

RESEARCH

Open Access



Diversity of flux distribution in central carbon metabolism of *S. cerevisiae* strains from diverse environments

Thibault Nidelet*, Pascale Brial, Carole Camarasa and Sylvie Dequin

Abstract

Background: *S. cerevisiae* has attracted considerable interest in recent years as a model for ecology and evolutionary biology, revealing a substantial genetic and phenotypic diversity. However, there is a lack of knowledge on the diversity of metabolic networks within this species.

Results: To identify the metabolic and evolutionary constraints that shape metabolic fluxes in *S. cerevisiae*, we used a dedicated constraint-based model to predict the central carbon metabolism flux distribution of 43 strains from different ecological origins, grown in wine fermentation conditions. In analyzing these distributions, we observed a highly contrasted situation in flux variability, with quasi-constancy of the glycolysis and ethanol synthesis yield yet high flexibility of other fluxes, such as the pentose phosphate pathway and acetaldehyde production. Furthermore, these fluxes with large variability showed multimodal distributions that could be linked to strain origin, indicating a convergence between genetic origin and flux phenotype.

Conclusions: Flux variability is pathway-dependent and, for some flux, a strain origin effect can be found. These data highlight the constraints shaping the yeast operative central carbon network and provide clues for the design of strategies for strain improvement.

Keywords: Diversity, Flux balance analysis, Metabolic flux, Modeling, *S. cerevisiae*

Background

Cellular metabolism entails a large number of reactions that are involved in the conversion of various resources into precursors and energy for biosynthesis and cellular compounds. The rates of these reactions, i.e. fluxes, reflect metabolic activity through the operative network. Fluxes are the combined result of regulation at many different biological levels, such as transcription, translation, post-translational protein modification and protein–protein interactions. Therefore, metabolic fluxes are a global representation of the cellular phenotype expressed under specific conditions; thus, analyzing flux distribution is a valuable approach to study cell metabolism [1].

While intracellular fluxes are difficult to measure experimentally, they can be predicted by different methods

that rely on constraint-based models (CBM) that formalize the metabolic network as a stoichiometry matrix. These CBM range from small networks focused on a specific aspect of cellular metabolism to genome-scale models that include all reactions of a given organism. The first step to solve these systems and predict fluxes from these networks is to add constraints on the input and output fluxes. Depending on the number of constraints and the size of the network, it is possible to estimate the fluxes in some cases; this approach is referred to as metabolic flux analysis (MFA). However, in most cases, adding constraints only on input and output data is not sufficient; therefore, there are two possibilities: the ^{13}C -MFA [2] and the flux balance analysis (FBA), [3]. In the ^{13}C -MFA approach, cells are fed ^{13}C -labeled glucose, and the analysis of the subsequent ^{13}C enrichment in different amino-acids generates experimental data that can be used to constrain internal fluxes and therefore estimate

*Correspondence: nidelet@supagro.inra.fr
SPO, INRA, SupAgro, Université de Montpellier, 34060 Montpellier, France

intracellular fluxes [1, 2]. By contrast, the FBA is based on the choice of an optimal solution in the space of possible solutions defined by the constraint stoichiometry matrix. This solution will optimize an objective function [3]; therefore, the predicted flux distribution depends on the objective function that is used [4–6]. Objective functions commonly used are maximization of ATP production [7], minimization of metabolic adjustment [8, 9] or, most frequently, maximization of biomass production [10, 11]. These objective functions appear to be more or less effective depending on the conditions, constraints and models, without one of them emerging in particular [6].

In a previous study, ^{13}C -MFA and FBA approaches have been used to predict intracellular fluxes of central carbon metabolism of *S. cerevisiae* in conditions where the intracellular redox balance is modified [12]. Comparable relative changes between environments were obtained regardless of the predicting method, even if some flux predictions differed, in particular for the pentose phosphate pathway (PPP) [12].

Understanding how metabolic fluxes are modulated by environmental and/or genetic perturbations is a central question to understanding cellular physiology. For example, the FBA approach has been used to study the flux distribution sensitivity of *S. cerevisiae* wine yeast to environmental conditions, including various glucose concentrations, temperature or acetoin levels [9, 13]. In these studies, the PPP was one of the most variable fluxes, while the glycolytic flux remained virtually unchanged. These approaches have also been widely used to study network robustness and the effects of deletion mutants [14–16]. For example, using a ^{13}C flux approach in *S. cerevisiae*, Blank et al. [17] have shown that network redundancy through duplicate genes is a major determinant of genetic network robustness (75 %), while alternative pathways contribute to a lesser extent (25 %). Using a similar approach, Velagapudi et al. [18] studied the effect of knockout strains on the rerouting of metabolic fluxes in glucose and galactose media, highlighting interesting links between pathways, such as a positive correlation between flux through the PPP and biomass yield.

Flux prediction has also been used to guide metabolic engineering and strain improvement strategies [19, 20]. For instance, Bro et al. used CBM to predict the best possible metabolic engineering strategies to increase ethanol yield [21]. Guided by a genome scale model, they developed a strain with a glycerol yield reduced by 40 % and an ethanol yield increased by 3 % without affecting growth. Other examples include the prediction of strategies to optimize the yields of purine [5], succinic acid [20, 22] or proline [23].

The estimation of metabolic fluxes was also used in a few studies to investigate the divergence of flux distribution among species. ^{13}C flux analysis has been used to compare flux distributions in central carbon metabolism for pairs of species, including *S. cerevisiae* and *Phaffia rhodozyma* [24] or *S. cerevisiae* and *Pichia stipitis* [25], highlighting differences in the relative flux distribution, especially for the PPP. Using ^{13}C flux analysis, Blank et al. [17] and Christen and Sauer [26] studied the diversity of flux distributions in fourteen and seven yeast species, respectively. In both studies, similar correlations were shown between metabolic pathways, in particular, a trade-off between glycolysis and TCA fluxes and a positive correlation between biomass production and flux through the PPP.

In recent years, tremendous knowledge has been gained regarding the genetic and phenotypic diversity of *S. cerevisiae* [27–34]. The phenotypic diversity in these studies has mainly been addressed by the comparison of growth rate patterns in various media. Several other studies have begun to characterize the diversity of more various phenotypic traits. Spor et al. [35] have studied the phenotypic diversity of six life-history traits and three metabolic traits of different strains of *S. cerevisiae*, and they have identified two main life-history strategies, the “ants” and “grasshoppers,” which are characterized by divergence in cell size, reproductive rate and carrying capacity. A wider phenotypic analysis, performed with 72 *S. cerevisiae* strains from different origins and studying seven life-history traits and eleven metabolic traits, showed that strain origin has a wide impact on phenotypes [36]. Other studies have focused on nitrogen availability [37] or bio-ethanol-related traits [38].

Thus, the intra-species diversity of flux distribution remains unexplored. Studying the diversity of metabolism, particularly of metabolic fluxes, is fundamental to understanding the constraints and regulations that shape strain phenotypes. The functional and regulatory properties of yeast central carbon metabolism (CCM) determine most of the phenotypic traits relevant for various industrial processes, including food and beverage production (wine, bread, beer, cheese etc.), bioethanol or the use of yeast as a cell factory. For example, the fermentation rate, ethanol yield or production of acetate, and even aroma production are all dependent on carbon metabolism.

Thus, understanding how metabolic constraints structure metabolic pathways may enable a better exploitation of this diversity for industrial biotechnology. The objective of this study was to characterize the diversity of metabolic fluxes in a large set of *S. cerevisiae* strains from different genetic and ecological origins. To this end, we used a FBA approach to predict flux distribution for 43 strains of *S. cerevisiae* from six different ecological

origins: bread, rum, wine, flor, Mediterranean and American oak. The analysis of flux distribution dataset enabled us to identify the most flexible/robust fluxes and several correlations or trade-offs between metabolic pathways. In addition, we analyzed the flux structuration to strain origin in order to observe a possible convergence.

Results

In this work, we used *DynamoYeast*, a previously developed constraint-based model of central carbon metabolism [9], to study the diversity of metabolic flux distributions for 43 strains of six different ecological origins: “Bread,” “Rum,” “Wine,” “Flor,” “Mediterranean Oak” (Med_Oak) and “American Oak” (Oak). This model comprises the cytosol, mitochondria and extracellular medium and includes upper and lower glycolysis, the PPP, the synthesis of glycerol, the synthesis of ethanol, and the reductive and oxidative branches of the TCA as the main metabolic pathways (Fig. 1).

