeq data of co-fit mutants. y-axis: normalized counts adjusted to zero at day zero (y-axis). **i)** ranking of co-fit mutants among all other mutants' wdev values for the 6 hours period.

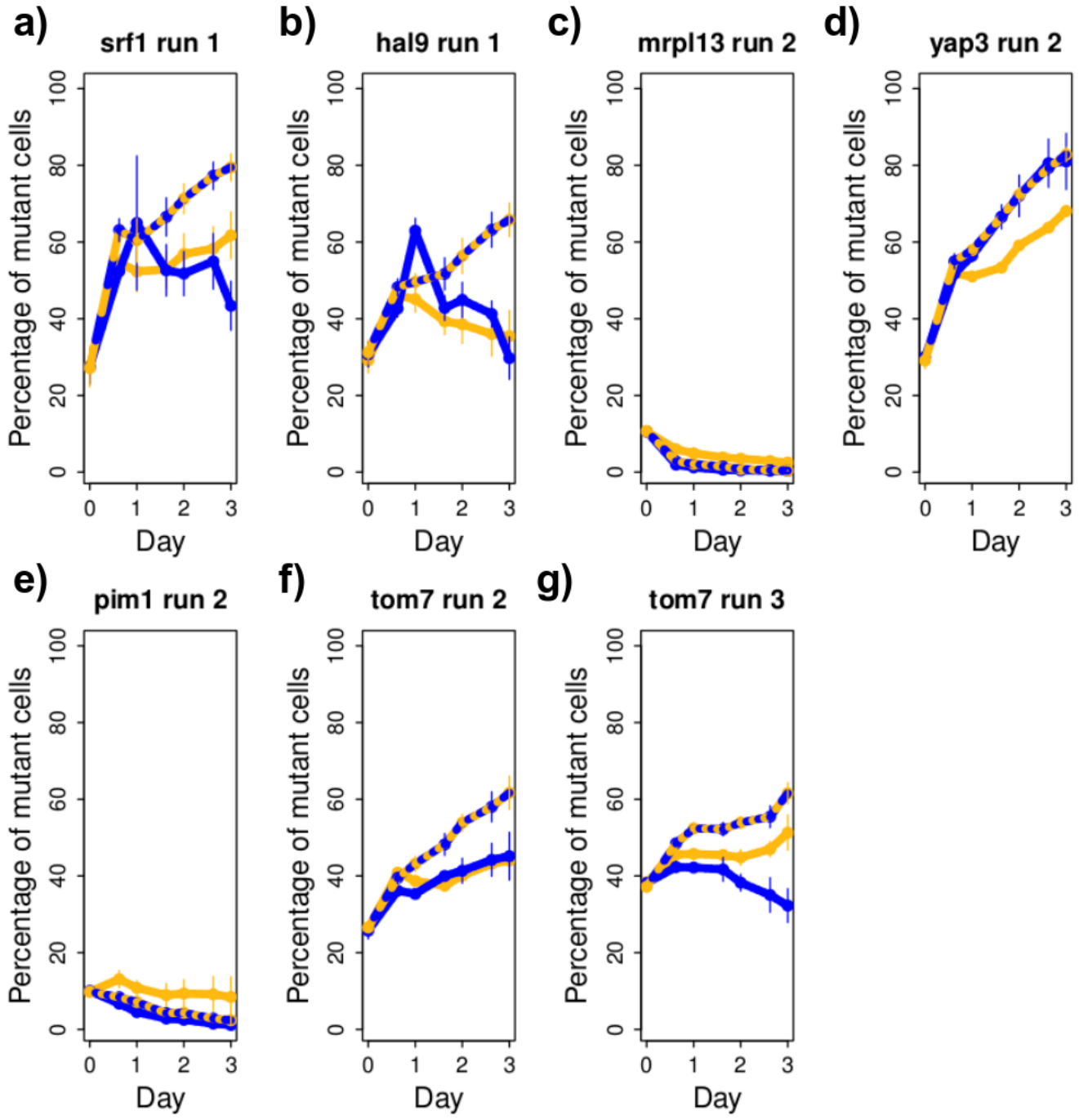


Fig V-A.2. *tom7* group flow-cytometry data.

B *trm1* group: no obvious cellular function

Qri5p is a mitochondrial inner membrane protein required for accumulation of spliced cox1 mRNA. The *qri5* mutant showed a very strong nonlinear winner effect at the short periods of fluctuations (Fig

V-B.1d). As above, I tried to infer the molecular mechanism of its nonlinear effect by looking at co-fitness databases. Three mutants (*trm1*, *ygr164w*, *ymr031w-a*) have a correlation with the *qri5* mutant higher than 0.6 on fitDB. The *rpl37b* mutant is missing on fitDB and also have a high correlation with *qri5*, on the FMI database. Those five mutants have a phenotype similar to mutants of the *tom7* group: few growth differences in steady conditions and a strong transgressive winner effect, especially at shorter periods (Fig V-B.1). Three other mutants (*mig1*, *mck1*, *ras2*) have a lower correlation with *qri5* on fitDB, and have a transgressive winner phenotype but with a different growth effect in steady conditions. All those mutants had very high wdev values at the 6 hours period (Fig V-B.1). In this group, *ras2*, *qri5* and *rpl37b* were not validated in the flow-cytometry assay (Fig V-B.1). *trm1* and *ygr164w* were validated in the flow-cytometry assay, but their complemented strains also showed a transgressive winner phenotype. Therefore, no evident cellular function can be attributed to this group yet. Since the *trm1* mutant showed a strong nonlinear phenotype in flow-cytometry assays (Fig IV-C.3c), below, I will refer to this group as the *trm1* group.

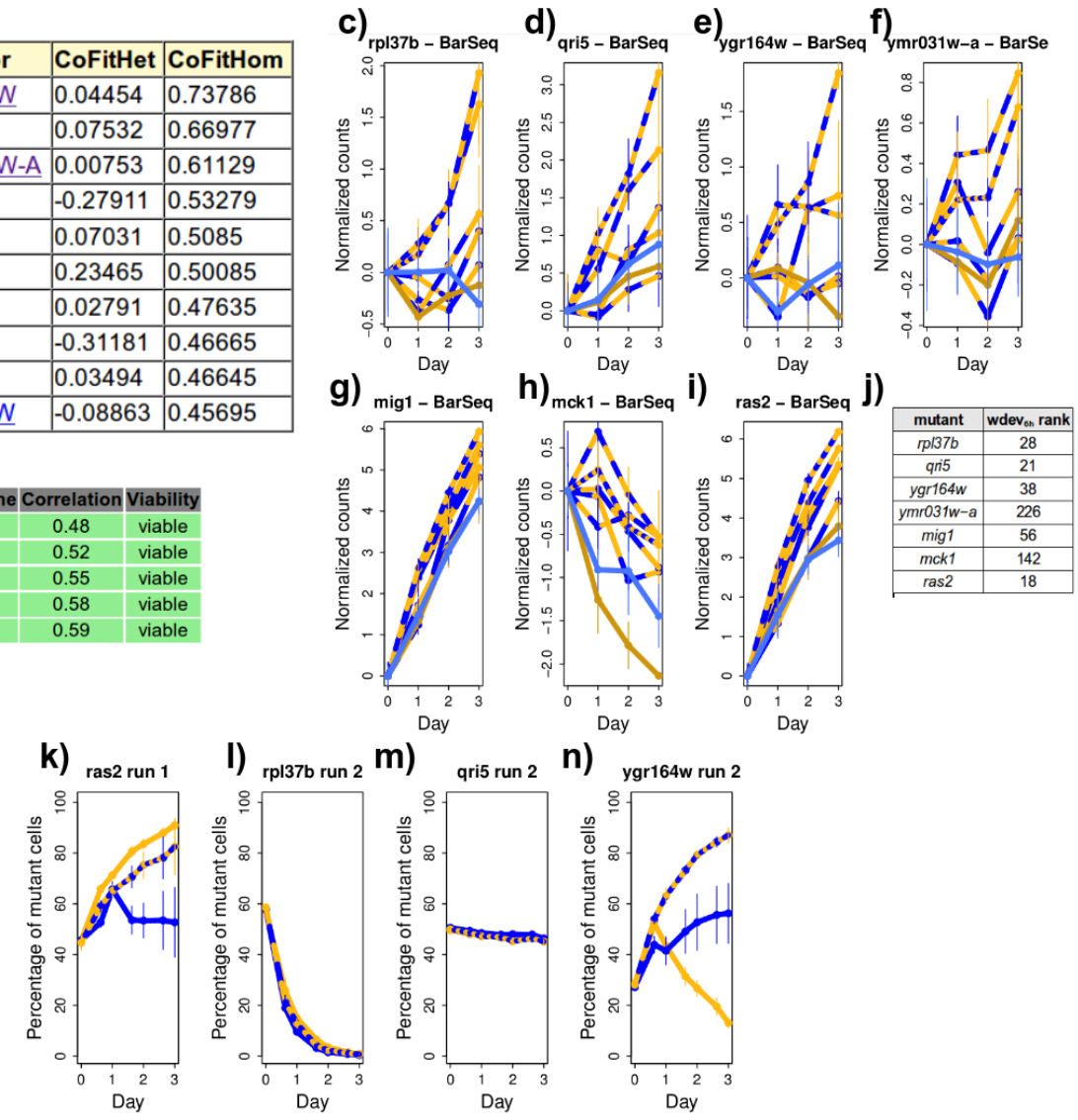
In 2002, Giaever *et al.* used differential interference contrast (DIC) microscopy to determine the shape and size of mutants in the homologous diploid YDL. They classified mutants in seven classes: ‘elongated’, ‘round’, ‘small’, ‘large’, ‘pointed’, ‘clumped’ and ‘other’. Interestingly, in both the *tom7* and the *trm1* group, there is a significant gene ontology enrichment (Robinson et al., 2002) of mutants with a large morphology: yap3, *tom7* and *hal9* in the *tom7* group (pvalue: 0.00069 after Bonferroni correction) and *rpl37b*, *mig1* and *qri5* in the *trm1* group (pvalue: 0.00044 after Bonferroni correction). Thus, there might be a link between this specific cell shape and fitness transgressivity in salt periodic fluctuations.

a)

Query	Interactor	CoFitHet	CoFitHom
QRI5	YGR164W	0.04454	0.73786
QRI5	TRM1	0.07532	0.66977
QRI5	YMR031W-A	0.00753	0.61129
QRI5	MIG1	-0.27911	0.53279
QRI5	MCK1	0.07031	0.5085
QRI5	CSF1	0.23465	0.50085
QRI5	PEX14	0.02791	0.47635
QRI5	RAS2	-0.31181	0.46665
QRI5	PEX1	0.03494	0.46645
QRI5	YCR085W	-0.08863	0.45695

b)

Gene ID	Gene Name	Correlation	Viability
YLR087C	CSF1	0.48	viable
YDR120C	TRM1	0.52	viable
YGR214W	RPS0A	0.55	viable
YDR500C	RPL37B	0.58	viable
YLR423C	ATG17	0.59	viable



j)

mutant	$wdev_{6h}$ rank
rpl37b	28
qri5	21
ygr164w	38
ymr031w–a	226
mig1	56
mck1	142
ras2	18

Fig V-B.1. *trm1* co-fitness group: homozygous mutants most co-fit to the *qri5* mutant on **a)** fitDB and **b)** FMI. Atg17 and rps0a are missing in my data. Rpl37b is missing on fitDB. **c-i)** BarSeq data of co-fit mutants. y-axis: normalized counts adjusted to zero at day zero (y-axis). **j)** ranking of co-fit mutants among all other mutants' wdev values for the 6 hours period. **k-n)** flow-cytometry data of co-fit mutants. y-axis: percentage of mutant cells vs a *wild-type* strain. x-axis: days. Flow cytometry data for the *trm1* mutant in a one-day experiment are shown in Fig IV-C.3a.

C *pde2* group: only one strain highly co-fit

Since the *pde2* mutant was validated and complemented, I tried to infer its function using co-fitness databases. The *pde2* mutant has only one mutant that has a strong correlation across conditions on fitDB: the *ira2* mutant (Fig V-C.1). Unfortunately, it was missing in my data, therefore no inference

could be made. Interestingly, the *tom7* mutant and the *ymr031w-a* mutant are among the top 10 strains most correlated to *pde2*. Thus, *pde2* might mediate its transgressive effect through regulation of genes in the *tom7* or *trm1* group.

Query	Interactor	CoFitHet	CoFitHom
PDE2	IRA2	0.4002	0.70373
PDE2	YJL217W	0.05504	0.53215
PDE2	FBP26	-0.20545	0.52537
PDE2	MDM36	0.25551	0.52162
PDE2	RAS2	-0.03129	0.51258
PDE2	GPB1	0.19955	0.50733
PDE2	FYV10	0.01311	0.49696
PDE2	YMR031W-A	-0.03635	0.4748
PDE2	YJL185C	0.13729	0.47172
PDE2	TOM7	-0.22528	0.46467

Fig V-C.1. *pde2* co-fitness group: Homozygous mutants most co-fit to the *pde2* mutant on fitDB.

D YCR group: mutations in a region of the genome results in a strong transgressive winner phenotype

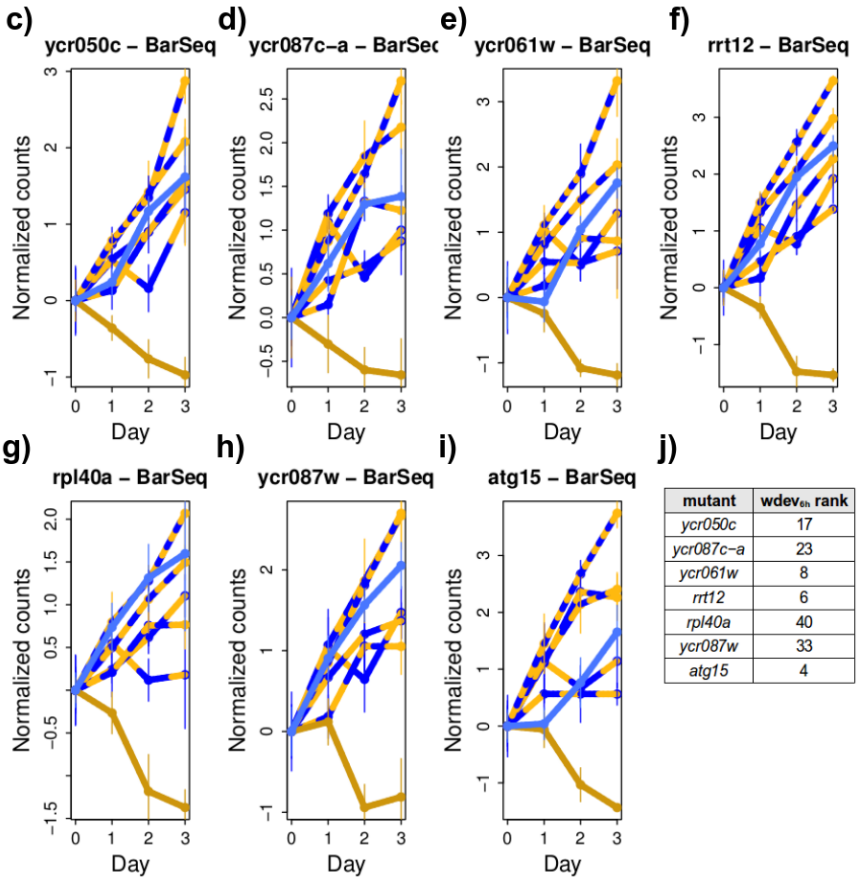
Ycr050cp is a non-essential protein of unknown function. The *ycr050c* mutant had a very strong nonlinear winner effect at the short periods of fluctuations. As above, I tried to infer the molecular mechanism of its nonlinear effect by looking at co-fitness databases. On fitDB there are 10 mutants that have a correlation with *ycr050c* higher than 0.6. In my data, those mutants look very similar; with a strong fitness decrease in S, and strong increase in N (therefore an Antagonistic Pleiotropy phenotype) and a shaded transgressive winner effect, stronger at shorter periods (Fig V-D.1). Those mutants have very high wdev values at the 6 hours period that are all in the top 50 of highest wdev values at the 6 hours period (Fig V-D.1). This transgressive winner phenotype was validated during the flow-cytometry assay for the six strains that I tested (Fig V-D.2). However, I did not do complementation assays for any of those strains. Many of those mutants have no known function for their deleted ORF. Curiously, several of those mutants have their deleted ORF very close on the genome: between ORF YCR026C and YCR087W on chromosome III (Fig V-D.3) (YCR group). Thus, I can hypothesize that mutations in those genes dis-regulate one key gene in this region. Interestingly, in the flow-cytometry assay, all of those mutants increase as fast in S than in N after one day (except for *ycr087w* than increases slower but longer) and then start to quickly decrease in S. Thus, they might be ignoring salt stress response completely. By looking at their localisation on the genome (Fig V-D.3), I found that they are localised in the proximity of the SSK22 gene, the MAPKKK of the Hog pathway. It could be an interesting candidate gene to explain the effect of mutants in this group. Three other mutants are highly co-fit to *ycr050c*, and are localised in the YBR region on chromosome II (Fig V-D.4) (YBR group). They are localised in a hotspot of transporter genes. The reason of their co-fitness with mutants of the YCR region remains to be determined.

a)

Query	Interactor	CoFitHet	CoFitHom
YCR050C	YCL001W-A	0.78044	0.73862
YCR050C	LUG1	0.74713	0.65661
YCR050C	YCR061W	0	0.65231
YCR050C	YCR045C	0.81978	0.64795
YCR050C	CTP1	0.71581	0.63862
YCR050C	MAL31	0.70637	0.62872
YCR050C	RPL40A	-0.15292	0.62522
YCR050C	NPP1	0.78572	0.6242
YCR050C	YCR087W	0.70907	0.61479
YCR050C	ATG15	0.79967	0.60821

b)

Gene ID	Gene Name	Correlation	Viability
YNL196C	SLZ1	0.52	viable
YBR006W	UGA2	0.53	
YBR291C	CTP1	0.53	viable
YCR059C	YIH1	0.53	viable
YCR026C	NPP1	0.54	viable
YIL148W	RPL40A	0.56	viable
YBR286W	APE3	0.56	viable
YCR049C	YCR049C	0.56	viable
YCR045C	RRT12	0.58	viable



j)

mutant	$wdev_{6h}$ rank
ycr050c	17
ycr087c–a	23
ycr061w	8
rrt12	6
rpl40a	40
ycr087w	33
atg15	4

Fig V-D.1. YCR co-fitness group: homozygous mutants most co-fit to the *ycr050c* mutant on **a)** fitDB, and **b)** FMI, **c-i)** BarSeq data of co-fit mutants. y-axis: normalized counts adjusted to zero at day zero (y-axis). **h)** ranking of co-fit mutants among all other mutants' wdev values for the 6 hours period. Lug1 is YCR087C-A. Atg15 is YCR068W. Rrt12 is YCR045C. Mal31 is YBR298C.

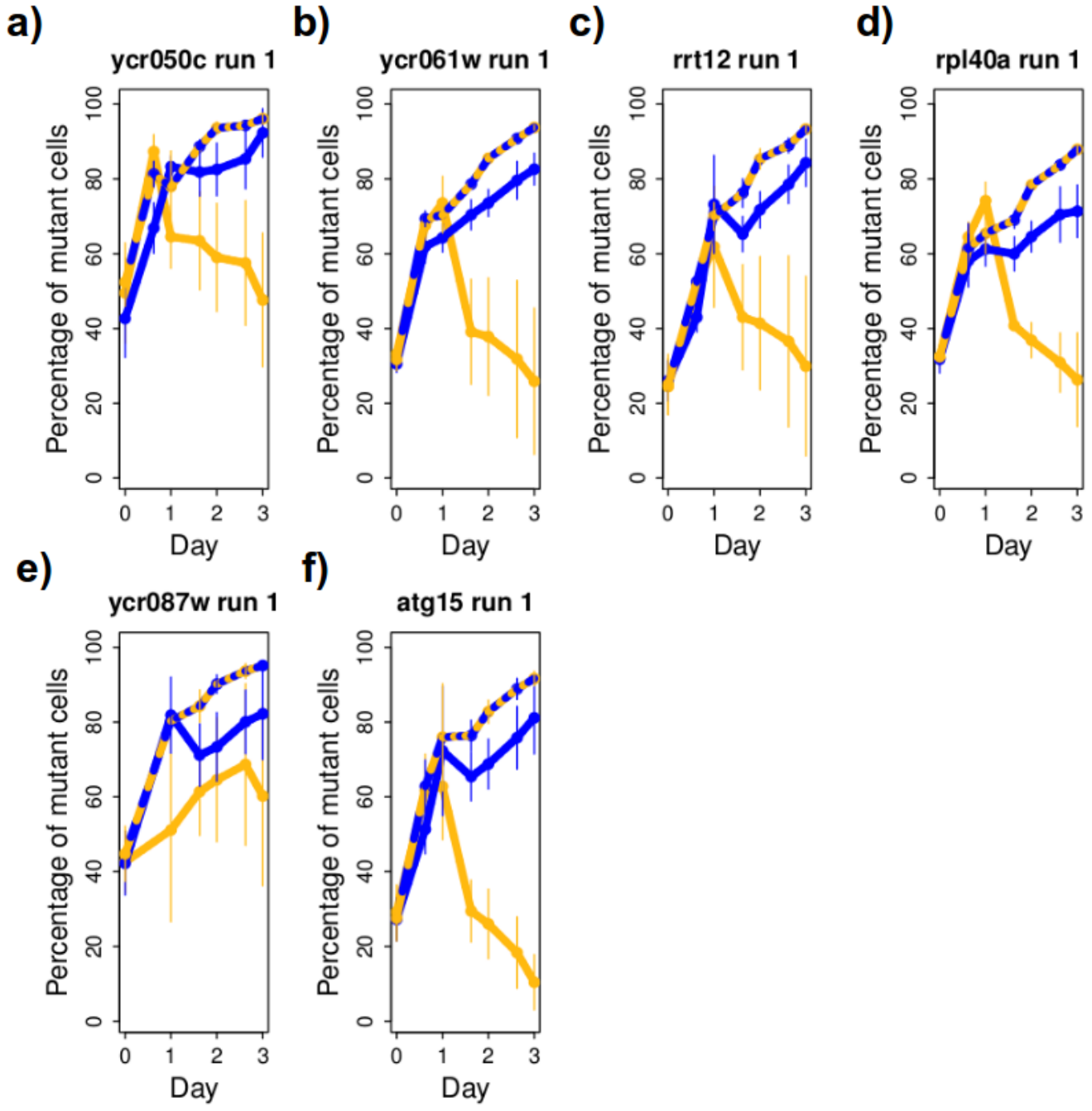


Fig V-D.2. YCR group flow-cytometry data.

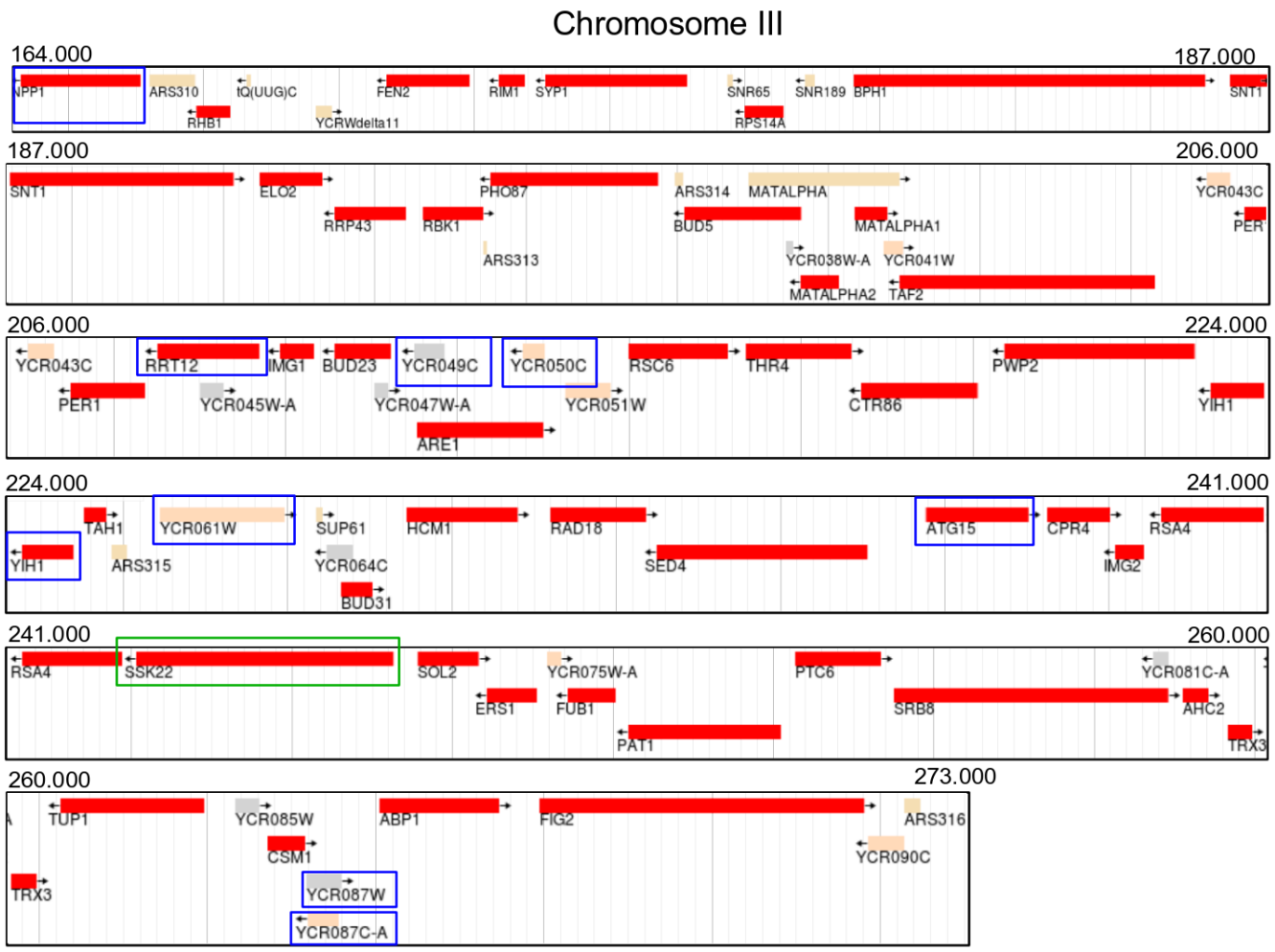


Fig V-D.3. Localisation of deleted genes of mutants in the YCR group. In blue are YCR mutants that are highly co-fit to *ycr050c*. In Green is a putative secondary mutation that might explain their common phenotype: SSK22 is a component of the Sln1 branch of the HOG pathway.

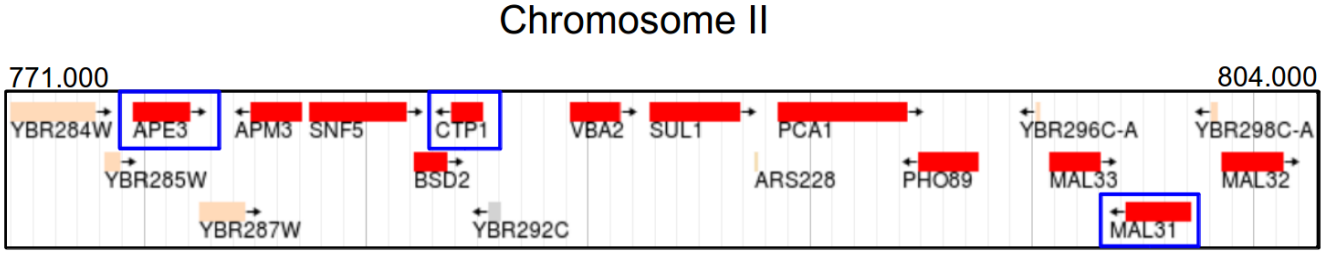


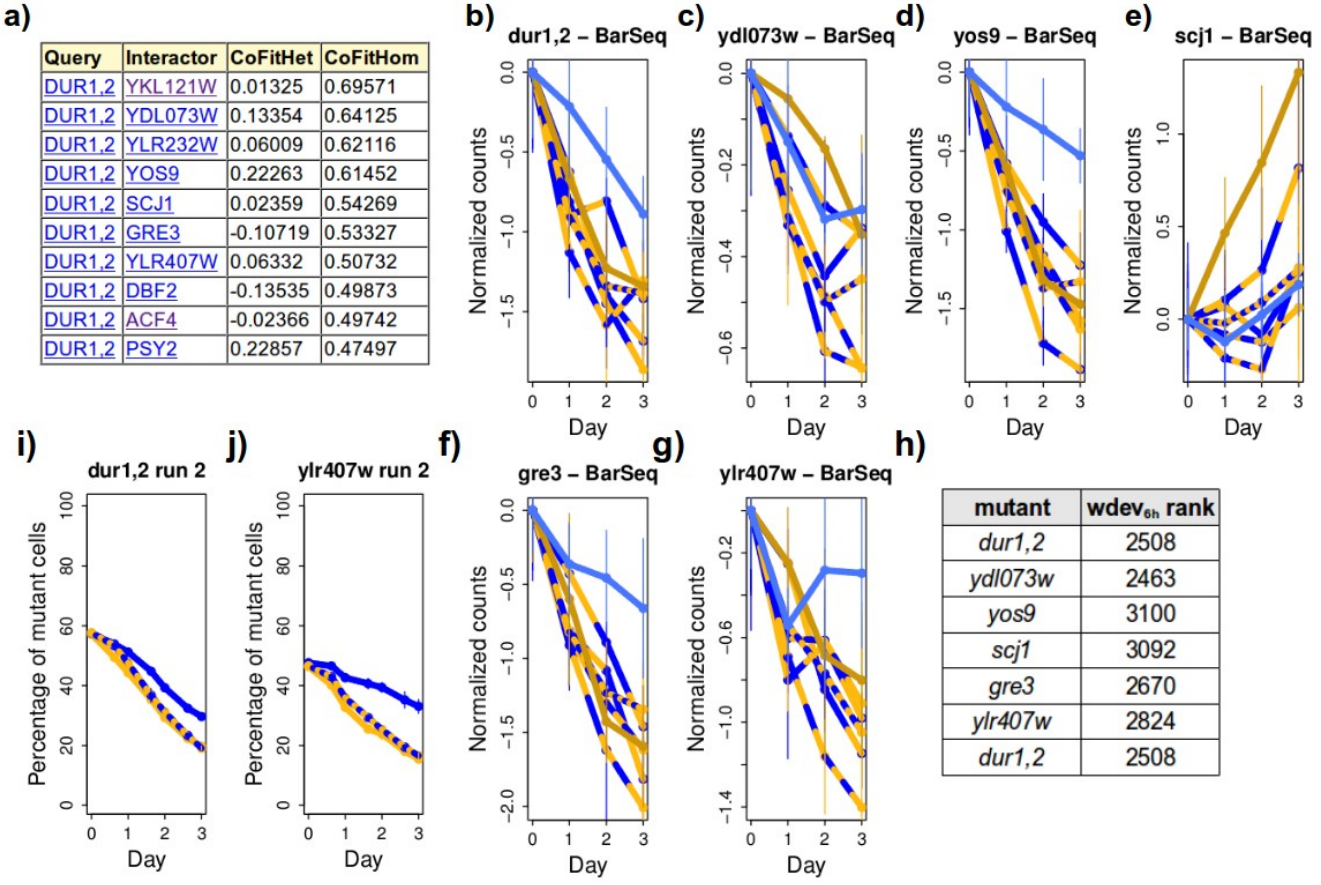
Fig V-D.4. YBR genes localisation: hotspot of transporters/permeases genes. In blue are YBR mutants that are highly co-fit to *ycr050c*.

E Loser group: no obvious cellular functions

I found a group of mutants that have a loser phenotype in periodic fluctuations and low wdev values at the 6 hours period (Fig V-E.1). Those mutants are all highly correlated on fitDB, and 2 of them were validated in the flow-cytometry experiment (Fig V-E.1i,j). Two genes in this group have functions related to protein maturation in the Endoplasmic Reticulum (*yos9* and *scj1*). However, no obvious function could be found for this group of genes.

a)

Query	Interactor	CoFitHet	CoFitHom
DUR1,2	YKL121W	0.01325	0.69571
DUR1,2	YDL073W	0.13354	0.64125
DUR1,2	YLR232W	0.06009	0.62116
DUR1,2	YOS9	0.22263	0.61452
DUR1,2	SCJ1	0.02359	0.54269
DUR1,2	GRE3	-0.10719	0.53327
DUR1,2	YLR407W	0.06332	0.50732
DUR1,2	DBF2	-0.13535	0.49873
DUR1,2	ACF4	-0.02366	0.49742
DUR1,2	PSY2	0.22857	0.47497



h)

mutant	$wdev_{6h}$ rank
dur1,2	2508
ydl073w	2463
yos9	3100
scj1	3092
gre3	2670
ylr407w	2824
dur1,2	2508

Fig V-E.1. Loser co-fitness group: a) homozygous mutants most co-fit to the *dur1,2* mutant on fitDB. **b-e)** BarSeq data of co-fit mutants. y-axis: normalized counts adjusted to zero at day zero (y-axis). x-axis: days. **i-g)** flow-cytometry data of co-fit mutants. y-axis: percentage of mutant cells vs a wild-type strain. x-axis: days. Flow cytometry data for the *trm1* mutant in a one-day experiment are shown in Fig IV-C.3a. **h)** ranking of co-fit mutants among all other mutants' wdev values for the 6 hours period.

VI Deviation between expected and measured fitness: costs and benefits of ignoring environmental stress

When looking at the plot of expected vs measured fitness in periodic salt (Fig VI.1), we can see that many mutants have much higher fitness than expected (as shown by the grey ellipse). Mutants highlighted in blue have been tested in at least one flow-cytometry experiment (Fig VI.1b-i). Those mutants are all transgressive winners with similar phenotypes: a sharp decrease in fitness in salt, a high

fitness without salt and an even higher fitness in the 6 hours periodic oscillations. Strikingly, an interesting phenomenon regularly occurred: fitness in salt was increasing as fast as in the other conditions until 15 or 24 hours. Only after it started to decrease regularly.

I developed the following hypothesis to explain those results (Fig VI.2). Those mutants are defective for functions needed for the response to salt stress. They are ignoring the cues indicating that there is a salt stress, and they continue to divide at a high rate. In contrast, wild-type cells are launching different stress responses (osmotic stress, ionic stress, ESR...), and experience a lag phase in order to adapt to salt, which have a cost on fitness. The concentration of salt used in this experiment is not toxic to cells until about one day of continuous exposure, which explains the non-monotonic growth of those mutants in the S condition. In rapid periodic conditions, the S condition does not last long enough for the stress to be toxic. Alternatively, detoxifying mechanisms may occur in the subsequent N medium. Thus, the growth rates of those mutants remain always high, as compared to the growth rate of the wild-type strain, and they outperform wild-type cells in competition assays.

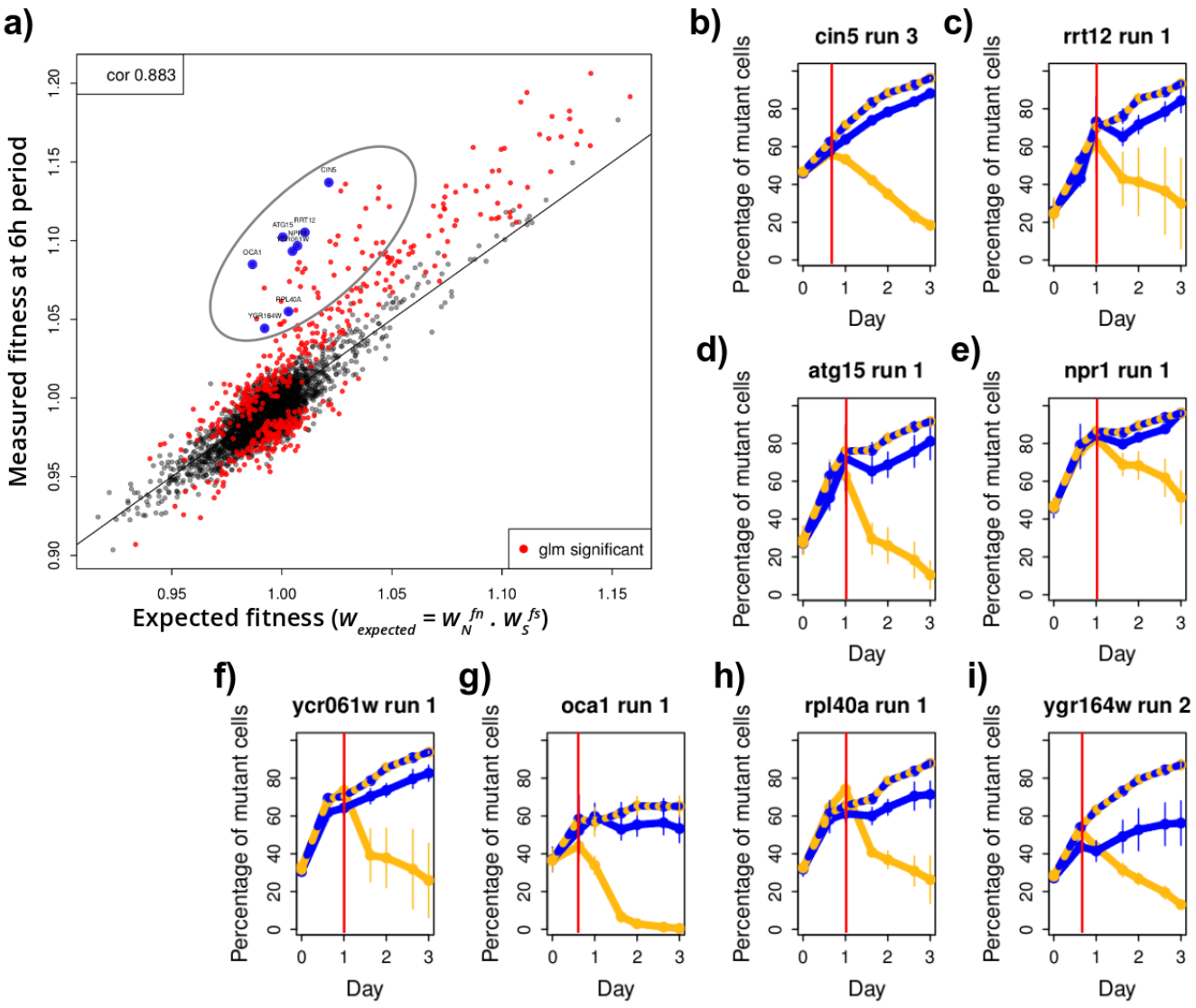


Fig VI.1. Most mutants have no fitness defect when exposed to salt for less than a day or half a day. a) Expected vs measured fitness in 6 hours periodic salt stress (Fig 1-E in the submitted publication). Mutants highlighted in blue are shown in **b-i)** Flow-cytometry data of transgressive winner mutants at the 6 hours period. The red line highlights the time where mutants start to decrease in relative abundance in the S condition.