Fermentation was performed for all strains in a synthetic medium simulating grape must, containing high sugar and low nitrogen concentrations. Typical wine fermentation comprises a lag phase, a growth phase of approximately 24–36 h followed by a stationary phase, during which most of the sugar is fermented (reviewed in Marsit and Dequin [39]). We measured the production of biomass and metabolites, including ethanol, glycerol, acetate, succinate, pyruvate and alpha-ketoglutarate during the growth phase (at 11 g/L CO₂ released), which can be considered as steady state (a prerequisite to CBM). These experimental data ($\pm 2.5\%$) were used to constrain the model as upper and lower bound to then perform a flux balance analysis (FBA).

The FBA consists of choosing the best solution for the objective function in the space of possible fluxes. Instead of using an optimization that maximizes biomass flux, which is frequently used in FBA studies, we chose to minimize the glucose input, allowing us to use the experimental biomass as a constraint for the model. By making this optimization choice, we considered that the yeasts were optimal, in that they used the least amount of resources (here the glucose input) to produce biomass and fermentation byproducts. This strategy also has the advantage of optimizing the modeling approach by maximizing the use of available experimental data. Using this approach, we obtained a flux distribution for 68 fluxes of the central carbon metabolism for each strain, expressed as relative fluxes normalized to the specific glucose uptake in the corresponding strain.

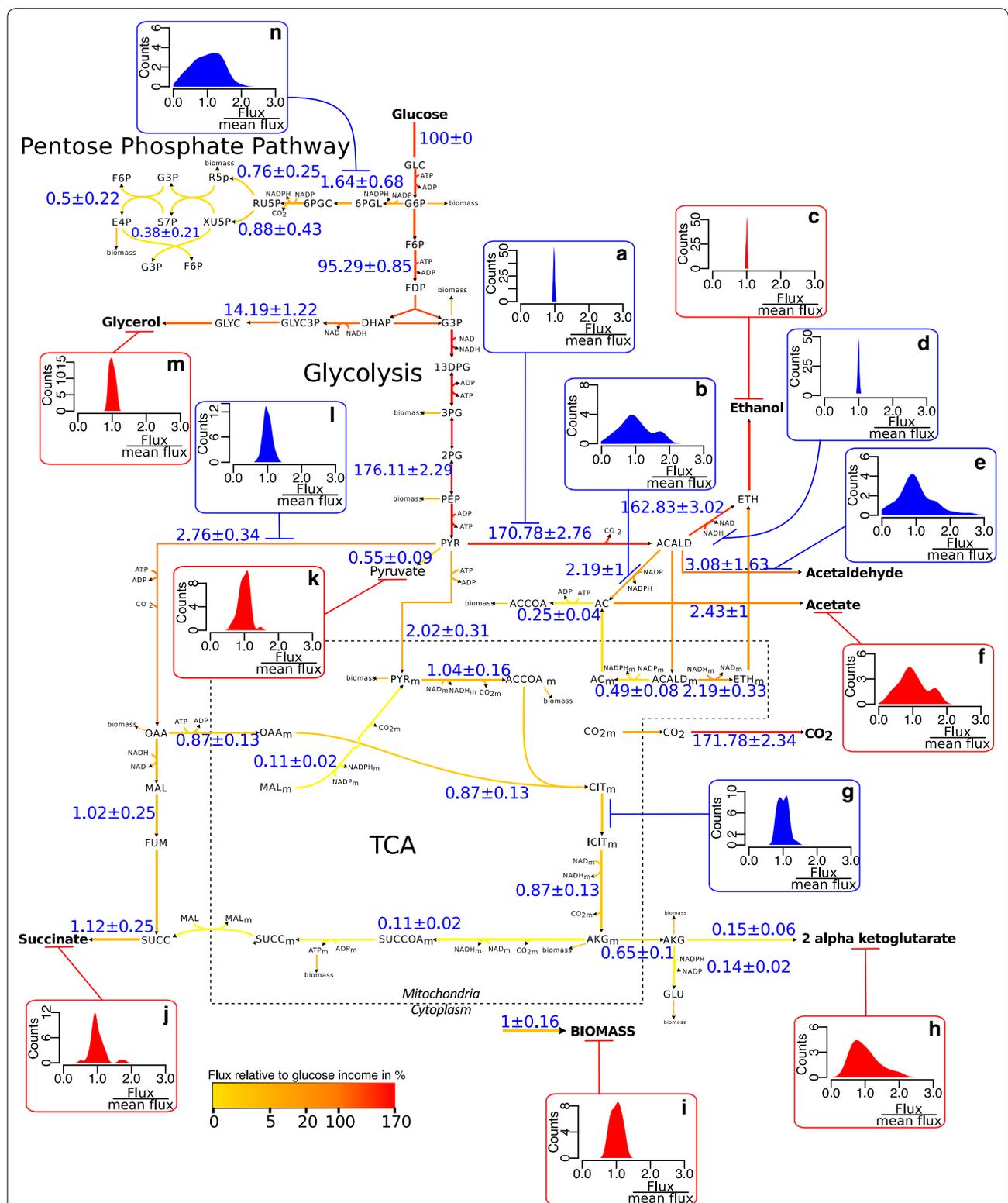
In this type of optimization, the given solution is often not the only one that meets the optimization criterion; i.e., different possible pathways are perfectly equivalent for the optimization criteria. We thus decided to

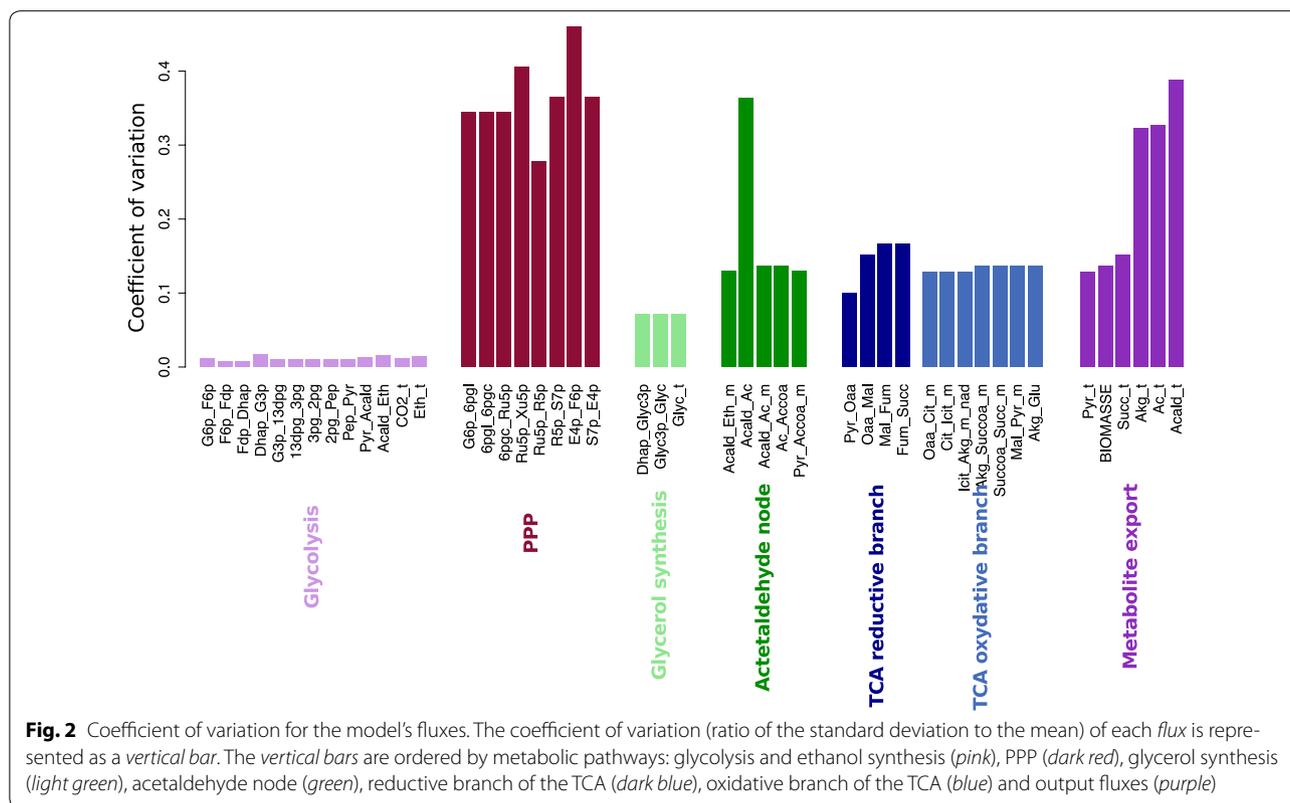
characterize all equivalent solutions to determine the fluxes that varied most between alternative solutions, which would therefore correspond to poorly predicted fluxes. To achieve this, we first fixed the input and output fluxes to the exact values predicted by the FBA, and we then used the “*enumerateOptimalSolution*” algorithm from the *cobra* toolbox [40] to identify all alternative solutions. For the large majority of fluxes, we found only one predicted value, except for the fluxes of the reductive branch of the TCA involved in the conversion of malate to fumarate and then to succinate, for which two solutions were identified. Indeed, these fluxes can be cytoplasmic or mitochondrial, which had no effect on the other fluxes predicted by the model, as the transport between these two compartments of the metabolites was free in our model. Setting either option to zero suppressed the alternative solution. We finally retained the solution going through the cytoplasm, which involved fewer reactions (no mitochondrial transport).