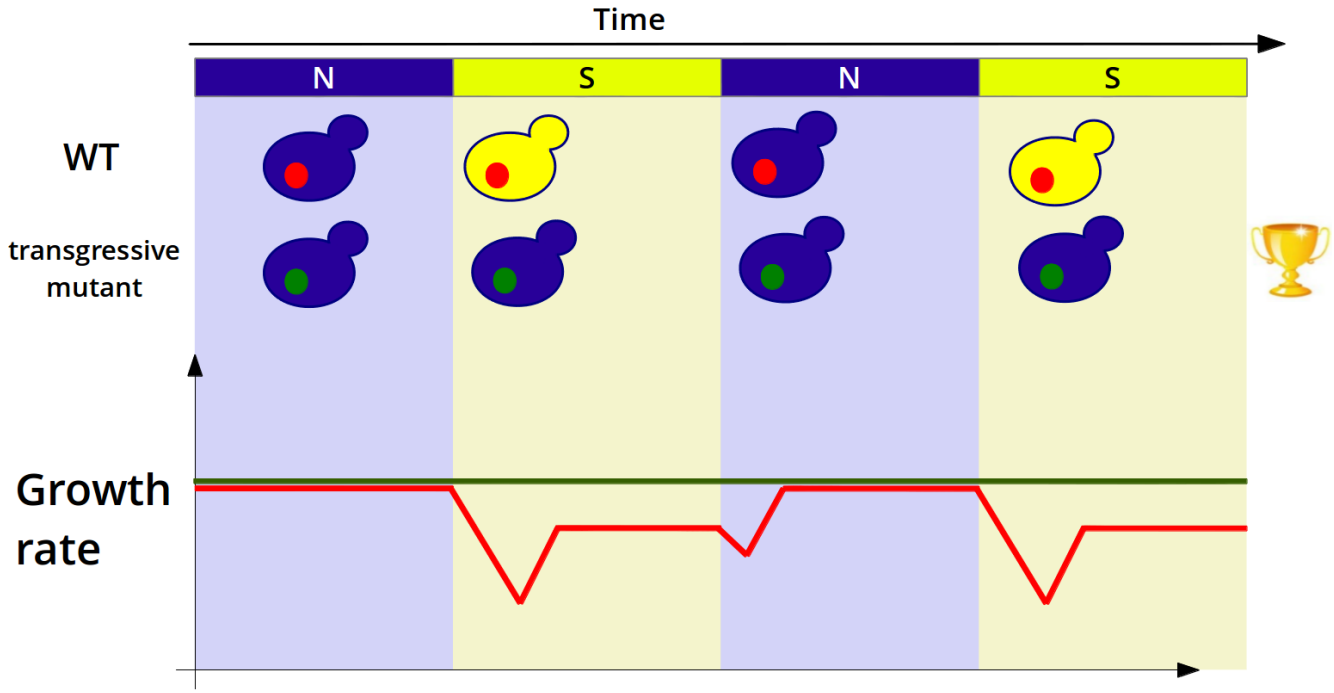


Fig VI.2. Environmental insensitivity can be an adaptive strategy in artificial periodic salt stress.

A theory Schematic illustration of the growth rate over time of two yeast strains (green and red) evolving in a fluctuating environment that oscillates between a stressful (S) and a non-stressful (N) condition. The red strain is hyper-activating stress-response pathways and experiences many events of lag phases and slower growth. On the contrary, the green strain is ignoring environmental stresses, and does not pay the cost of physiological adaptation. The environment changes quickly enough to not be toxic for the cells. If the two strains were in evolving in the same environment, the green strain would win the competition.

4 Discussion

I Environmental insensitivity can be an adaptive strategy in artificial periodic salt stress

I developed a working model to explain my most surprising result in periodic salt stress: mutants with high expected fitness in periodic stress had much higher measured fitness. The model is that the salt stress that we applied was low-enough that its toxicity started to impact the growth of yeast cells only after one day (or half a day) of continuous exposure (Fig VI.2). Wild-type cells activate stress responses as soon as they sense the apparition of the salt stress, in contrast to the nonlinear winner mutants that are defective for some stress response pathways/mechanisms. Activation of a stress response pathway is costly for yeast cells, as it requires energy to rewire the cell from a high-growth rate phenotype to a stress-resistance phenotype. The impact on growth rate is immediate since there is usually a lag phase during the early phase of adaptation to a stress. In addition, cellular energy (i.e. ATP) is invested for up-regulating hundreds of stress response genes.

In my artificial setup of 6h-periodic stress, all this investment is wasted, since the stress will not last long enough to impair the cells' growth rate. Thus, in periodic stress, the wild-type strain and most other mutants in the pool, have a lower growth rate than the mutants ignoring environmental stress. It is possible that some mutants have a slightly superior absolute fitness in periodic stress, as might be the case for the *trm1* mutant (Fig IV-B.7), or the *oca1* and *ygr164w* mutant (Fig IV-B.6). This could be due to an hormesis effect: slight doses of a stress increase an organism's defences and improve its overall fitness/health on the long-term (Mattson, 2008).

The optimal strategy of adaptation depends on environmental dynamics. For instance, once they sense a stress, if they sense it, yeast cells have to either commit to stress response pathways, or not. They might decide to commit more or less rapidly. A fast commitment might be non-adaptive (i.e. is not beneficial for competitive fitness) if the stress disappears quickly (i.e. if they sensed just a transient stimulus). It is always a bet to commit or not, since the cell cannot guess if the stress will remain, increase or disappear. A strain's genome has been shaped by the strain evolutionary history, and will largely determine its choice to commit or not in response to a given stimulus. In the ecology of wild yeasts, once an osmotic stress is detected (i.e. in a rotten fruit), it is common that it will last for some time, and gradually increase. For this reason, the yeast cells will launch a stress response as soon as they sense the stress. Even at a low concentration of 0.2 M NaCl, that impairs growth only after 1 day of exposure. Periodic salt stress is not common in nature, especially at short periods, and thus yeasts are not well-adapted to it. For this reason, it would have been surprising to observe many mutations that increase absolute transgressive fitness. Indeed, that would have implied that evolving such a phenotype is easy. And thus, that the wild-type yeast strain has often encountered such selective pressure during their evolutionary history. In fact, it has been proposed to use artificial periodic fluctuations, to characterize cellular mis-perceptions of the environments, in order to better understand the dynamics of cellular adaptation (Mitchell and Lim, 2016). Finally, dedicated directed evolution experiments could most

likely generate strains that have strong transgressive absolute fitness phenotypes. Indeed, it is just a question of re-wiring the genome to be perfectly adapted to the exact artificial period that the cell experiences.

In 2008, Kao and Sherlock evolved yeast cells at steady state under glucose limited conditions for 448 generations (Kao and Sherlock, 2008). In 2013, whole genome sequencing allowed them to identify 120 mutations over three replicate experiments, most of which were reproducible (Kvitek and Sherlock, 2013). They found that more than half of the mutations were in three key signaling networks that regulate growth control: glucose signaling, Ras/cAMP/PKA (mainly Gpb2, Ira2 and Pde2) and HOG. They conclude that the loss of environmental sensibility is adaptive in a steady environment, but maladaptive if the environment were to change. In contrast, in my experiment, I showed that a loss of environmental sensitivity can be adaptive in a changing environment. The exact specificities of the dynamics and composition of the changing environment should determine if the more adaptive strategy is to adapt to new environments, or to ignore them altogether and grow as fast as possible, or a combination of those two strategies (or other strategies such as bet-hedging).

II Low data quality in other experiments prevent from drawing general principles about fitness in periodic environments

The Yeast Deletion Library (YDL) can be used to screen for genes performing specific functions, or with a specific phenotype. Another interest of using the YDL is that it represents an interesting pool of strains with various phenotypes. Thus, a systems biology approach can be taken in order to determine the general behaviour of living organisms in various environmental conditions. In my project, I used a screening approach in order to find mutants with a transgressive phenotype. This was successful in most experiments, as shown by results of the *vhr1*, *thi2* or *met12* mutants. My aim was also to adopt a systems biology approach to draw general principles about fitness of yeasts in periodic environments. In order to make general conclusions about a phenomenon, a large number of cases must be examined. For this reason, a substantial number of mutants should have important fitness differences between the conditions (N, S, periodic conditions). This number depends mainly of two factors: the strength and the length of the stress applied.

The concentration of the stressful agent used should be high enough to detect many mutants with a difference in growth rate between the different conditions. If the difference in fitness between the two conditions is too small, then BarSeq will not be resolutive enough to measure precisely this difference. If the stress is not strong enough, only mutants for genes that are especially sensitive for the stress will have high-enough fitness differences between conditions to be detected by BarSeq.

Another aspect to consider is the diversity of mechanisms that are used by the cell to detoxify the stress that is applied. This diversity of mechanisms will likely influence the diversity of fitness values between mutants. Indeed, genes that are involved in similar detoxifying mechanisms may be more likely to share similar fitness values in steady and periodic conditions. For instance, salt stress corresponds to both an osmotic and an ionic stress. There are multiple processes within cells to detoxify ionic stress or to regulate osmolarity levels. For instance, the response to osmolar stress involves the regulation of hundreds of genes through transcription factors activated by the MAPK

Hog1p. Thus, this type of stress will affect the fitness of numerous mutants. In contrast, growing a Yeast Deletion Library in absence of methionine will likely affect the growth of a reduced panel of mutants that are mostly involved in methionine biosynthesis.

One way to estimate how many mutants have fitness differences between conditions in each media is to look at the number of Antagonistic Pleiotropic (AP) mutants. I found 48, 3, 16 and 1 AP mutants in the salt, the sodium metabisulfite, the glucose and the methionine experiment respectively (methods of the submitted publication). Thus, it seems that few mutants were affected by the stress applied in the glucose experiment, and even fewer in the methionine and the sodium metabisulfite experiments. This idea is confirmed by the PCA plots (Fig III-B.14-18), where we can see that the S condition is clearly discriminated from the other conditions in the salt experiment, barely in the glucose experiment, and not at all in the methionine and the sodium metabisulfite experiments. Similarly, in the OD plots (Fig III-A.3,4), we can see that at day 3 that the difference in OD between the N and S condition is of 0.5, 0.17, 0.3 and 0 for the salt, sodium metabisulfite, glucose and methionine experiments respectively. Finally, the glucose and methionine experiments suffered from a lower overall data quality (especially replicability) than the salt and the sodium metabisulfite experiments.

To conclude, while I applied the same type of genomic analysis as in my submitted publication to the other media fluctuations, I am not confident that I can truly make generalizations about fitness in fluctuating conditions from those datasets. I chose those stressful medium and concentration in order to have comparable results with another project. However, future experiments trying to address the question of the principles of fitness in periodic stress should precisely chose the stress used, by considering the number of affected cellular functions, and the appropriate concentration necessary to observe important phenotypes without losing all mutants in the S condition. As explained above, losing strains before the end of the experiment is an issue since it prevents from concluding about the linearity of fitness of the mutant (i.e. we don't know how much lower the fitness of the mutant could go down).

However, I can also look at flow-cytometry data in different media fluctuations to determine if my results in salt are condition-specific or not, at least for some targeted mutants. During one experiment (run 4) I tested 6 mutants with a transgressive winner phenotype in salt (*pde2*, *tom7*, *cin5*, *hal9*, *srf1*, *trm1*) in 4 types of media fluctuations (NaCl 0.2 M, NaCl 0.4 M, Sorbitol 0.4M and Raffinose 1.9%) (Fig IV-D.1 and data not shown). I found that, in raffinose fluctuations (carbon stress), only *tom7* was a transgressive winner, *cin5* and *pde2 were* nonlinear losers, and *srf1*, *hal9* and *trm1* were linear (IV-D.1d,h and data not shown). In sorbitol, most of those mutants were linear, some were nonlinear winners, and none were transgressive winners. Thus, it appears that the transgressive winner phenotype of those mutants is mostly specific to mild periodic salt stress. Thus, those preliminary analysis indicate that fitness in periodic stress appears to be largely gene and condition dependent.

III Hypothesis to explain the behavior of selected nonlinear mutants

A Mutants with a frequency-dependent phenotype

A Vhr1

The *vhr1* mutant showed a very special behavior: this mutant had similar fitness in N and S in the salt, glucose and methionine experiment, however, its fitness in periodic conditions was different. Why did I observe this phenomenon?

The VHR1 gene is a transcriptional activator of the high affinity biotin transporter gene VHT1 (Vitamin H Transporter) and of a biotin intermediate precursors importer gene BIO5 (biotin biosynthesis intermediate transporter). Vhr1p activates its target genes upon low biotin concentrations. Biotin (vitamin H) is an essential vitamin that is required for lipid metabolism, leucine metabolism and that acts as a substrate of the biotin protein ligase (BPL1) (Hall and Dietrich, 2007). The s288c strain (the background strain of my mutants) contains 3 core biotin biosynthesis genes (BIO2-4). However, it lacks the initial steps of biotin biosynthesis that are mediated by the BIO1 and BIO6 genes (Hall and Dietrich, 2007). Thus, it is auxotroph for biotin. However, since it contains the other steps of the pathway, it can be complemented by the addition of biotin vitamers, such as the keto 8-aminopelargonic acid (KAPA), which is imported by the product of the BIO5 gene (Phalip et al., 1999).

The expression level of biotin biosynthesis genes and of the biotin transporter, the Vht1p, are generally regulated in an antagonistic way by the environment. For instance, a decrease in extracellular biotin concentrations induces up-regulation of biotin biosynthesis genes (Pirner and Stolz, 2006), while an increase induces up-regulation of the VHT1 gene (Weider et al., 2006). Also, *vhr1* mutants have constitutively high expression levels of biotin biosynthesis genes (Pirner and Stolz, 2006). Another example is the case of iron stress. Biotin biosynthesis proteins contains irons, but not Vht1p. It has been shown that in the case of iron deprivation, biotin biosynthesis genes are down-regulated, whereas VHT1 is up-regulated by AFT1, the major iron-dependent transcription factor (Shakoury-Elizeh et al., 2004). The KAPA importer BIO5 is also up-regulated (Bellí et al., 2004), indicating that cells might store biotin precursors for future iron replete conditions. In contrast, in iron replete conditions, the VHT1 gene is down-regulated and biotin biosynthesis genes are up-regulated (Bayeva et al., 2013; Shakoury-Elizeh et al., 2004).

This shows that yeast cells optimize their growth by balancing production of biotin through import of biotin precursors and biosynthesis of biotin, or import of biotin, depending on their external environment. The SD-all medium, that I used for all my experiments, contains biotin (2 µg/L) but not its precursors KAPA. Thus, in my experiments, production of biotin relied exclusively on import through VHT1. Looking at transcriptomic data of the *vhr1* haploid mutant, I saw that the *vhr1* mutant is strongly down-regulating both the BIO2 and the VHT1 genes (Kemmeren et al., 2014). This transcriptome experiment was performed in standard conditions. Thus, the *vhr1* mutant is probably not able to import biotin, or very little, in standard conditions. This explains why, in my data, the *vhr1* mutant starts being depleted after two days in the N conditions (Fig III-F.6). Interestingly, iron depletion can induce the up-regulation of VHT1 through the iron-specific transcription factor AFT1. Thus, we can imagine that different types of stress can also up-regulate biotin import, through different stress-specific transcription factors. This could explain why we observe almost no growth defect in salt, in sodium metabisulfite or in methionine steady stress. If this hypothesis is correct, the different behavior in fluctuating conditions could be related to the different timing of actions of those transcription factors. Interestingly, the *vhr1* mutant was a loser at the longest periods in the three experiments (Fig III-F.8c,e,f). Thus, there seems to be a limit of survival of ~24 hours without biotin,

which is coherent with the decrease in fitness in the N condition after 2 days. This limit of 24h could be due to the number of generations until which the concentration of biotin of daughter cells is not limiting for their growth. That would be similar to the study of Lambert and Kussel, where cells could grow without any lag phase in periodic lactose fluctuations of up to 8 hours periods, thanks to the dilution of the Lac1p protein that was still above a certain threshold (Lambert and Kussell, 2014).

Importantly, when I tested the *vhr1* mutant during the flow-cytometry experiment, I observed a similar growth effect in all conditions, but with a far weaker intensity (submitted publication Sup Fig 2). I developed an hypothesis to explain this result. In theory, biotin is present at high concentrations in the growth media. Therefore, the *vhr1* mutant should be able to incorporate biotin through non-specific, low-affinity biotin transporters. However, biotin concentration could vary between the two experiments. In the flow-cytometry assay, only two strains were in competition: the *vhr1* mutant and the wild-type strain. Thus, about 50% of cells used high-affinity biotin transporters for biotin uptake. In contrast, during the BarSeq assay, there were more than 4000 mutants competing for resources. Thus, more than 99.99% of cells used high-affinity biotin transporters for biotin uptake. Consequently, there should be twice less biotin available in the primary screen. In addition, some strains might over-express the biotin high-affinity transporter gene vht1, which would result in a faster depletion of biotin in the pool of mutant. Finally, I didn't measure Optical Density during my flow-cytometry experiments. But it is possible that cultures were denser in the BarSeq assay than in the flow-cytometry assay, which would result in even higher biotin depletion in the pool of mutants. Future flow-cytometry experiments on the *vhr1* mutant could be performed by lowering the biotin concentration in the medium, to see if, as expected, it increases the differences in growth rate between conditions.

B Thi2

The *thi2* mutant had a strong frequency-dependent effect in the sodium metabisulfite experiment. THI2 is a transcriptional activator of thiamine regulatory genes (THI genes). Thiamine (vitamin B1) is phosphorylated to produce Thiamine PyroPhosphate (TPP). TPP is an important co-factor for amino acids and carbohydrate metabolisms. It is synthesized *de novo* in plants and microorganisms, but not by animals (Iosue et al., 2016). In *S. cerevisiae*, thiamine can be either imported or synthesized, through a large set of thiamine biosynthesis genes (Iosue et al., 2016). Thiamine regulation is different from biotin regulation, where transporters and biosynthesis genes are antagonistically regulated. Indeed, intracellular TPP is sensed. When it is high, expression of both thiamine high affinity transporter genes and thiamine biosynthesis genes are down-regulated (Iosue et al., 2016; Nosaka et al., 2012). At the molecular level, there are 3 transcriptional activators of thiamine (THI) genes: THI2, THI3 and PDC2. Thi2p and Pdc2p bind together or separately the THI genes promoters (Nosaka et al., 2012). Thi3p then binds the Thi2p-Pdc2p complex, which changes the conformation of the Pdc2p protein and induces the expression of THI genes (Nosaka et al., 2012). Thi3p is bound by TPP, making a negative feedback loop (Nosaka et al., 2012). Indeed, when TPP levels are low, free Thi3p is available to bind to Thi2p-Pdc2p complexes. And when TPP levels are high again expression of THI genes is down-regulated. Consistently, low thiamine concentration induces stronger Thi2p-Pdc2p associations (Nosaka et al., 2012).

Due to its antioxidant properties, sodium metabisulfite (Na2O5S2) is used as a food preservative, and

in the commercial wine making industry. Pediatric formulations contains mixture of vitamins, which usually include thiamine, that are essentials for providing neonates daily requirements or to supply possible deficiencies. Some commercial pediatric formulations use sodium metabisulfite as an antioxidant. Several studies measured the stability of vitamins in those pediatric formulations. It was found that thiamine is strongly degraded by sodium metabisulfite (Ribeiro et al., 2011; Scheiner et al., 1981).

The media that I used for my experiments contained thiamine (400 µg/L), which was thus depleted in the sodium metabisulfite experiment. In yeasts, thiamine deficiencies lead to severe growth defect (Wang et al., 2005). Therefore, in this experiment, yeasts relied on thiamine biosynthesis genes for maintaining a high growth rate. Consistently, in my data, the *thi2* and the *thi3* mutants were depleted quickly in steady sodium metabisulfite stress (data for pdc2 are missing) (Fig III-F.2 and III-F.9).

Looking at raw counts, it seems that the *thi3* mutant was depleted faster since it was already absent at day 1. However, it had 4 times less raw counts than *thi2* at day zero, thus it is unclear which mutants is more sensitive to thiamine deprivation. Interestingly, the *thi3* mutant behaved quite linearly, at least at day 1. In contrast, the *thi2* mutant had a frequency-dependent effect: it had a strong nonlinear winner phenotype at the shortest periods, a linear phenotype at intermediate periods, and a nonlinear loser phenotype at the longest periods. It is intriguing that two directly interacting transcriptional regulators mutants, having similar extreme fitness differences between the two steady conditions, have different fitness in periodic stress. One hypothesis could be that activation of Pdc2p alone by Thi3p, in the *thi2* mutant, could trigger a weak induction of THI genes. This low activation could be sufficient to survive short events of stress (3 hours or 6 hours, for respectively the 6 hours or 12 hours periodic conditions) but not enough when the stress is prolonged. In contrast, the *thi3* mutant would have no activation at all of its THI genes and thus behave linearly. Further experiments will be needed to elucidate the origin of the differences in fitness between the *thi2* and *thi3* mutants in dynamic conditions.

B Met12: a gene that may favor fitness in periodic methionine stress

The *met12* mutant showed a strong transgressive loser phenotype at most periods in the methionine experiment. The MET12 gene codes for a Methylenetetrahydrofolate reductase (MTHFR). This enzyme catalyzes the reduction of 5,10-methylenetetrahydrofolate (CH 2-THF) to 5-methyltetrahydrofolate (CH 3-THF). CH3-THF is then used with homocysteine to produce tetrahydrofolate (THF) and methionine. MTHFR deficiencies lead to several human diseases such as neonatal lethality, hyperhomocysteinemia (Froese et al., 2016), schizophrenia, depression (Gilbody et al., 2007), vascular diseases and neural tube defect (Raymond et al., 1999; Shan et al., 1999). In yeasts, there are two MTHFR genes: MET12 and MET13 (Raymond et al., 1999). However, despite the important therapeutic potential, literature on MTHFR yeast genes remain very scarce since the characterization of the MET12 gene in 1999 (Raymond et al., 1999). In this study, the authors showed that both Met12p and Met13p have MTHFR activity, but Met13p is responsible for most of the activity in the cell. They found that the *met12* mutant had no phenotype. In contrast, single *met13* mutant or double *met12*, *met13* mutants were auxotrophs for methionine. The auxotrophy of the *met12-met13* double mutant was complemented by a plasmid containing the MET13 gene, but not by a multicopy

plasmid containing the MET12 gene (Raymond et al., 1999). Complementation of the auxotrophy of the *met13* mutant was also successful by expression of its human homologue (Shan et al., 1999).

The *met13* mutant is missing in my data. This is potentially due to the stronger MTHFRs activity of the met13 gene, that makes the *met13* mutant inviable. It has been shown in *S. pombe*, that even a small reduction of MTHFR activity causes a methionine requirement (Naula et al., 2002). Consistently, I observed in my data that the *met12* mutant had no growth defect in N, but an important growth defect in S_{Met} (absence of methionine). This growth defect seemed to be exacerbated in fluctuating conditions, as the *met12* mutant had a transgressive loser phenotype at most of the fluctuation regimes tested. Importantly, several mutants for genes involved in homocysteine biosynthesis, through the sulfate accumulation pathway (*met14*, *met16*, *met10*, *met5*, *met8*, *met28*) had, as well, important fitness differences between steady conditions. However, their fitness in fluctuating conditions was linear (excepting met28 at the 6h period that was slightly nonlinear winner). *met31*, which is a transcriptional activator of sulfur metabolic genes, was a nonlinear loser at short periods, but was not transgressive. Thus, impaired production of methionine should not be the correct answer to explain the strong transgressive loser phenotype of *met12* at short periods. One possibility to explain this phenomenon could be that periods of methionine stress would trigger the production of homocysteine, that would accumulate due to to unmatched CH 3-THF levels, which would be toxic for the cells. While in S_{Met}, there could be a feedback mechanism to adjust the production of homocystein to the production of CH 3-THF levels. However, many other explanations are possible.

Finally, before any deeper analysis, careful confirmation experiments (i.e. by flow-cytometry or microfluidics) should be made, in order to confirm the exact phenotype of the *met12* mutant in periodic absence of methionine. Indeed, as mentioned previously, at day 1, we can see that the mutant is decreasing at periods equal or higher to 48 hours, while it is increasing in N (Fig III-F.11). This is inconstant since, at day 1 for long periods, the mutant experienced only the N medium. Nevertheless, the very strong decrease in fitness at short periods is promising.

C Groups of nonlinear mutants: implication of mitochondria?

I wondered why most of the mutants with the strongest transgressive behavior in fluctuating environments were strongly co-fit on databases of co-fitness across thousands of conditions (*tom7* and *trm1* group, Fig V-A.1 and V-B.1). The similar phenotypes of those mutants could not be explained by a close localization on the genome, as for the YCR group (V-D.3 and V-D.4). A possibility is that highly co-fit mutants share the same secondary mutation. In this case, chances are that the shared secondary mutation is, in fact, within the deleted gene of one of the co-fit mutants. This simple theory could be verified by sequencing, or measuring gene expression, of all deleted gene in all of the co-fit mutants of the same group. However, that would be fastidious. And this would not be fruitful if high correlations between mutants reflect true shared biological functions, like for instance the *hog1* group, or the *rim101* group (Fig V.1). I found that some mutants in the *trm1* and *tom7* groups, have deleted genes with very different functions. For instance, Srf1p is a regulator of phospholypase D, Yap3p is a transcription factor involved in resistance to benzene, Hal9p is a transcription factor involved in ions regulations, and Trm1p is a tRNA methyltransferase (Table 4-III-C1/2).

However, after a deeper analysis, I found that mutants in the *tom7* and *trm1* group may share similar phenotypes due to perturbed regulations within mitochondria (Table 4-III-C1/2, Fig 4-III-C3). In particular, I looked at the immediate neighbors on the genome of the mutants' deleted genes. And I found that several mutants were neighbor to genes playing a role within mitochondria. I made a link with mitochondria for four mutants in both the *trm1* and the *tom7* group Several mutants of the *tom7* group had a quite high correlation with mutants of the *trm1* group (and vice-versa). Thus, we can imagine that mutants from those two groups are involved in related functions.

Importantly, I found 3 mutants with a putative role in the electron transport chain: the *qri5* mutant (also called cox24), and the *psy4* and *trm1* mutants, whose deletion could miss-regulate their neighbor genes COR1 and COX26 (Fig 4-III-C3). Cox26p and Cox24p are both localized in the cytochrome c oxidase enzyme (complex IV), while Cor1p is localized in the cytochrome bc1 complex (complex III, or coenzyme Q : cytochrome c – oxidoreductase). In *S. cerevisiae*, the complex IV associates with the complex III into supercomplexes, for improved energy transduction (Levchenko et al., 2016), thus those genes might be directly interacting. Cox24p plays a role in the splicing of the cox1 transcript, however, it has at least one other unidentified function (Barros et al., 2006). The Cox1p assembles to form complex IV. RNA processing of cox1 transcripts is very complex and involves more than 20 genes (Barros et al., 2006). The COX24 gene is localized just upstream the 5' of the mss51 gene, a specific translational activator of cox1 transcripts. Thus, the COX24 and MSS51 gene might be regulated together, in an operon like fashion. Cor1p and Cox26p are subunits or respectively the complex III and the complex IV (Levchenko et al., 2016; Tzagoloff et al., 1986). In the *tom7* group, both Pim1p and Tom7p are important for the biogenesis of mitochondria. Mrpl13p and Mrps35p are both mitochondrial ribosomal proteins. Finally, by looking at their phenotypes on the SGD, I found that almost all mutants from the *tom7* and *trm1* groups had either a decreased rate of respiratory growth, or an absence of respiratory growth (missing data for few of them), which is consistent with a reduced mitochondrial activity (Cherry et al., 2012b).

Interestingly, the *ygr164w* and the *trm1* complemented strains were constructed and tested during the secondary screen (*ygr164w* was complemented by a tRNA present within the YGR164W ORF), but their transgressive phenotypes were still present (Fig IV-C.3). For the complementation assay, the goal was to see if I could abolish the phenotype of certain mutants. This was indeed the case for the *pde2* and *tom7* mutants. However, I cannot directly conclude that, for mutants for which complementation didn't abolish the phenotype, the transgressive effect is not due to the action of the deleted gene. To do so, one should sequence the full sequence inserted at the HO locus, to see if there are no mutations. Additionally, one should measure the expression levels of the complemented genes, for instance through Northern blot assays.

However, the successful complementation of the *tom7* mutant confirms the mitochondrial lead: perturbed mitochondrial regulations may be responsible for the transgressive phenotype of mutants from the *tom7* and *trm1* group. Next experiments to confirm this theory could consist in measuring the expression levels of the neighbor genes with suspected perturbed regulations (MSS51 in *qri5*, MRPS35 in *ygr164w*, COX26 in *trm1*, EIS1 in *ymr031w-a*, COR1 in *psy4*) as compared to wild-type expression levels. Interestingly, both the *cor1* and *eis1* mutant are transgressive losers in my data. We can imagine that deletion of their neighbor genes (respectively psy4 and ymr031w-a) increased their gene expression. This would mean that the gene dosage of those genes results in a transgressive loser or transgressive winner phenotype.

Perturbed TOM/TIM regulation induce perturbed protein import into the mitochondria (Pellegrino and Haynes, 2015). Since mitochondrial autophagy (mitophagy) and the mitochondrial unfolded protein response (UPRmt) are regulated by mitochondrial protein import efficiency (Pellegrino and Haynes, 2015), a possibility is that mutants of the *tom7* group might be involved in a mitochondrial quality control pathway. The *pde2* mutant has a quite high correlation of fitness (0.46) with the *tom7* mutant (Fig V-C.1), thus Pde2p and Tom7p might be involved in similar functions. In addition, it has been shown that the levels of cAMP/PKA signaling can modulate the activity of mitochondria (Leadsham and Gourlay, 2010). One of the most over-expressed genes in a *pde2* haploid mutant is the cit2 mRNA (Kemmeren et al., 2014). This transcript is often used as a reporter of the activity of the retrograde pathway, another type of mitochondrial quality control pathway (Jazwinski, 2013). *pde2* also over-expressed DLD3 and CPA2, two genes that are direct targets of the retrograde pathway (Cherry et al., 2012b; Kemmeren et al., 2014). In conclusion, protein quality control pathways might be involved in the transgressive phenotype that I observed in my data.

The Ytp1p protein is a probable type-III integral membrane protein of unknown function, that has regions of similarity to mitochondrial electron transport chain proteins (West Jr. et al., 1996). The *ytp1* mutant is also a transgressive winner mutant in my data (submitted publication, supplementary file: table dat.summary). On fitDB, it is not correlated to any mutant of the *trm1* or *tom7* group. However, on FMI, its most-highly correlated strains are *trm1* and *rpl37b*. However, the correlation is poor (~0.3 as compared to ~0.6 between *rpl37b* and *trm1*). Thus, the transgressive behavior of the *ytp1* mutant could also be linked to mitochondrial dysfunction and to the electron transport chain, as for the *trm1*, *qri5* and *psy4* mutants. Even though the *ytp1* mutant might have a more distant function.

Trm1 group	
Deleted gene	Informations on SGD
Rpl37b	Ribosomal 60S subunit protein L37B
Qri5 = Cox24	**Mitochondrial** inner membrane protein; required for accumulation of spliced **COX1** mRNA; may have an additional role in translation of COX1 mRNA
In 5' of Mss51	Specific translational activator for the **mitochondrial COX1** mRNA; loosely associated with the matrix face of the mitochondrial inner membrane; localizes to vacuole membrane in response to H2O2; influences both COX1 mRNA translation and Cox1p assembly into **cytochrome c** oxidase
YGR164W	Putative protein of unknown function; conserved across S. cerevisiae strains
In 5' of Mrps35	**Mitochondrial** ribosomal protein of the small subunit
Trm1	tRNA methyltransferase; two forms of protein are made by alternative translation starts; localizes to both nucleus and **mitochondrion** to produce modified base N2,N2-dimethylguanosine in tRNAs in both compartments
In 3' of Cox26	Protein of unknown function; co-purifies with **respiratory chain** supercomplexes composed of Complex III (ubiquinol-**cytochrome c** reductase) and Complex IV (cytochrome c oxidase)
YMR031W-A	Dubious open reading frame; unlikely to encode a functional protein
Overlaps 5' of Eis1	Component of the eisosome required for proper eisosome assembly; similar to Uso1p; authentic, non-tagged protein is detected in a phosphorylated state in highly purified **mitochondria** in high-throughput studies

Table 4-III-C1. Description on the Saccharomyces Genome Database (SGD) of genes in the *trm1* group (Cherry et al., 2012b). Mutants for the MRPS35 and COX26 genes are missing in my data. Mutant *eis1* is a transgressive loser in my data.

Tom7 group	
Deleted gene	Informations on SGD
Mrpl13	**Mitochondrial** ribosomal protein of the large subunit; not essential for mitochondrial translation
Psy4	Regulatory subunit of protein phosphatase PP4
In 3' of Cor1	Core subunit of the ubiquinol-**cytochrome c** reductase complex (bc1 complex)
Yap3	Basic leucine zipper (bZIP) transcription factor involved in ER stress (mutant is sensitive to benzene, arsenic, tunicamycin)
Srf1	Regulator of phospholipase D (Spo14p)
Hal9	salt tolerance through increased expression of the ENA1 (Na+/Li+ extrusion pump)
Pim1	ATP-dependent Lon protease; required for biogenesis and maintenance of **mitochondria**; involved in degradation of misfolded proteins in mitochondria
Tom7	Component of the TOM (translocase of outer membrane) complex; responsible for recognition and initial import steps for all **mitochondrially** directed proteins; promotes assembly and stability of the TOM complex

Table 4-III-C2. Description on the Saccharomyces Genome Database (SGD) of genes in the *tom7* group (Cherry et al., 2012b). Mutant for the COR1 gene is a transgressive loser in my data.

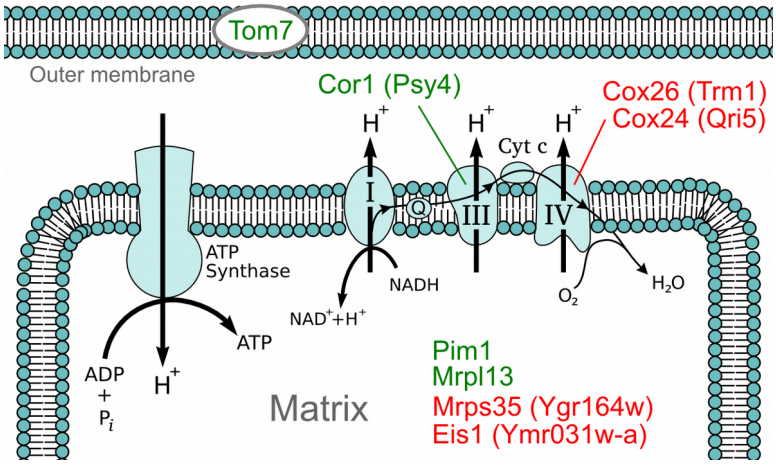


Fig 4-III-C3. Localization within mitochondria of the gene products corresponding to highly co-fit mutants, or their neighbors. In red are mutants from the *trm1* group and in green are mutants from the *tom7* group. This Figure was adapted from: "Electron transport chain" Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 24 May 2017. Web. 27 June 2017., en.wikipedia.org/wiki/Electron_transport_chain

D Interactions between time and environment for selected mutants: the Pde2p protein has a long half-life

My project – analysis of linearity of fitness at different frequencies – allowed to estimate the timescale at which genes are critical for fitness. Indeed, if a mutant is inhomogeneous in some conditions – it has interactions with time and the environment – then it will be nonlinear (see 4-IV-B-iii). For this reason, I thought it could be interesting to look at some genomic datasets of measures related to time. I found two interesting datasets for this purpose: one from a study where yeast cells were exposed to NaCl 0.8 M for up to 36 minutes. RNA synthesis and degradation rates were then measured (Miller et al., 2014). RNA half-lives are usually short, and thus not very informative as compared to the timescales in my experiment. However, RNA synthesis is a good proxy for protein levels. In contrast, to RNAs, protein can have very long half-lives. Therefore, I looked at a study where the half-life of the two third of the yeast proteins was measured (Christiano et al., 2014).