Then, we considered the biological variance between strains to identify the more robust and variable fluxes of the central carbon metabolism by studying the individual flux distributions (Fig. 1) and by comparing the variation coefficients (the ratio of the standard deviation to the mean) between fluxes (Fig. 2). Substantial differences were found in the variability of fluxes depending on the metabolite pathways (Fig. 2). The glycolysis and ethanol synthesis pathways displayed almost no variation (e.g. Pyr_Acald: $170.78 \pm 2.76\%$, Fig. 1a; Acald_Eth: $162.83 \pm 3.02\%$, Fig. 1d). The reductive and oxidative branches of the TCA (e.g. Cit_Icit_m: $1.02 \pm 0.24\%$, Fig. 1g; Pyr_Oaa: $2.76 \pm 0.34\%$, Fig. 1l), the glycerol synthesis pathway (e.g. Glyc_t: 14.41 ± 1.29 , Fig. 1m) and the biomass synthesis (BIOMASS: $1.02 \pm 0.18\%$, Fig. 1i) displayed a moderate variation. By contrast, the PPP pathway was the highest variable pathway (e.g. G6p_6pgl: $1.64 \pm 0.68\%$, Fig. 1n).

The acetaldehyde node displayed a particular pattern as it includes individual fluxes with very different variabilities (Fig. 2): besides the invariant synthesis of ethanol, the synthesis of acetate was highly variable with a wide bimodal distribution (Acald_Ac: $2.19 \pm 1\%$, Fig. 1b). The acetate output (Ac_t: $2.43 \pm 1\%$, Fig. 1f) and the excretion of acetaldehyde (Acald_t: $3.08 \pm 1.63\%$, Fig. 1e) were also highly variable.

Then, we searched for potential links between fluxes by studying all correlations between the model’s fluxes (Fig. 3). This approach first highlighted a “pathway block” structure, where fluxes were highly correlated to each other and operated almost like a single flux. For example, all the fluxes of the PPP displayed a Pearson correlation coefficient between them greater than 0.985 (Fig. 3). We identified seven blocks: upper glycolysis, lower glycolysis,



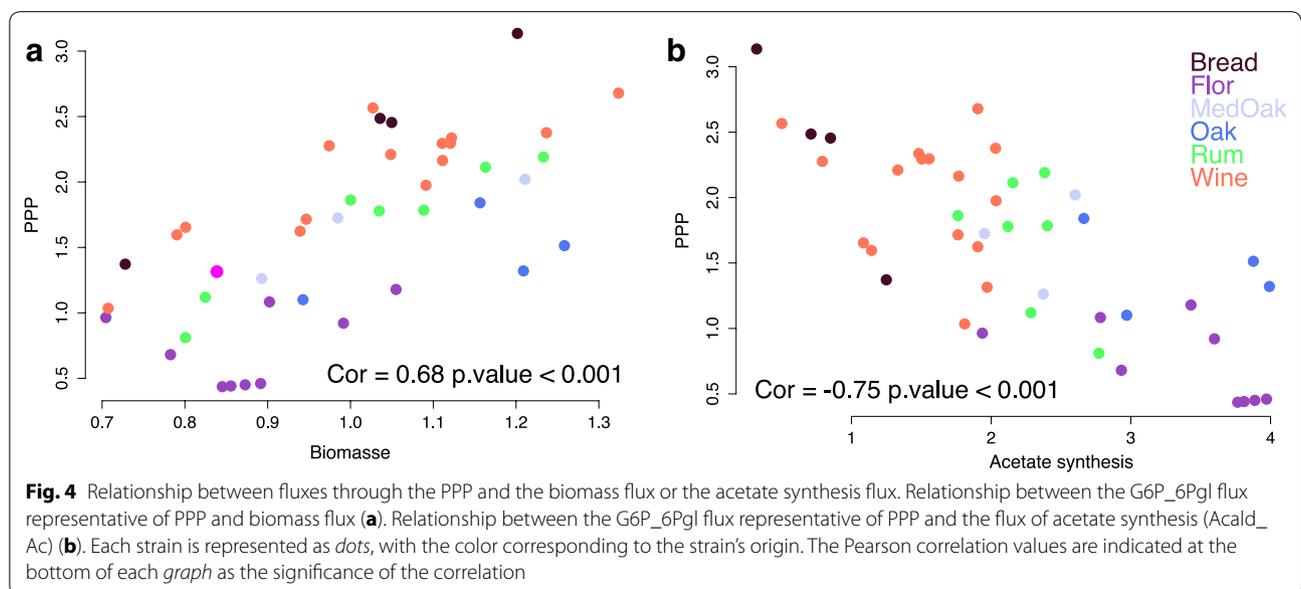
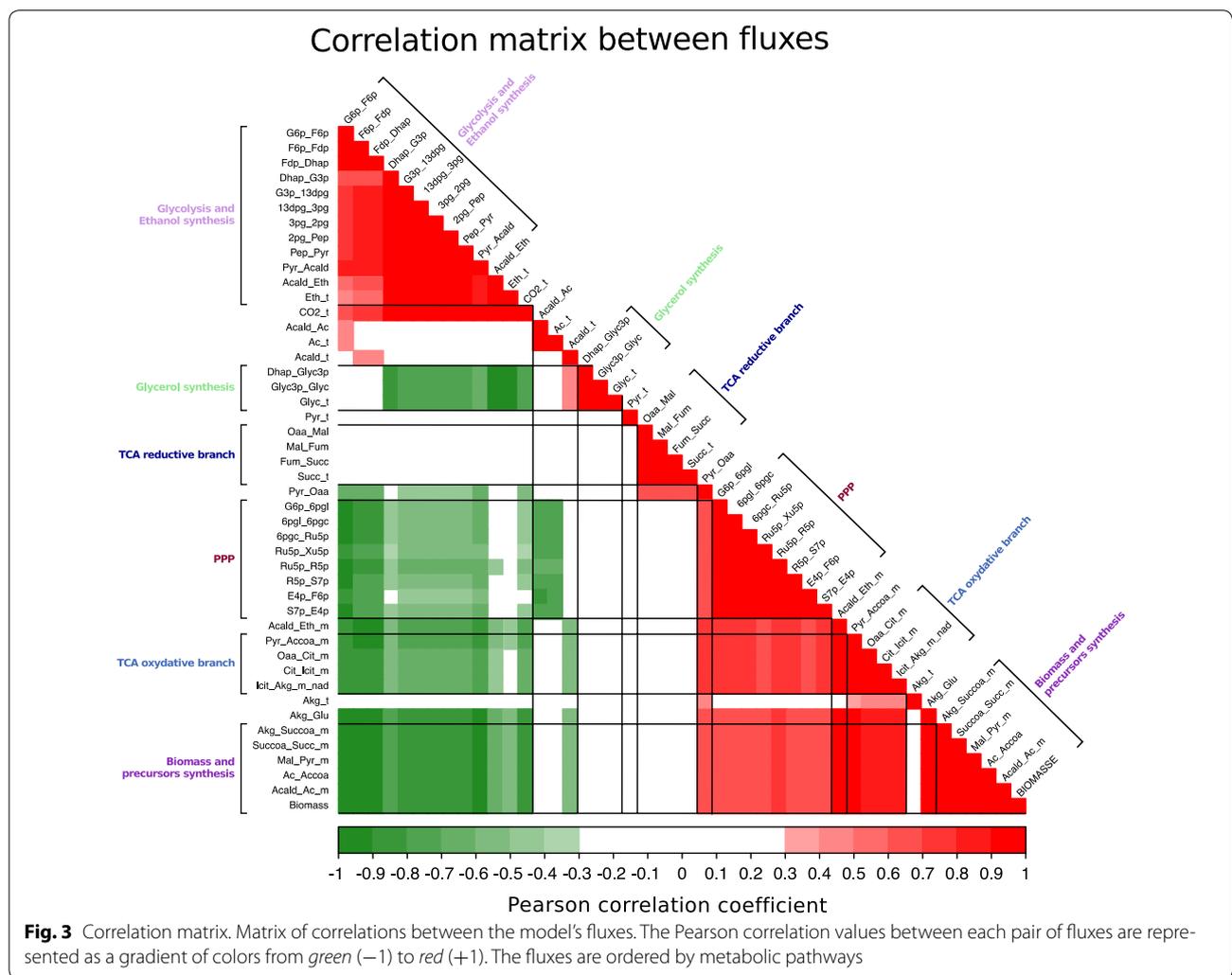