I found that mutants with strong nonlinear winner phenotypes in salt had interesting characteristics in those datasets. The two genes with the most pleiotropic nonlinear phenotype (*vhr1* and *cin5*, submitted publication Fig 3-C/D) also had very short protein half-lives (Cin5p: 0.7 hours, 47th shortest out of 3773 proteins; Vhr1p: 3.3 hours, 293th shortest) (Christiano et al., 2014). Oppositely, the *pde2* protein had one of the longest half-life (Pde2p: 39.4 hours, 90th longest). Unfortunately, data were missing for the *tom7* protein. The *bre2* mutant was the only one with a strongly linear behavior in my flow-cytometry assay. The half-life of the Bre2p protein was intermediate (Bre2p: 7.4 hours, 1212th shortest). 5 AP mutants were completely linear at the 6h period (*pfk26*, *aim26*, *ylr374c*, *bem4* and *stp4*) (submitted publication Fig 3C). Protein half-lives were available for only 2 of them and were intermediate (Pfk26p: 7.3 hours 1181th shortest, Bem4p: 9.5 hours 2299th shortest).

This very basic analysis suggests that there may indeed be a relationship between the linearity of fitness and the half-lives of proteins. However, I found no correlation between wdev and protein half-life at the genomic scale, which could be due to different reasons. Briefly, this includes biases specific to pooled assays, or imprecision of measurements for many mutants (see part 4-IV-B-ii for an extensive discussion). Now will follow some speculations about the link between the linearity effects of the *pde2*, *cin5* and *vhr1* mutants and the kinetic measurements described above.

The long half-life of pde2p is the easiest to interpret. First, the effect of the *pde2* gene on the transgressive phenotype was validated by the complementation assay. The Ras/cAMP/PKA is a major determinant of the trade-off between stress resistance and metabolism and cell cycle production in yeast (Park et al., 2005; Pescini et al., 2012). Pde2p is a high-affinity cyclic AMP phosphodiesterase that plays a dominant role in removing cyclic AMP to repress PKA activity (Park et al., 2005). Thus, *pde2* mutants have high PKA activity, which corresponds to fast proliferation and decreased stress resistance (Park et al., 2005). PDE2 transcripts are highly induced in the presence of salt (Miller et al., 2014). Thus, during the fluctuation experiments, wild-type cells constantly accumulates newly created Pde2p proteins and are thus growing slower than *pde2* mutant cells.

The short half-lives of the Vhr1p and Cin5p proteins are more challenging to explain. Interestingly, both the *cin5* and the *vhr1* mutants have a pleiotropic winner effect at most periods of fluctuations. And both are transcriptional activators. However, the origin of their effect in dynamic environments may be different. Indeed, the VHR1 gene is involved in biotin uptake, and *vhr1* is more fit in S than in N.

While the CIN5 gene is involved in salt stress, and *cin5* is more fit in N than in S. But, first I need to clarify results concerning the *cin5* mutant.

Indeed, the *cin5* complemented strain had the same transgressive behavior as the *cin5* mutant (Fig IV-C.3). Importantly, this does not prove that the CIN5 gene is not involved in the transgressive effect of the *cin5* mutant. Indeed, as explained above (see 4-III-C), there could be mutations in the DNA sequence that was inserted, or more critically the gene might not be expressed, or be expressed at low levels. Several lines of evidence indicate that such an issue may have happened for the *cin5* complemented strain. First, and most importantly, the *cin5* gene is a transcription factor that mediates pleiotropic drug resistance and salt tolerance (Hanlon et al., 2011; Mendizabal et al., 1998). This is consistent with the fact that the *cin5* mutant had a high fitness in N and a low fitness in S, in both the BarSeq and flow-cytometry experiments (submitted publication Sup Fig 2, flow-cytometry run 3). Second, the *cin5* mutant most highly cofit strain on fitDB is Yor029w, a neighbor ORF in 5' of cin5, with a correlation of 0.57. Other mutants have correlations bellow 0.41. This is not what would be expected if a secondary mutation was responsible of the phenotype of the *cin5* mutant. Lastly, the *cin5* and *yor029w* mutants have the two highest wdev values at the 6 hours period. They also make one separate group in the heatmap of AP mutants, the only group for which fitness is strongly non-linear at all periods (submitted publication, Figure 3C). Yor029wp is a putative protein of unknown function. Thus, deletion of the CIN5 gene or of its promoter through *yor029w* deletion is probably responsible of the unique phenotype of the *cin5* mutant in fluctuating conditions.

The transgressive-winner effect of the *cin5* mutant at short periods of fluctuations is a clear example of my theory that nonlinear winner mutants are shutting down their response to stress in order to grow faster (Fig VI.2). However, this does not explain the effect of the *cin5* mutant at long periods, and the relationship with the short half-life of the *cin5* protein. One hypothesis could be that salt stress completely stabilizes the *cin5* protein, and thus increases greatly its half-life. Thus, it would behave like a molecular memory of salt stress (which would be similar to pde2p's long half-life). Another possibility is that Cin5p levels are low at the beginning of each S events. And that stimulation of its target genes is much stronger when newly synthesized Cin5p binds to its promoter, than when Cin5p remains bound for a long duration (as in the S condition). This hyper-activation would result in repeated lag phases, and thus decreased growth rate. This would be similar to a study where periodic salt fluctuations with an 8 minutes period induced a much stronger transcriptional output than continuous salt exposure (Mitchell et al., 2015). In fact, the Cin5p protein might be responsible for the accumulation of the transcriptional output in periodic stimulations observed in this study.

The Vhr1p protein also has a short half-life. As proposed above (section 4-III-A-α), in the N condition, the *vhr1* mutant may be depleted because it is not importing any biotin from the environment. In the S and periodic conditions, the *vhr1* mutant is importing biotin, through stress response transcription factors that activate the biotin transporter gene VHT1. In periodic stress, the *vhr1* mutant grow well as long as the fluctuation period is not too long, with a limit of about 24 hours. To conclude, the dynamics of the other transcription factor(s) that likely induce(s) the over-expression of VHT1 in stressful conditions are probably important to explain the effect of the *vhr1* mutant. However, we cannot exclude that the short half-life of the Vhr1p protein does not play a role as well in the non-linear and pleiotropic behavior of the *vhr1* mutant.

IV Perspectives

A Single-cell approaches to characterize nonlinear mutants

Future single-cell studies could allow determining with a greater precision the dynamical properties of the nonlinear mutants. This would typically involve using a microfluidic device, to change the environment quickly and track single cells over time, coupled with a microscope, to record growth of each trapped cell. Those experiments would allow determining with a high precision the absolute growth rate of nonlinear mutants and wild-type cells. This would allow determining if some mutants have strong absolute transgressive fitness. In addition, the different phases of growth could be decomposed computationally, in order to determine if nonlinear winners have a reduced lag phase when transitioning from N to S or from S to N.

Strategies of adaptations could be studied in more detail. For instance, if we observe two subpopulations with different growth rates within a group of isogenic cells, then it could be indicative of a potential diversifying bet-hedging strategy, where some cells are well-adapted to the N medium, while others are well-adapted to the S medium. However, I doubt that many transgressive winner mutants from the salt experiment employ a type of bet-hedging strategy. Indeed, my theory is that most of them are insensitive mutants that grow as fast as they can and ignore the appearance of a stress (Fig VI.2). It is usually assumed that bet-hedging strategies are favored in unpredictable and severe environmental conditions. In my experiment, the apparition of periodic salt stress, especially at high periods (i.e. 6 hours), is unpredictable for yeast cells, as it is not common in their ecology. However, the stress is quite mild, and thus it might be more adaptive for cells to ignore environmental changes, as proposed above (see 4-I). Still, it could be interesting to study the impact of bet-hedging strategies in fluctuating environments, and some of my mutants might implement this strategy. Using a microfluidic device, various nonlinear mutants could be screened, and classified as either bet-hedging, sensing or fast growth strategists. Then, strains of each class could be grown in either predictable (i.e. steady stress, ramp increase), or unpredictable environments (i.e. non-periodic stress, periodic stress with short periods), to determine the link between growth strategies and environmental conditions.

Effects of priming or molecular memory of stress could also be interesting to study. For instance, one could determine if the growth rate, right after the transition from N to S, is lower at the first cycle than at subsequent cycles. If we observe this phenomenon in the wild-type, then we could screen for mutants that reverse this effect, and thus that mediate molecular memory of salt stress (or another stress if S varies). Also, one could think of tagging proteins of interest (i.e. Pde2p, Cin5p, Vhr1p) in the wild-type strain. This way it would be possible to monitor the amount of those proteins over time in fluctuating or steady conditions. This could give hints about potential effects of memory of stress mediated by long/short protein half-lives.

B Perspectives for the exploration of the impact of environment and time on phenotypes

i Periodic conditions can help to set up models more robust to environmental dynamics

In biology, researchers are summarizing knowledge about interactions between key players/molecules involved in a given process (i.e. salt stress response, mating…) in schematic models. However, links between molecules are quantitative and not qualitative. In order to go further, computational models are needed. Once an accurate model is built, one can rapidly explore the impact of large sets of parameters on the behavior of the model. If an interesting behavior is found, or if the model predicts well a set of parameters but badly another set, then experiments can be performed to improve the model, and thereby to increase knowledge on the studied process. Modeling also allows to evaluate hypothesis that would be tedious to test experimentally. For instance, Apostu and Mackey constructed a model of the GAL network that was capturing its bistable behaviour (Apostu and Mackey, 2012). They could then use their model to discriminate between two hypothesis: the dissociation or non-dissociation of the Gal3p-Gal80p complex at the GAL promoters.

However, this traditional white box modeling approach can be tedious and time-requiring to perform. Indeed, in order to initially fit the model, or to improve it, one needs to have identified the key players involved, and to have measures of their kinetics of interactions. Moreover, different types of measurements could be needed (i.e. interactions between molecules, production and degradation rates, protein conformation…), which would involve mastering a set of experimental skills and having access to necessary equipment. Finally, once the model is built, it could be accurate only for the specific environment used when performing the experiments. Performing all experiments in another environment may be time-consuming.

In contrast, methods of system identification and frequency response analysis do not require knowledge about the internal structure of the system. Instead, they rely on repeated stimulation of the system, in order to characterize its input-output behavior. In this project, I exposed yeast cells to periodic stress, and I measured their fitness as the output. I did not build a model of the response to salt stress, that would include the dynamics of action (or bandwidth) of the genes involved in the response to salt stress (i.e. the HOG pathway, ion detoxifying processes, and downstream targets), since it was not the focus of my study. The data that I generated could be helpful to scientists who would like to improve models of the dynamic HOG response. Though supplementary flow-cytometry assays might be necessary to validate the effect of all genes included in the model, in order to eliminate biases of the pooled assay. In my project, I made a Generalized Linear Model to predict fitness in periodic stress from fitness in steady conditions (expected fitness). However, many mutants had nonlinear fitness: deviations between expected and measured fitness. It could be interesting to try to determine if there are factors that can be incorporated in the model in order to improve its predictive power (see next section).

While I didn't go very far in the modeling aspects for this project, it allowed to have insights about the impact of both the environment and time on the fitness of mutants. The BarSeq experiment revealed that different mutants with similar fitness in steady conditions can have different fitness in dynamic

conditions (submitted publication Fig 1B). This was even true for mutants within the same pathway (submitted publication Sup Fig 1). Thus, this assay allowed to determine the dynamical range at which a gene impacts fitness in the tested condition. Interestingly, I found that the *vhr1* mutant had similar fitness in N and S in three experiments, but various fitness in fluctuating conditions (Fig III-F.7). Thus, periodic stimulation of null mutants can provide informations about the dynamical range at which a gene is important for fitness in response to a stress, which can be condition-dependent. This method has the advantage to be scalable, since different types of periodic environments can be easily tested once the protocol is established.

The phenotype measured during my experiments was the fitness of yeast null mutants. This phenotype was chosen since it is a key evolutionary trait that reflects the selective pressure on the strains. Therefore, it allows to envision future studies on the strategies of adaptation in fluctuating environments, for a set of selected mutants. In addition, it allowed to achieve very high-throughput using the technique of BarSeq. Previous approaches using periodic stimulations to characterize biological dynamics in yeast have relied on fusing GFP to a gene, reporter of the state of activation of the pathway under stimulation (table INT6-ii). These approaches relied on microfluidics devices since they allow to quickly change the growth medium, and to monitor in real time the output of the pathway. However, in the future, periodic stimulations could be used to study other phenotypes than fitness or pathway activation. Those phenotypes should be simple to measure, in order to be able to test many periodic conditions. For instance, using RNAseq, one could determine the impact of environment and time on gene expression. This would allow to better characterize the functions and dynamics of genes.

ii How to improve prediction of fitness in periodic conditions from fitness in steady conditions?

As mentioned above, in the salt experiment, my hypothesis is that many nonlinear mutants with a winner phenotype are shutting down their stress responses (see 4-I). However, this doesn't tell the whole story, as several mutants that shut down their stress responses have linear fitness. The Rim101 pathway is good example where mutants within the same pathway, all have similar defect in constant salt, but levels of linearity vary in periodic conditions (submitted publication Sup Fig 1). Additionally, several Antagonistic Pleiotropic mutants, for which there is the highest precision to estimate linearity due to the high differences in fitness between steady conditions, have linear fitness. Thus, there should be some factors that influence the levels of linearity of fitness.

I did some basic analysis in order to see if I could find such factors. I made correlations between my wdev values (the full dataset or only antagonistic pleiotropic mutants, wdev or absolute wdev values) with other datasets available online. I was particularly interested with datasets that measured time-related variables, such as: RNA synthesis and degradation rate, protein half-life. I also looked at factors that could influence gene expression regulation over long time scales, such as: gene length, number of transcription factors binding sites, presence or not of introns. However, none of them were correlated/enriched. Other factors that could be interesting to look at could be: the number of protein-protein interaction partners, genetic interactions, protein localization, or the mutants' transcriptome.

Importantly, the reason why correlations failed may not be due to the tested factors, but to the nature of

my BarSeq data. Indeed, there are two potential issues. The first issue, is the presence of many biases in those pooled assays, such as: interactions between mutants, or mutants for which the phenotype is not driven by the deleted gene, but by aneuploidies, secondary mutations or mis-regulation in cis. Those mutants would lower any correlation with other types of data that do not share those bias. The second issue, is that a portion of mutants have few fitness differences between conditions. And thus, their wdev values might simply reflect noise in measurements, or the absence of phenotypes. For those reasons, efforts to make correlations or comparisons between BarSeq data and other types of data should be made cautiously.

For this purpose, it would be highly valuable if measures of the quality of each mutant was available. Indeed, it has been shown that some mutants are more likely to acquire secondary mutations, and often acquire the same secondary mutation (Teng et al., 2013). In addition, one study has made extensive quality controls on ~1400 haploid mutants (Kemmeren et al., 2014). Before measuring the transcriptome of those mutants, they wanted to ensure of the quality of the strains. They measured consistency (phenotyping of strains with other strains in the same pathway/complex to detect unexpected phenotypes), aneuploidies and correct gene deletions. They re-constructed the strains for 101 low-quality mutants. It would be very useful to the community working with BarSeq in yeast, if one study were to measure the quality of all strains in the homozygous yeast deletion library (without the hassle of reconstructing them). This could allow to identify strains that have non-desired genetic regulations. Another useful experiment would be to compare the phenotype of each strain, in the pooled assay, and in competition with the wild-type, for few conditions. This way one could identify the strains that are especially advantaged/disadvantaged by interactions with other strains in the pool, such as the *vhr1* mutant.

Direct correlations of BarSeq data to other types of data may fail due to reasons cited above. For this reason, it may be better to trim datasets from BarSeq, in order to keep a subset of high-quality mutants for which correlations to other datasets would be more meaningful. If a comprehensive list of the quality of all mutants was available, only mutants with the highest quality could be kept for comparisons with other datasets. In addition, one should work with mutants that have an important differential in fitness between conditions of interest, in order to have measures above the noise of measurements. In my case, I could select only mutants that are AP between any of the 3 conditions (N, S and NS6). Since an important number of mutants should be kept in order to achieve decent correlations, it could be interesting to lower the threshold that I used to select AP mutants (there are only 48 AP mutants in the salt experiment). Alternatively, I could keep only mutants that have fitness differences that are significantly above the noise between at least two conditions. Other types of analysis than correlations could be performed. For instance, the set of studied mutants could be stratified into 3 different groups: nonlinear winners, linear or nonlinear losers. A number of online datasets could then be analyzed and specific enrichments within one group could be detected.

Finally, the most promising approach to discover nonlinearity factors may be to do Pathway Analysis (PA). PA has proven to be a successful way to give biological meaning to high-throughput data (García-Campos et al., 2015). It consists in determining if certain pathways or functions are enriched in a dataset. PA can be performed with a great diversity of methods, and using a large set of Pathway DataBases (PDB) (García-Campos et al., 2015). A pathway, in those databases, designate a group of genes that are involved in the same function/process. The outcome of pathway analysis are functions or biological processes that are enriched for the measured trait. This is much easier to interpret than a list of unrelated single genes. In addition, experimental evolution studies showed that replicated evolved

lines have few chances to evolve mutations in the same individual genes, but high chances in the same functional units (Lang and Desai, 2014; Tenaillon et al., 2012). Thus, it might be more relevant to explore genomics data at the level of pathways or functional units. Finally, as mentioned above, nonlinearity is condition dependent. This would fit well with various pathways that modulate linearity of fitness in different conditions.

To conclude, once one or few factors that modulate linearity of fitness have been discovered, they could be incorporated into my Generalized Linear Model. Those nonlinear rectifiers would increase the ability of the model to predict fitness in fluctuating environments from fitness in steady conditions. Importantly, some nonlinear rectifiers may be pleiotropic, while others may be specific to one periodic condition.

iii Linearity, homogeneity and additivity of fitness – towards novel assays to detect environmental interactions?

In this study, I used the expression "linearity of fitness" to describe mutants that have a fitness in periodic conditions that corresponds to the weighted average of their fitness in steady conditions. Linearity of fitness can give clues on the rate of environmental change until which the deleted gene plays a role on fitness. However, to be linear a mathematical function must satisfy the conditions of homogeneity and additivity (Fig I-V-C). I didn't prove that my mutants are either homogeneous or additive. However, I considered that if a mutant was linear, there were good chances that both conditions were satisfied. In the future, it could be interesting to develop experiments to test if linear mutants are indeed both additive and homogeneous. In addition, it could be interesting to determine if nonlinear mutants are nonlinear because they are non-homogeneous, non-additive, or both.