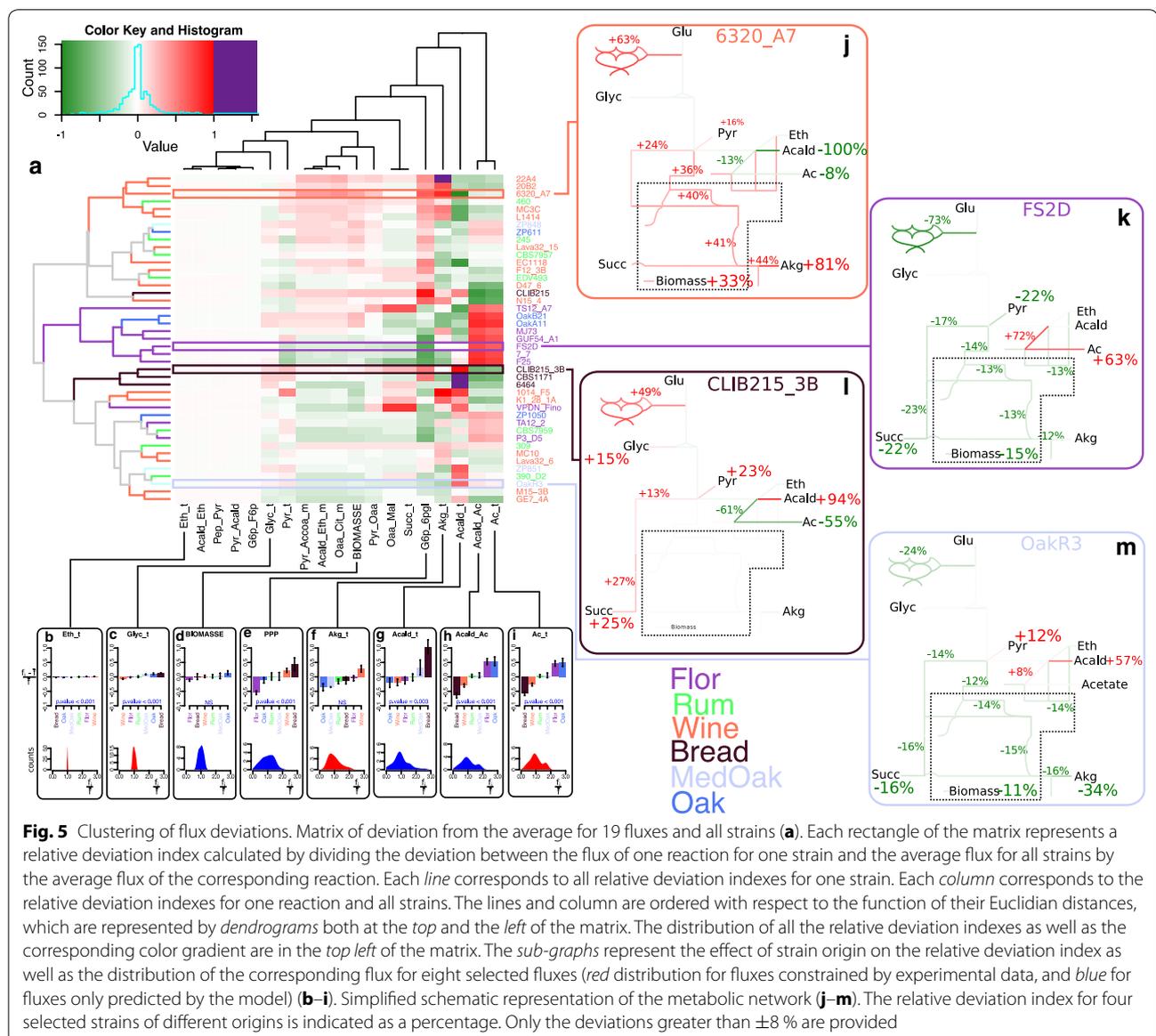
glycerol synthesis, the TCA reductive branch, the PPP, the TCA oxidative branch and the biomass block. The latter included the biomass synthesis reaction and all the fluxes that were only used to produce one of the biomass precursors. For example, cytoplasmic acetyl-CoA was only used in the model as a precursor of biomass (because the model never predicted its mitochondrial transport). Thus, the flux of acetyl-CoA synthesis (Ac_Accoa) was perfectly correlated with biomass synthesis (Fig. 3).

We also found correlations between blocks that had two main origins. In first case, these correlations were compulsory due to the model structure. For example, there was an expected negative correlation between the glycerol fluxes and the lower part of glycolysis because these two pathways diverged from the upper part of glycolysis. For the same reason, the flux through the PPP was negatively correlated with upper glycolysis. Positive correlations were also found between the PPP (Fig. 4a), the TCA oxidative branch and the biomass block, which could be connected to the synthesis of biomass precursors, such as Erythrose-4-phosphate (E4P), Ribose-5-phosphate (R5p) and alpha-ketoglutarate (AKG). Other correlations were independent of the network structure and emerged from the biological data. For example, a correlation was found between the fluxes through PPP and

acetate synthesis (Acald_Ac, Fig. 4b). This strong negative correlation was identified using the whole strain data set ($r = -0.76$, Fig. 4b). This trade-off could be linked to the synthesis of NADPH that can be achieved by these two pathways. Approximately 60 % of the NADPH demand is supplied by the PPP, but this proportion varied between 95.7 and 18.8 % depending on the strains, independently of the total production (Additional file 1: Figure S1). It is interesting to note that this trade-off did not appear in the model's null space of possible fluxes, which indicates that this correlation is independent of the network matrix and is purely biological.

Because the fluxes were mostly organized in blocks (Fig. 3), we decided to use only a subset of fluxes containing one representative flux for each block for further analysis. With this subset of 19 fluxes, we studied the deviation of each strain from the average for each flux. Then, we used a clustering method to classify the strains and fluxes as a function of their Euclidean distance (Fig. 5a). The fluxes that best separated strains were the most variable and also had binomial distributions, indicating very different behaviors across strains (Fig. 5b–i). The fluxes of acetate synthesis (Fig. 5h) and output (Fig. 5i) could separate one particular cluster of eight strains that was mainly characterized by a high production of acetate and a small flux through the PPP.





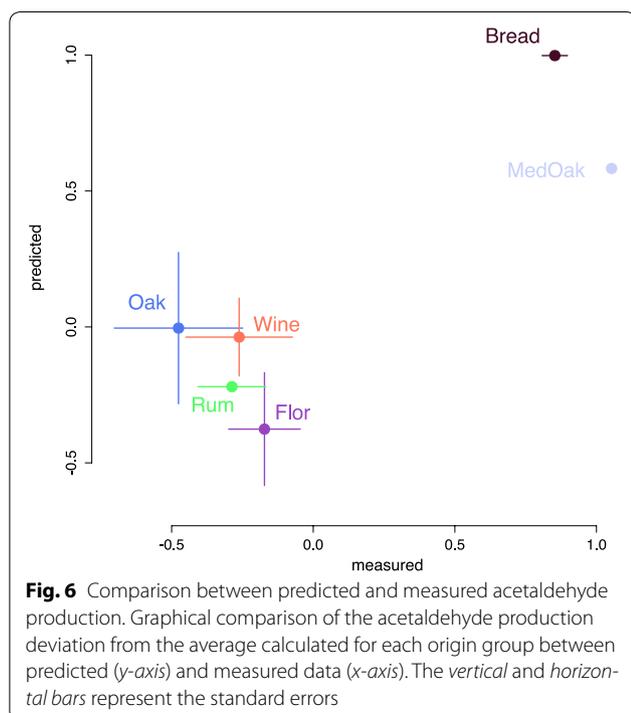
The strain FS2D (Fig. 5k) of this cluster had a small flux through the PPP (-73%), a small flux through both the TCA branch (-13 and -23%) and small production of biomass (-15%) but a high acetate synthesis and output ($+72$ and $+63\%$). Similarly, the flux of acetaldehyde output predicted by the model highlighted a cluster of three strains characterized by a very high production of acetaldehyde, of which Clib215_3B strain was a good example (Fig. 5l). This strain was mainly characterized by a high acetaldehyde output ($+94\%$), a high reductive branch of TCA ($+27\%$) and succinate output ($+25\%$), high glycerol output ($+15\%$) and a small acetate production and output (-61 and -55%). The other fluxes did not allow such a clear separation of

strains but illustrated small differences in similar global distributions.

Interestingly, these two particular clusters were overwhelmingly composed of strains having one ecological origin. The cluster characterized by a high production and output of acetate was composed of “Flor” strains, and the cluster with high acetaldehyde production was only composed of “Bread” strains. To better understand the effect of strain origin on flux distribution, we considered the mean fluxes by origin (Fig. 5b–i). The acetate synthesis and output fluxes (Fig. 5h, i) were approximately 50% higher for the “Flor” and “American Oak” (Oak) strains and approximately 50 and 25% lower for the Bread and Wine strains, respectively. This dichotomous behavior

explaining the bimodal distribution of these two fluxes also presented a significant effect of the ecological origin ($p < 0.001$ for both fluxes). Similarly, the very long tail in the flux distribution of acetaldehyde output (Acald_t) can be explained by the “Bread” strains that produce approximately 100 % more acetaldehyde than other strains (Fig. 5g, $p = 0.003$). Flux through the PPP (Fig. 5e, $p < 0.001$) and glycerol synthesis (Fig. 5c, $p < 0.001$) also presented significant effects of strain origin while having less variability. By contrast, fluxes with high variability and that well separated strains, such as the alpha-ketoglutarate output (Fig. 5f), presented no significant effect of strain origin. Thus, there was no link between the extent of flux distribution and its contribution to strain origin separation.

Thus, this analysis indicated interesting physiological differences between strains, some of which were related to the ecological origin. To experimentally confirm the higher production of acetaldehyde by the bread strains, we a posteriori measured the production of acetaldehyde for seventeen strains from various origins and compared the relative variations of production with flux prediction (Fig. 6). These experimental data confirmed our predictions, with the “Bread” strains producing $137.78 \pm 5.68 \text{ mg L}^{-1}$ of acetaldehyde on average, while the strains from other origins produced $59.88 \pm 35.51 \text{ mg L}^{-1}$ (p value < 0.001) at the fermentation time point of 11 g L^{-1} of CO_2 produced.

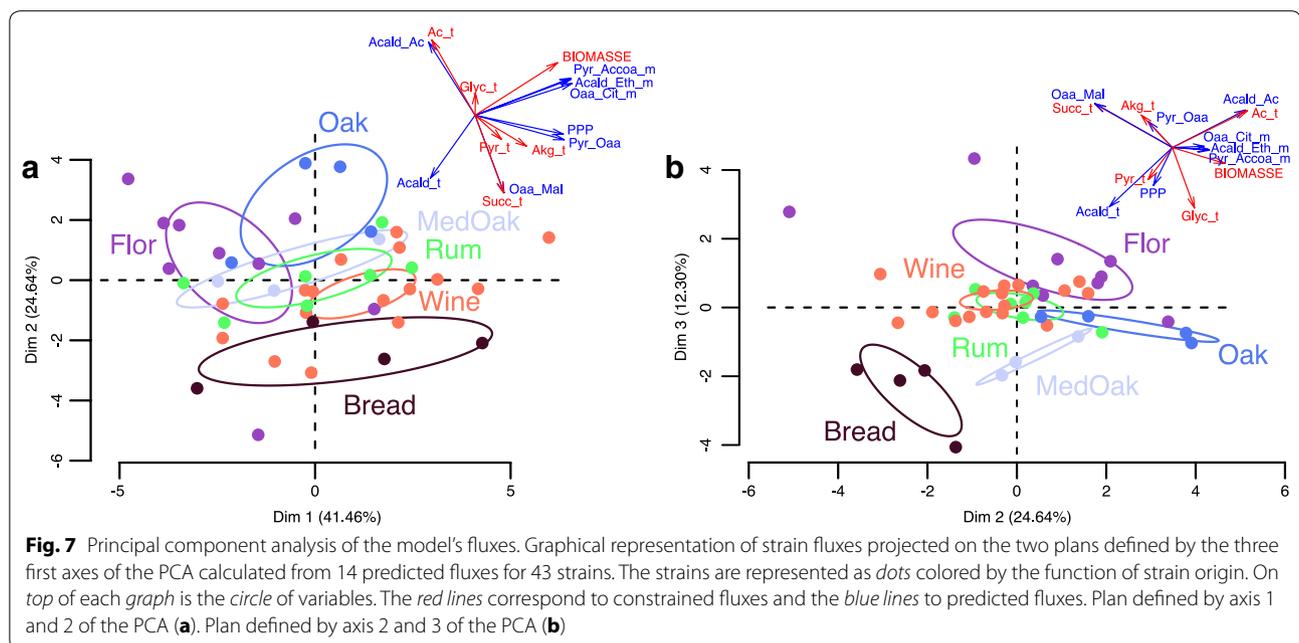


Moreover, a correlation was also found within groups of strains with similar ecological origins (Fig. 4) as well as for the proportion of the NADPH demand provided by the PPP or acetate synthesis. Indeed, the “Bread” and “Wine” strains mainly produced their NADPH by the PPP (approximately 84 and 72 %, respectively), while the six strains that predominantly produced NADPH by acetate synthesis were “Flor” strains, with only approximately 20 % of the NADPH demand produced by the PPP (Additional file 1: Figure S1).

Finally, to obtain an integrated vision of flux structuration, we performed a principal component analysis (PCA). For this, we selected the same subset of 19 fluxes, among which we excluded the fluxes of glycolysis and ethanol synthesis on the basis that they were stronger but also less variable fluxes, which would therefore give them too much importance in the PCA. A final subset of 14 fluxes was used to perform the PCA (Fig. 7). The first three axes of the PCA explained 41.46, 24.62 and 12.3 % of the variance. The PCA plan defined by the second and third axes was the one that better separated the strains according to their origins. The second axis significantly separated the “Bread” (+2.37) and the “Oak” (−2.4) strains, and the third axis significantly separated the “Flor” (+1.84), the “Wine” (+0.67), the “Med_oak” (−0.97) and the “Bread” (−1.95) strains. The “Bread” strains at the bottom left of this PCA plan were characterized by a high production of acetaldehyde and a small production of acetate. The oak strains (“Med_oak” and “Oak”) in the bottom right had high production of glycerol and small production of succinate. The “Flor” group at the top right had high production of acetate, a small flux through the PPP and small production of acetaldehyde. This group was almost symmetrically opposed to the “Bread” group. The two remaining groups, “Rum” and “Wine,” were more central and better separated by the plan determined by the two first axes of the PCA. Finally, it is interesting to highlight that the fluxes structuring the axis were in the same proportion predicted by the model and constrained by the experimental data.