Inhomogeneity indicates that the output does not increase monotonously and proportionally with time, thus that there is a time x environment interaction. There are plenty of ways this could happen biologically. For instance, there could be a certain lag time to initialize production and/or degradation of the output. Or there could be genetic circuits that attenuate (negative feedback loop) or increase (positive feedback loop) the output over time.

In fact, we can already have an idea of the homogeneity of some mutants, by looking at the data from my flow-cytometry experiments (submitted publication Sup Fig 2). We can see that 3 linear mutants (*glc8*, *hda2* and *bre3*) seemed to have homogeneous growth (excepting hda2 in N). In contrast, most transgressive winners had non-homogeneous and non-monotonous growth in the S condition (as shown in Fig VI.1). This is a good example of the interaction between time and the environment: the S medium is toxic for those yeast strains only after one day of continuous exposure. Most nonlinear losers (*mot3*, *ylr407w*, *emp46*, *put4*, *cka2*, *nrg1*) and nonlinear winners (*rim21*, *yap3*, *rrt106*) seemed to have homogeneous growth in all conditions. Thus, the nonlinear fitness effect of nonlinear winners and nonlinear losers may be explained by a non-additivity effects.

Additivity means that o(A+B) (the output of input A+B) is equal to o(A) + o(B). In fact, additivity of fitness has been widely studied in biology through genetic interactions assays (section 1-I-A), where the inputs are mutation of gene A and mutation of gene B and the output is fitness. This revealed to be a

highly informative approach to determine if two genes have similar functions, or are part of the same pathway or not. In the case of fluctuating environments, the property of additivity is respected if: o(A then B) = o(B then A) = o(A) + o(B). In this case the inputs A and B are the growth media, and the output is fitness. Let's consider an AP mutant such that: o(A) > 1 > o(B). If o(A then B) > o(A) + o(B) > o(B then A), it could mean that the first environment decides the growth strategy the mutant commits to. That is either fast growth if environment A comes first, or stress responses if environment B comes first. By analogy to the term of "genetic interaction", we could say that an environmental interaction, or an environment x environment interaction has been detected. Another type of environmental interaction is if o(A then B) = o(B then A) < o(A) + o(B). In this case the successions of the two environments is costly for the fitness of the mutant. This could be due to molecular factors that commits in opposite directions in both environments, hence a non-productive result when those environments occurs successively.

Finally, other factors may introduce nonlinearities in fluctuating environments. For instance, mechanisms of molecular memory of stresses can introduce nonlinearities starting from the second cycle of the periodic fluctuations. Those could as well be predictive factors for determining fitness in periodic stress from fitness in steady conditions.

5 Conclusions

To conclude, I found that a loss of environmental sensitivity is an adaptive strategy in periodic stress. This was unexpected, as fluctuating environments are often assumed to favor more plastic strategies such as bet-hedging or generalism. This result could be driven by the mildness of the stress that was applied periodically. In addition, I found that fitness in periodic conditions could often not be predicted from fitness in steady conditions (i.e. nonlinear fitness). For instance, some mutants involved in the same pathways and with similar growth in steady conditions had varying growth in fluctuating conditions. Interestingly, some mutants had very high fitness at short periods, and very low fitness at long periods. This behavior likely reflects a buffering of the stress at short periods, and a breakdown of the cells at longer periods. Finally, many mutants had a higher competitive growth in periodic stress than in steady conditions. Some of the most extreme mutants might mediate this effect by perturbed mitochondrial regulations. To support this hypothesis, a complementation assay in a *tom7* mutant strain succeeded in removing its high competitive growth phenotype in fluctuating environment.

While many theoretical studies proposed strategies of adaptation in fluctuating environments, very few experimental studies were performed. My results showed that a strategy that is usually considered better for growth in steady conditions – a loss of environmental sensitivity – can, in fact, be even more beneficial in certain fluctuating environments. I also found that fast fluctuating environments can increase the genetic variance in fitness, as compared to steady environments. Finally, I found numerous differences in the linearity of fitness in various mutants, which suggests a wild diversity in the speed at which different genes influence fitness in fluctuating conditions.

Future research could be aimed at discovering the factors that underlie nonlinearities in fitness, in a condition-specific or pleiotropic manner. Improving our predictive power to predict fitness in changing environments would translate in a better understanding of the dynamics of actions of genes within the cell, but also of the dynamics of stress responses. The nonlinear genome x environment x time interactions that occurs within cells have rarely been studied, and go a step further to mimicking natural environments. Similarly, novel assays could be developed to characterize the impact on fitness of yeast genes in single environments at different time points (time x environment interaction), or in pairs of two environments (environment x environment interaction).

During my PhD, I had the chance to work with a Yeast Deletion library, that allowed me to easily screen most yeast genes. However, BarSeq assays have some drawbacks, which render the analysis of results less straightforward. In this respect, quality control studies aiming at identifying biases of BarSeq studies using Yeast Deletion Libraries could be highly valuable for the community. Alternatively other types of yeast libraries may be constructed in the future. For instance, a Cre-Lox library where all genes would be mutated upon addition of a chemical, would reduce the number of secondary mutations at the moment where the experiment is performed. It could also be interesting to generate libraries with higher genome saturation, such as mutation of most of the yeast base pairs (one barcode and one mutation per strain), in order to test various levels of expression for each gene.

Finally, now that a set of mutants with various growth behavior in fluctuating environments have been identified, microfluidic approaches could help unravel the mechanisms that underlie these growth

effects. For instance, such studies could use reporter genes to measure the impact of long-lived proteins on fitness in changing environments (i.e. memory effect). One could also determine the phase of growth were a mutant has a growth advantage by measuring the lag phase and exponential phase of growth after each new environment. Finally, mutants with different strategies of growth could be identified and grown in various fluctuating environmental conditions. This would allow to test different theories about fitness in fluctuating conditions.

References

A. Sarhan, and H. Kokko (2007). Multiple Mating In The Glanville Fritillary Butterfly: A Case Of Within-Generation Bet-Hedging? Evolution.

Acar, M., Becskei, A., and van Oudenaarden, A. (2005). Enhancement of cellular memory by reducing stochastic transitions. Nature *435*, 228–232.

Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins. Cell *137*, 146–158.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Andrews, S. (2010). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data.

Ang, J., Ingalls, B., and McMillen, D. (2011). Chapter Ten - Probing the Input–Output Behavior of Biochemical and Genetic Systems: System Identification Methods from Control Theory. In Methods in Enzymology, M.L.J. and L. Brand, ed. (Academic Press), pp. 279–317.

Apostu, R., and Mackey, M.C. (2012). Mathematical model of GAL regulon dynamics in Saccharomyces cerevisiae. J. Theor. Biol. *293*, 219–235.

Aragon, A.D., Quiñones, G.A., Thomas, E.V., Roy, S., and Werner-Washburne, M. (2006). Release of extraction-resistant mRNA in stationary phase Saccharomyces cerevisiae produces a massive increase in transcript abundance in response to stress. Genome Biol. *7*, R9.

Arnoldini, M., Mostowy, R., Bonhoeffer, S., and Ackermann, M. (2012). Evolution of Stress Response in the Face of Unreliable Environmental Signals. PLoS Comput. Biol. *8*, e1002627.

Åström, K.J., and Murray, R.M. (2008). Feedback systems: an introduction for scientists and engineers (Princeton: Princeton University Press).

Attfield, P.V., Choi, H.Y., Veal, D.A., and Bell, P.J.L. (2001). Heterogeneity of stress gene expression and stress resistance among individual cells of Saccharomyces cerevisiae. Mol. Microbiol. *40*, 1000–1008.

Auld, J.R., Agrawal, A.A., and Relyea, R.A. (2010). Re-evaluating the costs and limits of adaptive phenotypic plasticity. Proc. R. Soc. Lond. B Biol. Sci. *277*, 503–511.

Avery, S.V. (2006). Microbial cell individuality and the underlying sources of heterogeneity. Nat. Rev. Microbiol. *4*, 577–587.

Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial Persistence as a Phenotypic Switch. Science 1622–1626.

Barros, M.H., Myers, A.M., Driesche, S.V., and Tzagoloff, A. (2006). COX24 Codes for a Mitochondrial Protein Required for Processing of the COX1 Transcript. J. Biol. Chem. *281*, 3743–3751.

Bar-Ziv, R., Voichek, Y., and Barkai, N. (2016). Chromatin dynamics during DNA replication. Genome Res. *26*, 1245–1256.

Battesti, A., Majdalani, N., and Gottesman, S. (2011). The RpoS-Mediated General Stress Response in *Escherichia coli* *. Annu. Rev. Microbiol. *65*, 189–213.

Baumgartner, B.L., Bennett, M.R., Ferry, M., Johnson, T.L., Tsimring, L.S., and Hasty, J. (2011). Antagonistic gene transcripts regulate adaptation to new growth environments. Proc. Natl. Acad. Sci. *108*, 21087–21092.

Bell, G. (2010). Fluctuating selection: the perpetual renewal of adaptation in variable environments. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *365*, 87–97.

Bennett, M.R., Pang, W.L., Ostroff, N.A., Baumgartner, B.L., Nayak, S., Tsimring, L.S., and Hasty, J. (2008). Metabolic gene regulation in a dynamically changing environment. Nature *454*, 1119–1122.

Berg, M.P., Kiers, E.T., Driessen, G., Van Der HEIJDEN, M., Kooi, B.W., Kuenen, F., Liefting, M., Verhoef, H.A., and Ellers, J. (2010). Adapt or disperse: understanding species persistence in a changing world. Glob. Change Biol. *16*, 587–598.

Berry, D.B., and Gasch, A.P. (2008). Stress-activated Genomic Expression Changes Serve a Preparative Role for Impending Stress in Yeast. Mol. Biol. Cell *4*, 4580–4587.

Berry, D.B., Guan, Q., Hose, J., Haroon, S., Gebbia, M., Heisler, L.E., Nislow, C., Giaever, G., and Gasch, A.P. (2011). Multiple means to the same end: the genetic basis of acquired stress resistance in yeast. PLoS Genet. *7*, e1002353.

Biggar, S.R., and Crabtree, G.R. (2001). Cell signaling can direct either binary or graded transcriptional responses. EMBO J. *20*, 3167–3176.

Bishop, A.L., Rab, F.A., Sumner, E.R., and Avery, S.V. (2007). Phenotypic heterogeneity can enhance rare-cell survival in “stress-sensitive” yeast populations: Phenotypic heterogeneity in yeast populations. Mol. Microbiol. *63*, 507–520.

Botstein, D., Chervitz, S.A., and Cherry, M. (1997). Yeast as a Model Organism. Science *277*, 1259–1260.

Breker, M., Gymrek, M., and Schuldiner, M. (2013). A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. J. Cell Biol. *200*, 839–50.

Brickner, J.H., and Walter, P. (2004). Gene Recruitment of the Activated INO1 Locus to the Nuclear Membrane. PLOS Biol. *2*, e342.

Brickner, D.G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P.-C., Widom, J., and Brickner,

J.H. (2007). H2A.Z-Mediated Localization of Genes at the Nuclear Periphery Confers Epigenetic Memory of Previous Transcriptional State. PLOS Biol. *5*, e81.

Bystrykh, L.V. (2012). Generalized DNA Barcode Design Based on Hamming Codes. PLOS ONE *7*, e36852.

Caudron, F., and Barral, Y. (2013). A Super-Assembly of Whi3 Encodes Memory of Deceptive Encounters by Single Cells during Yeast Courtship. Cell *155*, 1244–1257.

Chen, T., and Dent, S.Y.R. (2014). Chromatin modifiers and remodellers: regulators of cellular differentiation. Nat. Rev. Genet. *15*, 93–106.

Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., et al. (2012a). Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. *40*, D700–D705.

Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., et al. (2012b). Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. *40*, D700–D705.

Chevin, L.-M., and Lande, R. (2015). Evolution of environmental cues for phenotypic plasticity. Evolution *69*, 2767–2775.

Chevin, L.-M., Lande, R., and Mace, G.M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. PLoS Biol. *8*.

Childs, D.Z., Metcalf, C.J.E., and Rees, M. (2010). Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. Proc. Biol. Sci. *277*, 3055–64.

Chou, H.-H., Chiu, H.-C., Delaney, N.F., Segrè, D., and Marx, C.J. (2011). Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. Science *332*, 1190–1192.

Christiano, R., Nagaraj, N., Fröhlich, F., and Walther, T.C. (2014). Global Proteome Turnover Analyses of the Yeasts S. cerevisiae and S. pombe. Cell Rep. *9*, 1959–1965.

Cohen, D. (1966). Optimizing reproduction in a randomly varying environment. J. Theor. Biol. *12*, 119–129.

Cooper S., W., and Kaplan H., R. (1982). Adaptive "Coin-flipping": a Decision-theoretic Examination of Natural Selection for Random Individual Variation. J Theor Biol.

Danforth, B.N. (1999). Emergence dynamics and bet hedging in a desert bee , Perdita portalis. Proc R Soc Lond B.

Dannenberg, J.-H., and Berns, A. (2010). Drugging Drug Resistance. Cell *141*, 18–20.

DeWitt, T.J., Sih, A., and Wilson, D.S. (1998). Costs and limits of phenotypic plasticity. Trends Ecol. Evol. *13*, 77–81.

Dey, S., Proulx, S.R., and Teotónio, H. (2016). Adaptation to Temporally Fluctuating Environments by the Evolution of Maternal Effects. PLOS Biol *14*, e1002388.

Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, G., Castel, D., Estelle, J., et al. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Brief. Bioinform. *14*, 671–683.

Donaldson-matasci, M.C., Lachmann, M., and Bergstrom, C.T. (2008). Phenotypic diversity as an adaptation to environmental uncertainty. Evol. Ecol. Res. 493–515.

Dubnau, D., and Losick, R. (2006). Bistability in bacteria. Mol. Microbiol. *61*, 564–572.

D'Urso, A., and Brickner, J.H. (2014). Mechanisms of epigenetic memory. Trends Genet. *30*, 230–236.

D'Urso, A., Takahashi, Y., Xiong, B., Marone, J., Coukos, R., Randise-Hinchliff, C., Wang, J.-P., Shilatifard, A., and Brickner, J.H. (2016). Set1/COMPASS and Mediator are repurposed to promote epigenetic transcriptional memory. eLife *5*, e16691.

Duveau, F., Metzger, B.P.H., Gruber, J.D., Mack, K., Sood, N., Brooks, T.E., and Wittkopp, P.J. (2014). Mapping Small Effect Mutations in Saccharomyces cerevisiae: Impacts of Experimental Design and Mutational Properties. G3 GenestextbarGenomestextbarGenetics *4*, 1205–1216.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic Gene Expression in a Single Cell. Science *297*, 1183–1186.

Ezard, T.H.G., Prizak, R., and Hoyle, R.B. (2014). The fitness costs of adaptation via phenotypic plasticity and maternal effects. Funct. Ecol. *28*, 693–701.

Fehrmann, S., Bottin-Duplus, H., Leonidou, A., Mollereau, E., Barthelaix, A., Wei, W., Steinmetz, L.M., and Yvert, G. (2013). Natural sequence variants of yeast environmental sensors confer cell-to-cell expression variability. Mol. Syst. Biol. *9*, 695.

Ferrell Jr., J.E. (2012). Bistability, Bifurcations, and Waddington's Epigenetic Landscape. Curr. Biol. *22*, R458–R466.

Fisher, R.A. (1930). The Genetical Theory Of Natural Selection (At The Clarendon Press).

Frankenhuis, W.E., Panchanathan, K., and Belsky, J. (2016). A mathematical model of the evolution of individual differences in developmental plasticity arising through parental bet-hedging. Dev. Sci. *19*, 251–274.

French, B.W., Coates, B.S., and Sappington, T.W. (2014). Inheritance of an extended diapause trait in the Northern corn rootworm, Diabrotica barberi (Coleoptera: Chrysomelidae). J. Appl. Entomol. *138*, 213–221.

Froese, D.S., Huemer, M., Suormala, T., Burda, P., Coelho, D., Guéant, J.-L., Landolt, M.A., Kožich, V., Fowler, B., and Baumgartner, M.R. (2016). Mutation Update and Review of Severe Methylenetetrahydrofolate Reductase Deficiency. Hum. Mutat. *37*, 427–438.

Gallie, J., Libby, E., Bertels, F., Remigi, P., Jendresen, C.B., Ferguson, G.C., Desprat, N., Buffing, M.F., Sauer, U., Beaumont, H.J.E., et al. (2015). Bistability in a Metabolic Network Underpins the De Novo Evolution of Colony Switching in Pseudomonas fluorescens. PLOS Biol. *13*, e1002109.

Garmendia-Torres, C., Skupin, A., Michael, S.A., Ruusuvuori, P., Kuwada, N.J., Falconnet, D., Cary, G.A., Hansen, C., Wiggins, P.A., and Dudley, A.M. (2014). Unidirectional P-Body Transport during the Yeast Cell Cycle. PLOS ONE *9*, e99428.

Gasch, A.P., and Werner-Washburne, M. (2002). The genomics of yeast responses to environmental stress and starvation. Funct. Integr. Genomics *2*, 181–192.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes □. Mol. Biol. Cell 4241–4257.

Geisel, N., Vilar, J.M.G., and Rubi, J.M. (2011). Optimal resting-growth strategies of microbial populations in fluctuating environments. PloS One *6*, e18622.

Giaever, G., and Nislow, C. (2014). The Yeast Deletion Collection: A Decade of Functional Genomics. Genetics *197*, 451–465.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., et al. (2002a). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387–391.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., et al. (2002b). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387–391.

Gibney, P. a, Lu, C., Caudy, A. a, Hess, D.C., and Botstein, D. (2013). Yeast metabolic and signaling genes are required for heat-shock survival and have little overlap with the heat-induced genes. Proc. Natl. Acad. Sci. U. S. A. *110*, E4393–402.

Gilbody, S., Lewis, S., and Lightfoot, T. (2007). Methylenetetrahydrofolate Reductase (MTHFR) Genetic Polymorphisms and Psychiatric Disorders: A HuGE Review. Am. J. Epidemiol. *165*, 1–13.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., et al. (1996). Life with 6000 Genes. Science *274*, 546–567.

Graham, J.K., Smith, M.L., and Simons, A.M. (2014). Experimental evolution of bet hedging under manipulated environmental uncertainty in Neurospora crassa. Proc. R. Soc. B Biol. Sci. *281*, 20140706–20140706.

Grilly, C., Stricker, J., Pang, W.L., Bennett, M.R., and Hasty, J. (2007). A synthetic gene network for tuning protein degradation in Saccharomyces cerevisiae. Mol. Syst. Biol. *3*, 127.

Guan, Q., Haroon, S., Bravo, D.G., Will, J.L., and Gasch, A.P. (2012). Cellular Memory of Acquired Stress Resistance in Saccharomyces cerevisiae. Genetics *192*, 495–505.

Halfmann, R., Alberti, S., and Lindquist, S. (2010). Prions, protein homeostasis, and phenotypic diversity. Trends Cell Biol. *20*, 125–133.

Hallet, B. (2001). Playing Dr Jekyll and Mr Hyde : combined mechanisms of phase variation in bacteria. Curr. Opin. Microbiol. 570–581.

Hallsson, L.R., and Björklund, M. (2012). Selection in a fluctuating environment leads to decreased genetic variation and facilitates the evolution of phenotypic plasticity. J. Evol. Biol. *25*, 1275–1290.

Hamming, R.W. (1950). Error Detecting and Error Correcting Codes. Bell Syst. Tech. J. *29*, 147–160.

Hernday, A.D., Lohse, M.B., Fordyce, P.M., Nobile, C.J., DeRisi, J.L., and Johnson, A.D. (2013). Structure of the transcriptional network controlling white-opaque switching in Candida albicans. Mol. Microbiol. *90*, 22–35.

Hersen, P., McClean, M.N., Mahadevan, L., and Ramanathan, S. (2008). Signal processing by the HOG MAP kinase pathway. Proc. Natl. Acad. Sci. *105*, 7165–7170.

Hilker, M., Schwachtje, J., Baier, M., Balazadeh, S., Bäurle, I., Geiselhardt, S., Hincha, D.K., Kunze, R., Mueller-Roeber, B., Rillig, M.C., et al. (2016). Priming and memory of stress responses in organisms lacking a nervous system: Priming and memory of stress responses. Biol. Rev. *91*, 1118–1133.

Hillenmeyer, M.E., Fung, E., Wildenhain, J., Pierce, S.E., Hoon, S., Lee, W., Proctor, M., St.Onge, R.P., Tyers, M., Koller, D., et al. (2008). The Chemical Genomic Portrait of Yeast: Uncovering a Phenotype for All Genes. Science *320*, 362–365.

Hoepfner, D., Helliwell, S.B., Sadlish, H., Schuierer, S., Filipuzzi, I., Brachat, S., Bhullar, B., Plikat, U., Abraham, Y., Altorfer, M., et al. (2014). High-resolution chemical dissection of a model eukaryote reveals targets, pathways and gene functions. Microbiol. Res. *169*, 107–120.

Holland, S.L., Reader, T., Dyer, P.S., and Avery, S.V. (2014). Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. Environ. Microbiol. *16*, 1729–1740.

Hopper, K.R., States, U., Serice, A.R., Insect, B., and California, U. (2003). Within-generation bet hedging : a seductive explanation ? Oikos 219–222.

Howlett, N.G., and Avery, S.V. (1999). Flow cytometric investigation of heterogeneous copper-sensitivity in asynchronously grown Saccharomyces cerevisiae. FEMS Microbiol. Lett. *176*, 379–386.

Iosue, C.L., Attanasio, N., Shaik, N.F., Neal, E.M., Leone, S.G., Cali, B.J., Peel, M.T., Grannas, A.M., and Wykoff, D.D. (2016). Partial Decay of Thiamine Signal Transduction Pathway Alters Growth Properties of Candida glabrata. PLOS ONE *11*, e0152042.

J. Ripa, H. Olofsson, and N. Jonzen (2009). What is bet-hedging, really? Proc R Soc B 1149–1151.

Jazwinski, S.M. (2013). The retrograde response: When mitochondrial quality control is not enough. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1833*, 400–409.

K. Thumm, and M. Mahony (2002). Hatching dynamics and bet-hedging in a temperate frog , Pseudophryne australis (Anura : Myobatrachidae). Amphib.-Reptil. 433–445.

Kachroo, A.H., Laurent, J.M., Yellman, C.M., Meyer, A.G., Wilke, C.O., and Marcotte, E.M. (2015). Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science *348*, 921–925.

Keleman, K., Krüttner, S., Alenius, M., and Dickson, B.J. (2007). Function of the Drosophila CPEB protein Orb2 in long-term courtship memory. Nat. Neurosci. *10*, 1587–1593.

Kemmeren, P., Sameith, K., van de Pasch, L.A.L., Benschop, J.J., Lenstra, T.L., Margaritis, T., O'Duibhir, E., Apweiler, E., van Wageningen, S., Ko, C.W., et al. (2014). Large-Scale Genetic Perturbations Reveal Regulatory Networks and an Abundance of Gene-Specific Repressors. Cell *157*, 740–752.

Ketola, T., Mikonranta, L., Zhang, J., Saarinen, K., Ormälä, A.-M., Friman, V.-P., Mappes, J., and Laakso, J. (2013). Fluctuating temperature leads to evolution of thermal generalism and preadaptation to novel environments. Evol. Int. J. Org. Evol. *67*, 2936–44.

Kobayashi, T.J., and Sughiyama, Y. (2015). Fluctuation Relations of Fitness and Information in Population Dynamics. Phys. Rev. Lett. *115*, 238102.

Kokko, H., and López-Sepulcre, A. (2006). From Individual Dispersal to Species Ranges: Perspectives for a Changing World. Science *313*, 789–791.

Krueger, F., Andrews, S.R., and Osborne, C.S. (2011). Large Scale Loss of Data in Low-Diversity Illumina Sequencing Libraries Can Be Recovered by Deferred Cluster Calling. PLOS ONE *6*, e16607.

Kundu, S., and Peterson, C.L. (2010). Dominant Role for Signal Transduction in the Transcriptional Memory of Yeast GAL Genes. Mol. Cell. Biol. *30*, 2330–2340.

Kundu, S., Horn, P.J., and Peterson, C.L. (2007). SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. Genes Dev. *21*, 997–1004.

Kussell, E., and Leibler, S. (2005). Phenotypic Diversity , Population Growth , and Information in Fluctuating Environments. Science 2075–2079.

Kussell, E., Kishony, R., Balaban, N.Q., and Leibler, S. (2005). Bacterial persistence: a model of survival in changing environments. Genetics *169*, 1807–14.

Kvitek, D.J., and Sherlock, G. (2013). Whole Genome, Whole Population Sequencing Reveals That Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant Environment. PLoS Genet. *9*, e1003972.

Lambert, G., and Kussell, E. (2014). Memory and Fitness Optimization of Bacteria under Fluctuating Environments. PLOS Genet. *10*, e1004556.

Lande, R. (2009). Adaptation to an extraordinary environment by evolution of phenotypic plasticity and

genetic assimilation. J. Evol. Biol. *22*, 1435–46.

Lang, G.I., and Desai, M.M. (2014). The spectrum of adaptive mutations in experimental evolution. Genomics *104*, 412–416.

Leadsham, J.E., and Gourlay, C.W. (2010). cAMP/PKA signaling balances respiratory activity with mitochondria dependent apoptosis via transcriptional regulation. BMC Cell Biol. *11*, 92.

Lee, A.Y., St.Onge, R.P., Proctor, M.J., Wallace, I.M., Nile, A.H., Spagnuolo, P.A., Jitkova, Y., Gronda, M., Wu, Y., Kim, M.K., et al. (2014). Mapping the Cellular Response to Small Molecules Using Chemogenomic Fitness Signatures. Science *344*, 208–211.

Levchenko, M., Wuttke, J.-M., Römpler, K., Schmidt, B., Neifer, K., Juris, L., Wissel, M., Rehling, P., and Deckers, M. (2016). Cox26 is a novel stoichiometric subunit of the yeast cytochrome c oxidase. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1863*, 1624–1632.

Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions and reversals. Sov. Phys. Dokl. *10*, 707–710.

Levy, S.F., Ziv, N., and Siegal, M.L. (2012). Bet Hedging in Yeast by Heterogeneous, Age-Correlated Expression of a Stress Protectant. PLoS Biol. *10*, e1001325.

Levy, S.F., Blundell, J.R., Venkataram, S., Petrov, D.A., Fisher, D.S., and Sherlock, G. (2015). Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature *519*, 181–186.

Light, W.H., Freaney, J., Sood, V., Thompson, A., D'Urso, A., Horvath, C.M., and Brickner, J.H. (2013). A Conserved Role for Human Nup98 in Altering Chromatin Structure and Promoting Epigenetic Transcriptional Memory. PLOS Biol. *11*, e1001524.

Lisman, J.E. (1985). A mechanism for memory storage insensitive to molecular turnover: a bistable autophosphorylating kinase. Proc. Natl. Acad. Sci. *82*, 3055–3057.

Lof, M.E., Reed, T.E., McNamara, J.M., and Visser, M.E. (2012). Timing in a fluctuating environment: environmental variability and asymmetric fitness curves can lead to adaptively mismatched avian reproduction. Proc. Biol. Sci. *279*, 3161–9.

López-Maury, L., Marguerat, S., and Bähler, J. (2008). Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat. Rev. Genet. *9*, 583–93.

Love, M.I., Huber, W., and Anders, S. (2014a). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Love, M.I., Huber, W., and Anders, S. (2014b). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Macia, J., Regot, S., Peeters, T., Conde, N., Solé, R., and Posas, F. (2009). Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal transduction. Sci. Signal. *2*, ra13.

Martin, G., Elena, S.F., and Lenormand, T. (2007). Distributions of epistasis in microbes fit predictions

from a fitness landscape model. Nat. Genet. *39*, 555–560.

Miller, C., Schwalb, B., Maier, K., Schulz, D., Dumcke, S., Zacher, B., Mayer, A., Sydow, J., Marcinowski, L., Dolken, L., et al. (2014). Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. Mol. Syst. Biol. *7*, 458–458.

Mitchell, A., Romano, G.H., Groisman, B., Yona, A., Dekel, E., Kupiec, M., Dahan, O., and Pilpel, Y. (2009). Adaptive prediction of environmental changes by microorganisms. Nature *460*, 220–224.

Monod, J., and Jacob, F. (1961). General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation. Cold Spring Harb. Symp. Quant. Biol. *26*, 389–401.

Mugler, A., Kittisopikul, M., Hayden, L., Liu, J., Wiggins, C.H., Süel, G.M., and Walczak, A.M. (2016). Noise Expands the Response Range of the Bacillus subtilis Competence Circuit. PLOS Comput. Biol. *12*, e1004793.

Murren, C.J., Auld, J.R., Callahan, H., Ghalambor, C.K., Handelsman, C.A., Heskel, M.A., Kingsolver, J.G., Maclean, H.J., Masel, J., Maughan, H., et al. (2015). Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. Heredity *115*, 293–301.

Naula, N., Walther, C., Baumann, D., and Ernst Schweingruber, M. (2002). Two non-complementing genes encoding enzymatically active methylenetetrahydrofolate reductases control methionine requirement in fission yeast Schizosaccharomyces pombe. Yeast *19*, 841–848.

Neupert, W., and Herrmann, J.M. (2007). Translocation of Proteins into Mitochondria. Annu. Rev. Biochem. *76*, 723–749.

New, A.M., Cerulus, B., Govers, S.K., Perez-Samper, G., Zhu, B., Boogmans, S., Xavier, J.B., and Verstrepen, K.J. (2014). Different Levels of Catabolite Repression Optimize Growth in Stable and Variable Environments. PLoS Biol. *12*, e1001764.

Newby, G.A., and Lindquist, S. (2013). Blessings in disguise: biological benefits of prion-like mechanisms. Trends Cell Biol. *23*, 251–259.

Newman, J.R.S., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature *441*, 840–846.

Nicol-Benoit, F., Amon, A., Vaillant, C., le Goff, P., le Dréan, Y., Pakdel, F., Flouriot, G., Valotaire, Y., and Michel, D. (2011). A Dynamic Model of Transcriptional Imprinting Derived from the Vitellogenesis Memory Effect. Biophys. J. *101*, 1557–1568.

Nicol-Benoît, F., Le-Goff, P., Le-Dréan, Y., Demay, F., Pakdel, F., Flouriot, G., and Michel, D. (2012). Epigenetic memories: structural marks or active circuits? Cell. Mol. Life Sci. *69*, 2189–2203.

Niepel, M., Spencer, S.L., and Sorger, P.K. (2009). Non-genetic cell-to-cell variability and the consequences for pharmacology. Curr. Opin. Chem. Biol. *13*, 556–561.

Norman, T.M., Lord, N.D., Paulsson, J., and Losick, R. (2015). Stochastic Switching of Cell Fate in Microbes. Annu. Rev. Microbiol. *69*, 381–403.

Nosaka, K., Esaki, H., Onozuka, M., Konno, H., Hattori, Y., and Akaji, K. (2012). Facilitated recruitment of Pdc2p, a yeast transcriptional activator, in response to thiamin starvation. FEMS Microbiol. Lett. *330*, 140–147.

Novick, A., and Weiner, M. (1957). Enzyme induction as an all-or-none phenomenon. Proc. Natl. Acad. Sci. *43*, 553–566.

O'Duibhir, E., Lijnzaad, P., Benschop, J.J., Lenstra, T.L., Leenen, D. van, Koerkamp, M.J.G., Margaritis, T., Brok, M.O., Kemmeren, P., and Holstege, F.C. (2014). Cell cycle population effects in perturbation studies. Mol. Syst. Biol. *10*, 732.

Olofsson, H., Ripa, J., and Jonzén, N. (2009a). Bet-hedging as an evolutionary game: the trade-off between egg size and number. Proc. R. Soc. Lond. B Biol. Sci. *276*, 2963–2969.

Olofsson, H., Ripa, J., and Jonzén, N. (2009b). Bet-hedging as an evolutionary game: the trade-off between egg size and number. Proc. R. Soc. Lond. B Biol. Sci. *276*, 2963–2969.

O'Rourke, S.M., Herskowitz, I., and O'Shea, E.K. (2002). Yeast go the whole HOG for the hyperosmotic response. Trends Genet. TIG *18*, 405–412.

Pande, K., Chen, C., and Noble, S.M. (2013). Passage through the mammalian gut triggers a phenotypic switch that promotes Candida albicans commensalism. Nat. Genet. *45*, 1088–1091.

Pastor, V., Luna, E., Mauch-Mani, B., Ton, J., and Flors, V. (2013). Primed plants do not forget. Environ. Exp. Bot. *94*, 46–56.

Pellegrino, M.W., and Haynes, C.M. (2015). Mitophagy and the mitochondrial unfolded protein response in neurodegeneration and bacterial infection. BMC Biol. *13*.

Petrenko, N., Chereji, R.V., McClean, M.N., Morozov, A.V., and Broach, J.R. (2013). Noise and interlocking signaling pathways promote distinct transcription factor dynamics in response to different stresses. Mol. Biol. Cell *24*, 2045–2057.

Philippi, T., and Seger, J. (1989). Hedging one's evolutionary bets, revisited. Trends Ecol. Evol. *4*, 41–44.

Pierce, S.E., Davis, R.W., Nislow, C., and Giaever, G. (2007). Genome-wide analysis of barcoded Saccharomyces cerevisiae gene-deletion mutants in pooled cultures. Nat. Protoc. *2*, 2958–2974.

Qian, W., Ma, D., Xiao, C., Wang, Z., and Zhang, J. (2012). The Genomic Landscape and Evolutionary Resolution of Antagonistic Pleiotropy in Yeast. Cell Rep. *2*, 1399–1410.

Raser, J.M. (2004). Control of Stochasticity in Eukaryotic Gene Expression. Science *304*, 1811–1814.

Ratcliff, W.C., and Denison, R.F. (2010). Individual-level bet hedging in the bacterium Sinorhizobium meliloti. Curr. Biol. CB *20*, 1740–4.

Raymond, R.K., Kastanos, E.K., and Appling, D.R. (1999). Saccharomyces cerevisiae Expresses Two Genes Encoding Isozymes of Methylenetetrahydrofolate Reductase. Arch. Biochem. Biophys. *372*, 300–308.

Razinkov, I.A., Baumgartner, B.L., Bennett, M.R., Tsimring, L.S., and Hasty, J. (2013). Measuring Competitive Fitness in Dynamic Environments. J. Phys. Chem. B *117*, 13175–13181.

Reed, T.E., Waples, R.S., Schindler, D.E., Hard, J.J., and Kinnison, M.T. (2010). Phenotypic plasticity and population viability: the importance of environmental predictability. Proc. R. Soc. Lond. B Biol. Sci. *277*, 3391–3400.

Rees, M., Jessica, C., Metcalf, E., and Childs, D.Z. (2010). Bet-hedging as an evolutionary game: the trade-off between egg size and number. Proc. Biol. Sci. *277*, 1149–51.

Richard, M., and Yvert, G. (2014). How does evolution tune biological noise? Front. Genet. *5*.

Rinott, R., Jaimovich, A., and Friedman, N. (2011). Exploring transcription regulation through cell-to-cell variability. Proc. Natl. Acad. Sci. *108*, 6329–6334.

Robinson, D.G., Chen, W., Storey, J.D., and Gresham, D. (2014). Design and Analysis of Bar-seq Experiments. G3 GenesGenomesGenetics *4*, 11–18.

Roesch, A., Fukunaga-Kalabis, M., Schmidt, E.C., Zabierowski, S.E., Brafford, P.A., Vultur, A., Basu, D., Gimotty, P., Vogt, T., and Herlyn, M. (2010). A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. Cell *141*, 583–594.

Romero-Santacreu, L., Moreno, J., Perez-Ortin, J.E., and Alepuz, P. (2009). Specific and global regulation of mRNA stability during osmotic stress in Saccharomyces cerevisiae. RNA *15*, 1110–1120.

Saarikangas, J., and Barral, Y. (2016). Protein aggregation as a mechanism of adaptive cellular responses. Curr. Genet. *62*, 711–724.

Sæther, B.-E., and Engen, S. (2015). The concept of fitness in fluctuating environments. Trends Ecol. Evol. *30*, 273–281.

Saito, H., and Posas, F. (2012). Response to Hyperosmotic Stress. Genetics *192*, 289–318.

Salathé, M., Van Cleve, J., and Feldman, M.W. (2009). Evolution of stochastic switching rates in asymmetric fitness landscapes. Genetics *182*, 1159–64.

Salaun, L., Snyder, L.A., and Saunders, N.J. (2003). Adaptation by phase variation in pathogenic bacteria. Adv. Appl. Microbiol. *52*, 263–302.

Satory, D., Gordon, A.J.E., Halliday, J. a, and Herman, C. (2011). Epigenetic switches: can infidelity govern fate in microbes? Curr. Opin. Microbiol. *14*, 212–7.

Scaduto, C.M., and Bennett, R.J. (2015). Candida albicans the chameleon: transitions and interactions between multiple phenotypic states confer phenotypic plasticity. Curr. Opin. Microbiol. *26*, 102–108.

Schiffers, K., Bourne, E.C., Lavergne, S., Thuiller, W., and Travis, J.M.J. (2013). Limited evolutionary rescue of locally adapted populations facing climate change. Phil Trans R Soc B *368*, 20120083.

Seger, J., and Brockman, J. (1987). What is bet-hedging?