Discussion

In this work, we used a constraint-based model of yeast fermentative central carbon metabolism to study the diversity of flux distribution among 43 strains of different origins. We used a whole set of experimental data (ethanol, glycerol, succinate, acetate, pyruvate, alpha-ketoglutarate and biomass production) to constrain the model and a FBA approach with minimization of the glucose input to predict the distribution of metabolic fluxes. This method allowed us to optimize the modeling process by using all the available biological information. We first considered the variability of the predictions to



determine the confidence of the estimates. Considering alternate optimal solutions led us to conclude that the *DynamoYeast* model was very well determined, with only small variations in the reductive branch of the TCA due to free mitochondrial transport of the involved metabolites (malate, fumarate and succinate). This very low level of variability between alternate optimal solutions for a given set of constraints was the main advantage of using a reduced model. Indeed, the same constraints used with a genome-scale model (6th version of the consensus model, [41]) led to predicted flux distribution predictions with many alternative solutions, some of which were biologically irrelevant (data not shown).

The main objective of this study was to characterize the variability of flux distributions between *S. cerevisiae* strains from different origins. We found that this variability was strongly pathway-dependent. The glycolysis and ethanol synthesis pathways, despite being the stronger fluxes, showed almost no variability between strains. In contrast, flux through the PPP was the most variable, with a coefficient of variation more than two times higher than that of other pathways. This high variability of the PPP is in accordance with a previous study stressing high variability of the specific activity of the first enzyme of the PPP, glucose-6-phosphate dehydrogenase, in eleven *S. cerevisiae* strains [42]. This, in addition to the finding that the PPP was one of the most variable fluxes in different environments [13], suggests high flexibility of this pathway depending on environmental and genetic factors.

Our study also highlighted several correlations between metabolic pathways. The PPP produces around 2/3 of

the NADPH demand and displays a strong trade-off with the cytoplasmic synthesis of acetate from acetaldehyde (Acald_Ac in our model), the other main reaction generating NADPH. An indication of a link between these two pathways was found in previous studies. For example, in a study comparing the flux distributions of *S. cerevisiae* during respiro-fermentative growth in different conditions of pH and NaCl concentration, Heyland et al. [43] found an inverse variation between the fluxes through acetate production and PPP, unfortunately with too few points to test for a significant correlation. Predicted fluxes between an evolved strain of *S. cerevisiae* and its ancestor showed a similar trade-off: an increased flux through the PPP and a decreased production of acetate in the evolved strain [44].

Interestingly, among the intra-species correlations that we identified in this study, some have also previously been found when different yeast species were compared. The positive correlation between PPP and biomass fluxes (which we linked to biomass precursor synthesis) was also found in a comparative ^{13}C -flux analysis of seven yeast species [26] and of fourteen other hemiascomycetous yeasts [17]. Between these fourteen hemiascomycetous, the proportion of NADPH demand produced by the PPP varied between 60 % for *S. cerevisiae* and 90 % for *P. angusta* [17]. Similarly, in our work, the mean percentage of NADPH produced by the PPP was 59 % (Additional file 1: Figure S1). A higher level of flux through the PPP was found for *S. cerevisiae* in the Blank study compared to this work (10 versus 2 %); this discrepancy between fluxes predicted by ^{13}C -MFA or FBA is common [12]. Another correlation found in our work as in other studies

was the negative correlation between glycolysis and the TCA fluxes, which have been associated with a down regulation of glycolytic genes [43].

Another issue addressed in this study is the contribution of strain origin to intra-species metabolic diversity. For the variable fluxes, the flux distribution was divergent in broadness and could also be mono-, bi- or multimodal, indicating dichotomous behavior between strains. We could explain these different patterns of distribution by strain origin peculiarities. For example, the long tail of the acetaldehyde output distribution can be explained by the four “Bread” strains that produce twice as much acetaldehyde (Fig. 5g) and the bimodal distribution of the production and output of acetate by the contrasted behavior of the “Flor” and “Bread” strains. Further, using the predicted fluxes rather than only the experimental data helps to distinguish the strains according to their origins (Additional file 1: Figure S2). Indeed, among the five fluxes (G6p_6pgl, Acald_t, Akg_t, Acald_Ac, Ac_t) that best distinguished strains from each other (especially the “Bread” and “Flor” strains), two were only accessible by the model (G6p_6pgl, Acald_t), which highlights the potential of the flux analysis approach. Interestingly, some fluxes, such as flux through the PPP, were by themselves able to separate strains by origin.

Such knowledge on the most flexible fluxes and strain-dependent flux variability could be very useful for metabolic engineering strategies aimed at rerouting metabolic fluxes. Numerous studies [44–54] have attempted to modify yeast flux distributions using metabolic or evolutionary engineering approaches or hybridization to exploit natural diversity for various biotechnological applications. Our study shows almost no diversity in the flux distributions of glycolysis or ethanol synthesis, suggesting strong constraints on these fluxes, either evolutionary or metabolic. By contrast, the fluxes through glycerol synthesis [54–57] or the PPP [42, 44] were more flexible, which makes them more interesting targets to redirect metabolic fluxes. In addition, the availability of strain-specific maps of metabolic flux distribution will provide a framework for the selection of the most relevant strains for metabolic engineering strategies.

Conclusion

Overall, this work highlights the potential of flux analysis to identify the most variable and robust nodes of central carbon metabolism within a species and to provide information on the metabolic or evolutionary constraints that shape flux distribution. This knowledge will help to identify relevant targets and yeast strains for metabolic engineering. In addition, the availability of whole genome sequences for the strains used in this study offers a framework to decipher the links between flux distribution and

strain genotypes. In particular, the finding of a strain origin effect on the distribution of various fluxes opens the way for flux quantitative trait loci (QTL) detection (fQTL) to elucidate the genetic basis of flux distribution.

Methods

Strains and culture conditions

The 43 *S. cerevisiae* strains of six different ecological origins (4 “Bread,” 7 “Rum,” 16 “Wine,” 9 “Flor,” 3 “Medoak” and 4 “Oak”) used in this study are listed in Additional file 2: Table S1. These strains were conserved at -80°C and transferred to YPD agar plates 48 h before fermentation. Initial cultures (12 h, in 50 ml YPD medium, 28°C) were used to inoculate fermentation at a density of 106 cells/ml. Fermentation was carried out in synthetic MS medium, which contained 240 g/L sugars (equimolar mixture of glucose and fructose), 6 g/L malic acid, 6 g/L citric acid and 200 mg/L nitrogen in the form of amino acids (148 mg N/L) and NH_4Cl (52 mg N/L), at pH 3.5 (5). Ergosterol (1.875 mg/L), oleic acid (0.625 mg/L) and Tween 80 (0.05 g/L) were provided as anaerobic growth factors. Fermentation took place in 1.1-liter fermentors equipped with fermentation locks to maintain anaerobiosis, at 28°C , with continuous magnetic stirring (500 rpm). CO_2 release was followed by automatic measurements of fermentor weight loss every 20 min. The amount of CO_2 released allowed us to monitor the progress of the fermentation. Samples were harvested for further analysis when the released CO_2 reached approximately 11 g. The dry weight of the yeast was measured by filtering 50 mL of culture through a 0.45-mm-pore Millipore nitrocellulose filter, which was washed twice with 50 mL distilled water and dried for 24 h at 105°C . Metabolites in the supernatant (acetate, succinate, glycerol, alpha-ketoglutarate, pyruvate and ethanol) were analyzed by high-pressure liquid chromatography [36]. Acetaldehyde production was determined with an enzymatic UV method [58].

Fermentation was carried out in duplicate spread over various fermentation blocks. Data (six metabolites, biomass) were first normalized by the released CO_2 . We then used a linear mixed model (Rstudio, nlme package) to correct measures for “block” effects, and the average values between the two replicates were calculated. From these normalized and corrected data, we recalculated the biomass and metabolite concentrations corresponding to 11 g/L of CO_2 .

Model

Metabolite concentrations (in mmol ml^{-1}) and dry weight (g L^{-1}) were used to constrain DynamoYeast, a previously developed dedicated constraint-based model of yeast fermentative central carbon metabolism [9]. This model is composed of three compartments: the

cytoplasm, mitochondria and extracellular medium, and includes 61 metabolites (Additional file 2: Table S2 for full name and abbreviations) and 68 reactions (Additional file 2: Table S3). For each of the 43 strains, we used the corrected metabolite concentrations to constrain the corresponding output flux of the model and the measured dry weight to constrain the flux of biomass (Additional file 2: Table S1). We used the experimental measures $+2.5$ and -2.5 % at the upper and lower flux boundaries, respectively. Then, we performed a flux balance analysis (FBA) minimizing the flux of glucose input (Glc_t) to obtain the flux distribution through the metabolic network [9]. In contrast to other standard constraint-based methods that compute flux distribution based on the derivation of mass data, here we directly computed mass distribution, as in Celton et al. [9].

We considered that all sugars were glucose (instead of glucose and fructose) for the modeling approach, as this assumption did not impact the flux predictions. For all strains, we used the biomass composition previously determined for the EC1118 strain [9] and set the cytosolic isocitrate dehydrogenase reaction (*IDP2*, YLR174W), the mitochondrial glutamate dehydrogenase reaction (*GDH2*, YDL215C) and the futile cycle around glycerol [9] to 0.

All predictions were performed with Matlab R2010b. The flux balance analysis (FBA) was performed with the “optimizeCbModel” function from the cobra toolbox [59] and the GLPK solver. The evaluation of the number of alternative solutions was done with the “enumerateOptimalSolution” algorithm [40] from a model where all input and output fluxes had been constrained by their exact predicted value from the FBA optimization.

Statistical analysis

For each strain, we obtained a prediction of the flux distribution through the metabolic network. However, the predicted glucose uptake was different for each strain. To compare flux distributions between strains, we normalized each flux to the specific glucose uptake in the corresponding strain and expressed it as a percentage. We calculated the coefficient of variation for each flux: the standard deviation divided by the mean flux of all strains.

On a subset of 19 fluxes, we calculated the relative deviation from the average $\left(\frac{Flux_i - Flux_{mean}}{Flux_{mean}}\right)$, which gave an idea of how far a given strain was from the average distribution. To analyze the effect of strain origin on selected relative deviations, we used a linear model with a fixed effect of origins and ANOVA.

Principal component analysis of flux values was performed with fourteen fluxes that were representative of the entire model's network, with the exception of the glycolysis and ethanol synthesis fluxes. All analysis and

graphical representations were performed with RStudio [60] and with the following packages: “FactoMineR,” “corrplot,” “gplots” and “XML.” The graphical representations were later modified with Inkscape (<http://www.inkscape.org>) for visual ameliorations.

Additional files

Additional file 1. Table S1: Strains' origin and experimental data. **Table S2:** Model's metabolites abbreviations. **Table S3:** Model's reactions.

Additional file 2. Fig. S1: Proportion of the NADPH demand produced by the PPP and the acetate synthesis for each strain. Proportion of the NADPH demand produced by the PPP (red) and the acetate synthesis (green) represented for each strain as a vertical bar. The horizontal black line represents the average proportion of the NADPH demand produced by the PPP. **Fig. S2:** Comparison of the second and third axis of PCA made from the fluxes or the biological data. Graphical representation of strain fluxes projected on the two plans defined by the second and third axes of the PCA. The strains are represented as dots colored in function of strain origin. On top of each graph is the circle of variables. PCA calculated from 14 predicted fluxes (a). PCA calculated from 7 experimental data (b).

Abbreviations

CBM: constraint-based model; MFA: metabolic flux analysis; FBA: flux balance analysis; PPP: pentose phosphate pathway; CCM: central carbon metabolism; E4P: erythrose-4-phosphate; R5p: ribose-5-phosphate; AKG: alpha-ketoglutarate.

Authors' contributions

PB and CC produced the experimental data. TN performed the statistical analysis. TN performed the modeling and flux predictions. TN, CC and SD discussed the data and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank JL Legras, S Casaregola, G Loiseau and JP Sampaio for providing yeast strains. The authors acknowledge JL Legras for helpful comments and discussions and Isabelle Sanchez for help in statistical analysis.

Competing interests

The authors declare that they have no competing interests.

Received: 10 February 2016 Accepted: 23 March 2016

Published online: 05 April 2016

References

- Nielsen J. It is all about metabolic fluxes. *J Bacteriol.* 2003;185:7031–5.
- Wiechert W. 13C metabolic flux analysis. *Metab Eng.* 2001;3:195–206.
- Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nat Biotechnol.* 2010;28:245–8.
- Schuetz R, Kuepfer L, Sauer U. Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol Syst Biol.* 2007;3:119.
- Burgard AP, Maranas CD. Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnol Bioeng.* 2003;82:670–7.
- García Sánchez CE, Torres Sáez RG, CE. Comparison and analysis of objective functions in flux balance analysis. *Biotechnol Prog.* 2014;30:985–91.
- Ramakrishna R, Edwards JS, McCulloch A, Palsson BO. Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. *Am J Physiol Regul Integr Comp Physiol.* 2001;280:R695–704.