Shan, X., Wang, L., Hoffmaster, R., and Kruger, W.D. (1999). Functional Characterization of Human Methylenetetrahydrofolate Reductase in Saccharomyces cerevisiae. J. Biol. Chem. *274*, 32613–32618.

Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., et al. (2010). A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. Cell *141*, 69–80.

Shorter, J., and Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. Nat. Rev. Genet. *6*, 435–450.

Si, K., Lindquist, S., and Kandel, E.R. (2003). A Neuronal Isoform of the Aplysia CPEB Has Prion-Like Properties. Cell *115*, 879–891.

Simons, A.M. (2009). Fluctuating natural selection accounts for the evolution of diversification bet hedging. Proc. R. Soc. B Biol. Sci. *276*, 1987–1992.

Smith, A.M., Heisler, L.E., Mellor, J., Kaper, F., Thompson, M.J., Chee, M., Roth, F.P., Giaever, G., and Nislow, C. (2009). Quantitative phenotyping via deep barcode sequencing. Genome Res. *19*, 1836–1842.

Smith, A.M., Heisler, L.E., St.Onge, R.P., Farias-Hesson, E., Wallace, I.M., Bodeau, J., Harris, A.N., Perry, K.M., Giaever, G., Pourmand, N., et al. (2010). Highly-multiplexed barcode sequencing: an efficient method for parallel analysis of pooled samples. Nucleic Acids Res. *38*, e142–e142.

Specchia, V., Piacentini, L., Tritto, P., Fanti, L., D'Alessandro, R., Palumbo, G., Pimpinelli, S., and Bozzetti, M.P. (2010). Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. Nature *463*, 662–665.

Steffen, P.A., and Ringrose, L. (2014). What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. Nat. Rev. Mol. Cell Biol. *15*, 340–356.

Stockwell, S.R., Landry, C.R., and Rifkin, S.A. (2014). The yeast galactose network as a quantitative model for cellular memory. Mol BioSyst *11*, 28–37.

Sumner, E.R., and Avery, S.V. (2002). Phenotypic heterogeneity: differential stress resistance among individual cells of the yeast Saccharomyces cerevisiae. Microbiology *148*, 345–351.

Svardal, H., Rueffler, C., and Hermisson, J. (2015). A general condition for adaptive genetic polymorphism in temporally and spatially heterogeneous environments. Theor. Popul. Biol. *99*, 76–97.

Szamecz, B., Boross, G., Kalapis, D., Kovács, K., Fekete, G., Farkas, Z., Lázár, V., Hrtyan, M., Kemmeren, P., Koerkamp, M.J.A.G., et al. (2014). The Genomic Landscape of Compensatory Evolution. PLOS Biol. *12*, e1001935.

Tagkopoulos, I., Liu, Y.-C., and Tavazoie, S. (2008). Predictive Behavior Within Microbial Genetic Networks. Science *320*, 1313–1317.

Teng, X., Dayhoff-Brannigan, M., Cheng, W.-C., Gilbert, C.E., Sing, C.N., Diny, N.L., Wheelan, S.J., Dunham, M.J., Boeke, J.D., Pineda, F.J., et al. (2013). Genome-wide Consequences of Deleting Any Single Gene. Mol. Cell *52*, 485–494.

Tuljapurkar, S., and Istock, C. (1993). Environmental Uncertainty and Variable Diapause. Theor. Popul. Biol. *43*, 251–280.

Tzagoloff, A., Wu, M.A., and Crivellone, M. (1986). Assembly of the mitochondrial membrane system. Characterization of COR1, the structural gene for the 44-kilodalton core protein of yeast coenzyme QH2-cytochrome c reductase. J. Biol. Chem. *261*, 17163–17169.

Vardi, N., Levy, S., Assaf, M., Carmi, M., and Barkai, N. (2013). Budding Yeast Escape Commitment to the Phosphate Starvation Program Using Gene Expression Noise. Curr. Biol. *23*, 2051–2057.

Veening, J.-W., Smits, W.K., and Kuipers, O.P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. Annu. Rev. Microbiol. *62*, 193–210.

Venable, D.L. (2007). Bet hedging in a guild of desert annuals. Ecology *88*, 1086–1090.

Venail, P.A., Kaltz, O., Olivieri, I., Pommier, T., and Mouquet, N. (2011). Diversification in temporally heterogeneous environments: effect of the grain in experimental bacterial populations. J. Evol. Biol. *24*, 2485–2495.

Venkataram, S., Dunn, B., Li, Y., Agarwala, A., Chang, J., Ebel, E.R., Geiler-Samerotte, K., Hérissant, L., Blundell, J.R., Levy, S.F., et al. (2016). Development of a Comprehensive Genotype-to-Fitness Map of Adaptation-Driving Mutations in Yeast. Cell *166*, 1585–1596.e22.

Viney, M., and Reece, S.E. (2013). Adaptive noise. Proc. R. Soc. B Biol. Sci. *280*, 20131104–20131104.

Wang, F., and Higgins, J.M.G. (2013). Histone modifications and mitosis: countermarks, landmarks, and bookmarks. Trends Cell Biol. *23*, 175–184.

West Jr., R.W., Crivellone, M.D., Ma, J., and Thomas, S. (1996). Sequence of the Saccharomyces cerevisiae YTP1 gene encoding a deduced novel type-III integral membrane protein with domains of sequence similarity to mitochondrial electron-transport enzymes. Gene *169*, 119–124.

Wolf, D.M., Vazirani, V.V., and Arkin, A.P. (2005). Diversity in times of adversity: probabilistic strategies in microbial survival games. J. Theor. Biol. *234*, 227–53.

van der Woude, M.W. (2011). Phase variation: how to create and coordinate population diversity. Curr. Opin. Microbiol. *14*, 205–11.

Xiong, W., and Ferrell, J.E. (2003). A positive-feedback-based bistable “memory module” that governs a cell fate decision. Nature *426*, 460–465.

Yamamichi, M., Yoshida, T., and Sasaki, A. (2011). Comparing the Effects of Rapid Evolution and Phenotypic Plasticity on Predator-Prey Dynamics. Am. Nat. *178*, 287–304.

Yang, J., McCormick, M.A., Zheng, J., Xie, Z., Tsuchiya, M., Tsuchiyama, S., El-Samad, H., Ouyang, Q., Kaeberlein, M., Kennedy, B.K., et al. (2015). Systematic analysis of asymmetric partitioning of yeast proteome between mother and daughter cells reveals “aging factors” and mechanism of lifespan asymmetry. Proc. Natl. Acad. Sci. *112*, 11977–11982.

Zacharioudakis, I., Gligoris, T., and Tzamarias, D. (2007). A Yeast Catabolic Enzyme Controls Transcriptional Memory. Curr. Biol. *17*, 2041–2046.

Zadeh, L.A., and Deoser, C.A. (1976). Linear system theory (Robert E. Krieger Publishing Company Huntington).

Zakrzewska, A., van Eikenhorst, G., Burggraaff, J.E.C., Vis, D.J., Hoefsloot, H., Delneri, D., Oliver, S.G., Brul, S., and Smits, G.J. (2011). 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 *22*, 4435–46.

Zhang, X., and Rainey, P.B. (2010). Bet hedging in the underworld. Genome Biol. 10–12.

Zhang, Z., Qian, W., and Zhang, J. (2009). Positive selection for elevated gene expression noise in yeast. Mol. Syst. Biol. *5*.

Zuk, O., Hechter, E., Sunyaev, S.R., and Lander, E.S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. Proc. Natl. Acad. Sci. *109*, 1193–1198.

Résumé de chaque partie en Français

1. Introduction

1-I. Génétique de la levure

La levure est un organisme modèle très répandu du fait de sa similarité avec les cellules humaines et de la simplicité et puissance des expériences réalisées avec cet organisme unicellulaire. Un certain nombres d'outils permettent d'étudier les gènes de la levure. Notamment, plusieurs banques de délétions systématiques (Yeast Deletion Library) ont été créées. Celles-ci contiennent des milliers de souches (haploïdes ou diploïdes) dans lequel un gène a été supprimé ou atténué.

1-II. Adaptation a un seul changement environnemental

Lorsqu'ils sont confronté à de nouveaux environnements, les organismes s'adaptent. Ce processus d'adaptation peut être rapide, par un ajustement physiologique du métabolisme. Ou bien long, par l'apparition successive d'organismes avec des mutation bénéfiques, c.a.d. qui leur donnent un avantage sélectif. Pour les organismes unicellulaires, cet avantage sélectif est mesuré par la prolifération cellulaire (fitness) d'une souche dans un environnement donné. La baisse des coûts du séquençage a permit de séquencer des génomes entiers (Whole Genome Sequencing) avant et après sélection dans un environnement donné afin d'identifier les mutations bénéfiques.

1-III. Adaptation a des environnements fluctuants

Dans la nature, les organismes sont confrontés à des environnements qui fluctuent. Une stratégie évolutive très répandues pour faire face aux environnements fluctuants est la répartition des risques (bet-hedging). Pour un organisme, cette stratégie consiste à minimiser ses chances d'extinctions en réduisant le coût sélectif des environnements défavorables, au détriment d'une fitness moins élevés dans certains environnements.

Par ailleurs, un facteur important pour la survie des espèces unicellulaires en environnements changeants est la mémoire des événements passés. Ainsi des mécanismes moléculaires permettent aux levures de garder en mémoire l'apparition d'un environnement stressant, ceci afin de réagir plus vite lors de la prochaine confrontation avec cet environnement. L'anticipation cellulaire, via l'expression de gènes en anticipation d'un potentiel future stress, pourrait également être un facteur important pour la survie en environnements fluctuants.

1-IV. Des fluctuations artificielles afin de caractériser des dynamiques biologiques

Un certains nombres d'études ont cultivés des drosophiles en conditions statiques ou hétérogènes pendant des centaines de générations. Malgré des résultats contrastés, plusieurs études ont montré que la variance génétique, et donc la vitesse d'adaptation, était plus élevée en conditions dynamiques.

Les méthodes d'identification des systèmes (SI) s'appliquent sur des systèmes linéaires, c.a.d. des systèmes qui respectent les propriétés d'homogénéité (pour une fonction f : $f(A * \alpha) = \alpha * f(A)$) et d'additivité ($f(A) + f(B) = f(A+B)$). Elles consistent à stimuler périodiquement un système afin de déterminer son comportement entrées-sorties. Une poignée d'études ont appliquées les méthodes de SI

sur la levure, afin de caractériser les dynamiques d'actions de certains réseaux biologiques bien connus (Hog, Gal). Ainsi ces études ont montré que le réseau de réponse au stress hyper-osmotique (HOG) a une plus grande réactivité que le réseau de métabolisme du galactose.

1-V. Génomique de la croissance en environnements fluctuants periodiques

Faire pousser des microbes dans des environnements fluctuants de manière periodique a permit de révéler plusieurs phénomènes tels que: la mémoire de l'adaptation, l'amplification du temps de latence (lag time) pour l'adaptation a un environnement, ou des stratégies de répartition des risques (bet-hedging). Mon projet a consisté à mesurer la croissance (fitness) d'une banque de délétion de levures homozygote en conditions de fluctuation environnementales périodiques, mais également en conditions statiques. L'objectif principal était de déterminer si certains gènes sont avantageux et/ou désavantageux en conditions dynamiques. Par ailleurs, une autre objectif était de déterminer si la fitness en conditions dynamiques peut être prédite par la fitness en conditions statiques.

2. Méthodes

Description de méthodes non utilisées dans la publications soumise.

3 Résultats

3-I. Publication Soumise

La banque de délétion a été cultivée en conditions de léger stress salin périodique, avec 5 périodes de fluctuations (de 6h à 48h) et dans deux conditions statiques (milieu salin ou standard). Plus de la moitié des mutants avaient une fitness non prédictible en conditions fluctuantes, dont la plupart avaient une fitness plus haute qu'attendue en environnements dynamiques. Les mutants avec des fitness les plus inattendues en conditions périodiques étaient des mutants pour des gènes impliqués dans des fonctions variées. Cependant un enrichissement fut observe pour des gènes impliqué dans la régulation transcriptionnelle, la réponse au stress, et le stress osmotique. Deux mutants ont vu leur avantage en environnement fluctuant disparaître après l'insertion du gène supprimé à un autre locus, ce qui valide que la fonction du gène est importante pour ce phénotype. Les gènes concernés sont: TOM7, un gène important pour la biogenèse du complexe Tom qui régule l'importation de protéines dans les mitochondries, et PDE2, la principale phospho-diesterase chez la levure qui contrôle le taux d'AMP cyclique et donc l'activité de la protéine kinase A (PKA).

3-II. Résumé du projet

Le projet avait deux étapes expérimentales, avec tout d'abord un criblage a haut débit d'une banque de délétion homozygote par le biais de la technique de BarSeq. Cette technique, à fort multiplexage, consiste a séquencer le code barre unique de chaque mutant afin de déterminer son abondance dans une population de mutants cultivés ensemble. La deuxième étape consistait a mettre en compétition des souches mutantes d'intérêt avec une souche sauvage et de mesurer la croissance de ces souches via un cytomètre en flux. Cette expérience est plus précise que le séquençage et a permit de valider les résultats génomiques. Si pour un mutant, la fitness en condition périodiques est la moyenne géométrique de sa fitness en conditions statiques, alors ce mutant est désigné comme étant linéaire. Sinon celui-ci est non-linéaire.

3-III. Quatre cribles génomiques

L'expérience a été répété dans 3 autre types de milieux fluctuants (fluctuations en sodium métabisulfite, glucose et méthionine) et à d'autres périodes de fluctuations (de 6h a 60h). Différentes étapes de normalisation et de contrôles qualités ont été appliqué. Deux rondes de séquençages ont été effectuées afin d'identifier des mutants avec des effets subtils. Un Modèle Linéaire Généralisé (GLM) a été mis au point afin de d'identifier les mutants non-linéaires. Un fort enrichissement en mutants non-linéaire a été observé pour les périodes les plus courtes, uniquement pour l'expérience de fluctuations salines. Certains mutants ont montre un phénotype intéressant en conditions fluctuantes, tel que: une fitness plus basse qu'attendue aux fluctuations rapides, une fitness haute aux fluctuations rapides et basses aux fluctuations lentes, ou un fitness similaire dans différentes paires d'environnements statiques mais différentes en conditions fluctuantes.

3-IV. Crible secondaire: mesures de mutants individuels en stress salin périodique

Des expériences utilisant la cytométrie en flux ont permit de détailler l'avantage de croissance de nombreux mutants observé en fluctuations salines. Ces expériences suggèrent que cet avantage est un avantage relatif et non absolu, ou alors un avantage absolu léger. Autrement dit, ces mutants ont une croissance plus élevée en environnement standard qu'en environnement fluctuant, cependant ils ont une croissance plus élevé en environnements fluctuants que la souche sauvage. Par ailleurs, l'avantage des mutants *pde2* et *tom7* en fluctuations saline n'était pas systématique lorsque la composition des milieux fluctuants variait.

3-V. Analyse de données de co-fitness pour grouper des mutants avec des profiles phénotypiques similaires

Des études ont exposé les banques de délétions à des milliers de petites molécules et ont mesuré leur croissance. Les mutants qui ont un comportement similaire dans un grand nombre de conditions (mesuré par la corrélation des valeurs de fitness, la co-fitness) sont mutants pour des gènes souvent impliqués dans des fonctions similaires. J'ai analysé ces jeux de données publiés afin de faire de l'inférence fonctionnelle pour les mutants avec la plus grande non-linéarité de la fitness. La plupart des mutants avec les plus grandes valeurs de non-linéarité de la fitness étaient fortement corrélés entre eux, avec plusieurs groupes distincts de mutants corrélés. Le fonction commune des mutants dans chaque groupe, si elle existe, reste largement a élucider. Cependant, un de ces groupes pourrait regrouper des mutants pour des gènes impliqués dans les mitochondries. Un autre groupe semble regrouper des mutants pour des gènes partageant une localisation proche dans le génome.

3-VI. Déviations entre la fitness attendue et mesurée: coûts et bénéfices d'ignorer le stress environnemental

De manière intéressante un certain nombre de mutant non-linéaires présentaient une croissance élevées en condition saline après 1 jour, avant de voir leur croissance diminuer fortement par la suite. Cet effet suggère que ces mutants ignore les signaux indiquant un changement d'environnement, et ne lancent pas les voies de réponses au stress salin.

4. Discussion

4-I. L'insensibilité environnementale peut être une stratégie d'adaptation en stress salin periodique artificiel

Une adaptation rapide à un nouvel environnement n'est pas toujours bénéfique; notamment si ce nouvel environnement disparaît rapidement. Mes résultats suggèrent qu'ignorer le stress salin peut être une stratégie efficace lorsque l'environnement change rapidement.

4-II. Une qualité de données faible dans les autres expériences empêchent de tirer des conclusions sur les principes généraux de la fitness en environnements périodiques

Mon projet avait deux objectifs principaux: identifier des gènes avec une fitness fortement non-linéaire, et déterminer le niveau de linéarité de la fitness au niveau génomique. En conditions salines, un grand nombre de mutants ont une fitness non-linéaire. Il est difficile de comparer ces résultats avec les autres type de fluctuations de milieux testés car la qualité de données était sous-optimale (expériences en méthionine et glucose) ou bien le stress utilisé était un peu trop spécifique à certains fonction cellulaires et ne modulait pas la fitness d'un grand nombre de mutants (expérience en sodium métabisulfite).

4-III. Hypothèses pour expliquer le comportement de certains mutants non-linéaires

Différents mécanismes sont proposé pour expliquer le comportement de certains mutants non-linéaires. Une hypothèse proposée est qu'une réduction de la respiration cellulaire pourrait être avantageuse en fluctuations salines périodiques. Plusieurs mutants non-linéaires ont des gènes supprimé qui codent pour des protéines avec une courte ou longue demi-vie. Des hypothèses sont proposées afin d'expliquer cette observation.

4-IV. Perspectives

Ce projet a permit d'identifier différents mutants avec une fitness fortement non-linéaire en fluctuations salines. De futures études en cellules-uniques permettront de caractériser plus finement le comportement de ces mutants, dont notamment : leurs stratégies d'adaptations, ainsi que leur implication dans des phénomènes de mémoire du stress. Des directions de recherches pour identifier des facteurs qui modulent la non-linéarité de la fitness sont proposées. Finalement, de nouveaux types d'expériences sont proposées afin d'étudier les interactions entre les gènes, l'environnement et le temps.

5. Conclusions

Pour conclure, de nombreux mutants non-linéaire ont été identifiés en fluctuations salines. De manière inattendue, ignorer la présence du stress peut être une stratégie avantageuse pour les levures dans ces conditions artificielles. Ce crible à haut débit ouvre la porte a plusieurs questions mécanistiques. Notamment, des études en cellules uniques pourraient permettre de mieux caractériser certains mutants non-linéaires, et les facteurs qui modulent la non-linéarité de la fitness.