8. Segrè D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci*. 2002;99:15112–7.
9. Celton M, Goelzer A, Camarasa C, Fromion V, Dequin S. A constraint-based model analysis of the metabolic consequences of increased NADPH oxidation in *Saccharomyces cerevisiae*. *Metab Eng*. 2012;14:366–79.
10. Kauffman KJ, Prakash P, Edwards JS. Advances in flux balance analysis. *Curr Opin Biotechnol*. 2003;14:491–6.
11. Feist AM, Palsson BO. The biomass objective function. *Curr Opin Microbiol*. 2010;13:344–9 (**Ecology and industrial microbiology • special section: systems biology**).
12. Celton M, Sanchez I, Goelzer A, Fromion V, Camarasa C, Dequin S. A comparative transcriptomic, fluxomic and metabolomic analysis of the response of *Saccharomyces cerevisiae* to increases in NADPH oxidation. *BMC Genom*. 2012;13:317.
13. Quirós M, Martínez-Moreno R, Albiol J, Morales P, Vázquez-Lima F, Barreiro-Vázquez A, Ferrer P, Gonzalez R. Metabolic flux analysis during the exponential growth phase of *Saccharomyces cerevisiae* in wine fermentations. *PLoS ONE*. 2013;8:e71909.
14. Kerkhoven EJ, Lahtvee P-J, Nielsen J. Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Res*. 2014;15:1–13.
15. Oberhardt MA, Palsson BØ, Papin JA. Applications of genome-scale metabolic reconstructions. *Mol Syst Biol*. 2009;5:320.
16. Österlund T, Nookaew I, Nielsen J. Fifteen years of large scale metabolic modeling of yeast: developments and impacts. *Biotechnol Adv*. 2012;30:979–88.
17. Blank LM, Lehmbeck F, Sauer U. Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts. *FEMS Yeast Res*. 2005;5:545–58.
18. Velagapudi VR, Wittmann C, Schneider K, Heinzle E. Metabolic flux screening of *Saccharomyces cerevisiae* single knockout strains on glucose and galactose supports elucidation of gene function. *J Biotechnol*. 2007;132:395–404 (**In Memoriam W.-D. Deckwer: merging process engineering and systems biology**).
19. Patil KR, Nielsen J. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc Natl Acad Sci USA*. 2005;102:2685–9.
20. Agren R, Otero JM, Nielsen J. Genome-scale modeling enables metabolic engineering of *Saccharomyces cerevisiae* for succinic acid production. *J Ind Microbiol Biotechnol*. 2013;40:735–47.
21. Bro C, Regenberg B, Förster J, Nielsen J. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab Eng*. 2006;8:102–11.
22. Otero JM, Cimini D, Patil KR, Poulsen SG, Olsson L, Nielsen J. Industrial systems biology of *Saccharomyces cerevisiae* enables novel succinic acid cell factory. *PLoS ONE*. 2013;8:e54144.
23. Bundy JG, Papp B, Harmston R, Browne RA, Clayson EM, Burton N, Reece RJ, Oliver SG, Brindle KM. Evaluation of predicted network modules in yeast metabolism using NMR-based metabolite profiling. *Genome Res*. 2007;17:510–9.
24. Cannizzaro C, Christensen B, Nielsen J, von Stockar U. Metabolic network analysis on *Phaffia rhodozyma* yeast using ¹³C-labeled glucose and gas chromatography-mass spectrometry. *Metab Eng*. 2004;6:340–51.
25. Fiaux J, Çakar ZP, Sonderegger M, Wüthrich K, Szyperski T, Sauer U. Metabolic-flux profiling of the yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*. *Eukaryot Cell*. 2003;2:170–80.
26. Christen S, Sauer U. Intracellular characterization of aerobic glucose metabolism in seven yeast species by ¹³C flux analysis and metabolomics. *FEMS Yeast Res*. 2011;11:263–72.
27. Fay JC, Benavides JA. Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. *PLoS Genet*. 2005;1:e5.
28. Legras J-L, Ruh O, Merdinoglu D, Karst F. Selection of hypervariable microsatellite loci for the characterization of *Saccharomyces cerevisiae* strains. *Int J Food Microbiol*. 2005;102:73–83.
29. Legras JL, Merdinoglu D, Cornuet J, Karst F. Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol*. 2007;16:2091–102.
30. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, Davey RP, Roberts IN, Burt A, Koufopanou V. Population genomics of domestic and wild yeasts. *Nature*. 2009;458:337–41.
31. Schacherer J, Shapiro JA, Ruderfer DM, Kruglyak L. Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature*. 2009;458:342–5.
32. Cromie GA, Hyma KE, Ludlow CL, Garmendiatordes C, Gilbert TL, May P, Huang AA, Dudley AM, Fay JC. Genomic sequence diversity and population structure of *Saccharomyces cerevisiae* assessed by RAD-seq. *G3 Genom Genet*. 2013;3:2163–71.
33. Warringer J, Zörgö E, Cubillos FA, Zia A, Gjuvsland A, Simpson JT, Forsmark A, Durbin R, Omholt SW, Louis EJ, Liti G, Moses A, Blomberg A. Trait variation in yeast is defined by population history. *PLoS Genet*. 2011;7:e1002111.
34. Strope PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Magwene PM, Dietrich FS, McCusker JH. The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Res*. 2015;25:762–74.
35. Spor A, Nidelet T, Simon J, Bourgeois A, de Vienne D, Sicard D. Niche-driven evolution of metabolic and life-history strategies in natural and domesticated populations of *Saccharomyces cerevisiae*. *BMC Evol Biol*. 2009;9:296.
36. Camarasa C, Sanchez I, Briál P, Bigey F, Dequin S. Phenotypic landscape of *Saccharomyces cerevisiae* during wine fermentation: evidence for origin-dependent metabolic traits. *PLoS ONE*. 2011;6:e25147.
37. Barbosa C, Lage P, Vilela A, Mendes-Faia A, Mendes-Ferreira A. Phenotypic and metabolic traits of commercial *Saccharomyces cerevisiae* yeasts. *AMB Express*. 2014;4:39.
38. Mukherjee V, Steensels J, Lievens B, de Voorde IV, Verplaetse A, Aerts G, Willems KA, Thevelein JM, Verstrepen KJ, Ruyters S. Phenotypic evaluation of natural and industrial *Saccharomyces* yeasts for different traits desirable in industrial bioethanol production. *Appl Microbiol Biotechnol*. 2014;98:9483–98.
39. Marsit S, Dequin S. Diversity and adaptive evolution of *Saccharomyces* wine yeast: a review. *FEMS Yeast Res*. 2015;15:67.
40. Reed JL, Palsson BØ. Genome-scale in silico models of *E. coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. *Genome Res*. 2004;14:1797–805.
41. Heavner BD, Smallbone K, Price ND, Walker LP. Version 6 of the consensus yeast metabolic network refines biochemical coverage and improves model performance. *Database*. 2013;2013:bat059.
42. Heux S, Cadiere A, Dequin S. Glucose utilization of strains lacking PGI1 and expressing a transhydrogenase suggests differences in the pentose phosphate capacity among *Saccharomyces cerevisiae* strains. *FEMS Yeast Res*. 2008;8:217–24.
43. Heyland J, Fu J, Blank LM. Correlation between TCA cycle flux and glucose uptake rate during respiro-fermentative growth of *Saccharomyces cerevisiae*. *Microbiology*. 2009;155:3827–37.
44. Cadière A, Ortiz-Julien A, Camarasa C, Dequin S. Evolutionary engineered *Saccharomyces cerevisiae* wine yeast strains with increased in vivo flux through the pentose phosphate pathway. *Metab Eng*. 2011;13:263–71.
45. Dequin S. The potential of genetic engineering for improving brewing, wine-making and baking yeasts. *Appl Microbiol Biotechnol*. 2001;56:577–88.
46. Donalies UEB, Nguyen HTT, Stahl U, Nevoigt E. Improvement of *Saccharomyces* Yeast Strains Used in Brewing, Wine Making and Baking. In: Stahl U, Donalies UEB, Nevoigt E, editors. *Food Biotechnology*. Springer: Berlin; 2008. p. 67–98 (**Advances in biochemical engineering/biotechnology, vol. 111**).
47. Husnik JJ, Volschenk H, Bauer J, Colavizza D, Luo Z, van Vuuren HJJ. Metabolic engineering of malolactic wine yeast. *Metab Eng*. 2006;8:315–23.
48. Schuller D, Casal M. The use of genetically modified *Saccharomyces cerevisiae* strains in the wine industry. *Appl Microbiol Biotechnol*. 2005;68:292–304.
49. Ehsani M, Fernández MR, Biosca JA, Julien A, Dequin S. Engineering of 2,3-Butanediol dehydrogenase to reduce acetoin formation by glycerol-overproducing, low-alcohol *Saccharomyces cerevisiae*. *Appl Environ Microbiol*. 2009;75:3196–205.
50. Kutyna DR, Varela C, Stanley GA, Borneman AR, Henschke PA, Chambers PJ. Adaptive evolution of *Saccharomyces cerevisiae* to generate strains with enhanced glycerol production. *Appl Microbiol Biotechnol*. 2011;93:1175–84.
51. Steensels J, Snoek T, Meersman E, Nicolino MP, Voordeckers K, Verstrepen KJ. Improving industrial yeast strains: exploiting natural and artificial diversity. *FEMS Microbiol Rev*. 2014;38:947–95.

52. Varela C, Kutyna DR, Solomon MR, Black CA, Borneman A, Henschke PA, Pretorius IS, Chambers PJ. Evaluation of gene modification strategies for the development of low-alcohol-wine yeasts. *Appl Environ Microbiol.* 2012;78:6068–77.
53. Michnick S, Roustan J-L, Remize F, Barre P, Dequin S. Modulation of glycerol and ethanol yields during alcoholic fermentation in *Saccharomyces cerevisiae* Strains overexpressed or disrupted for GPD1 encoding glycerol 3-phosphate dehydrogenase. *Yeast.* 1997;13:783–93.
54. Tilloy V, Ortiz-Julien A, Dequin S. Reducing ethanol and improving glycerol yield by adaptive evolution of *Saccharomyces cerevisiae* wine yeast under hyperosmotic conditions. *Appl Environ Microbiol.* 2014. doi:10.1128/AEM.03710-13.
55. Eglinton JM, Heinrich AJ, Pollnitz AP, Langridge P, Henschke PA, de Barros Lopes M. Decreasing acetic acid accumulation by a glycerol overproducing strain of *Saccharomyces cerevisiae* by deleting the ALD6 aldehyde dehydrogenase gene. *Yeast.* 2002;19:295–301.
56. Remize F, Andrieu E, Dequin S. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae*: role of the cytosolic Mg²⁺ and mitochondrial K⁺ acetaldehyde dehydrogenases Ald6p and Ald4p in acetate formation during alcoholic fermentation. *Appl Environ Microbiol.* 2000;66:3151–9.
57. Nevoigt E, Stahl U. Reduced pyruvate decarboxylase and increased glycerol-3-phosphate dehydrogenase [NAD⁺] levels enhance glycerol production in *Saccharomyces cerevisiae*. *Yeast.* 1996;12:1331–7.
58. Noble J, Sanchez I, Blondin B. Identification of new *Saccharomyces cerevisiae* variants of the MET2 and SKP2 genes controlling the sulfur assimilation pathway and the production of undesirable sulfur compounds during alcoholic fermentation. *Microb Cell Fact.* 2015;14:1–16.
59. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, Kang J, Hyduke DR, Palsson BØ. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat Protoc.* 2011;6:1290–307.
60. Racine JS, RStudio. A Platform-Independent IDE for R and Sweave. *J Appl Econom.* 2012;27:167–72.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

