

OPEN

# QTL mapping of modelled metabolic fluxes reveals gene variants impacting yeast central carbon metabolism

Matthias Eder<sup>1</sup>, Thibault Nidelet<sup>1</sup>, Isabelle Sanchez<sup>1,2</sup>, Carole Camarasa<sup>1</sup>, Jean-Luc Legras<sup>1</sup> & Sylvie Dequin<sup>1\*</sup>

The yeast *Saccharomyces cerevisiae* is an attractive industrial microorganism for the production of foods and beverages as well as for various bulk and fine chemicals, such as biofuels or fragrances. Building blocks for these biosyntheses are intermediates of yeast central carbon metabolism (CCM), whose intracellular availability depends on balanced single reactions that form metabolic fluxes. Therefore, efficient product biosynthesis is influenced by the distribution of these fluxes. We recently demonstrated great variations in CCM fluxes between yeast strains of different origins. However, we have limited understanding of flux modulation and the genetic basis of flux variations. In this study, we investigated the potential of quantitative trait locus (QTL) mapping to elucidate genetic variations responsible for differences in metabolic flux distributions (fQTL). Intracellular metabolic fluxes were estimated by constraint-based modelling and used as quantitative phenotypes, and differences in fluxes were linked to genomic variations. Using this approach, we detected four fQTLs that influence metabolic pathways. The molecular dissection of these QTLs revealed two allelic gene variants, *PDB1* and *VID30*, contributing to flux distribution. The elucidation of genetic determinants influencing metabolic fluxes, as reported here for the first time, creates new opportunities for the development of strains with optimized metabolite profiles for various applications.

The yeast *Saccharomyces cerevisiae* has been used for millennia for the production of various fermented foods and beverages<sup>1</sup>. In modern times, yeast has become popular in new applications, ranging from the biosynthesis of ethanol (for biofuels) and other raw materials (for chemical syntheses) to the production of fine chemicals used as fuel additives, flavours and fragrances, or medical components<sup>2</sup>. Advances in metabolic engineering are constantly expanding the range of yeast's applications.

Many phenotypic traits of *S. cerevisiae* relevant to industrial processes are dependent on the functional and regulatory properties of central carbon metabolism (CCM)<sup>3</sup>. The yeast metabolic network that involves a large number of intracellular reactions is highly conserved and has evolved to be organized as a bowtie structure, meaning that all carbon sources are converted to 12 different precursor metabolites, which are then used by the cell for the biosynthesis of macromolecules that compose cellular biomass<sup>4,5</sup>. In addition, these precursors form the basis for the synthesis of extracellular metabolites, e.g., succinic acid, which is industrially produced by yeast fermentation as a building block for polymer production<sup>6,7</sup>. Furthermore, they serve as starting points for the heterologous production pathways of renewably produced fine chemicals<sup>8,9</sup>. The tight reduction of cellular carbon core metabolism to a small number of important metabolites results in a high carbon flux through these compounds<sup>10</sup>, controlled by a complex regulation on genetic (transcription, translation, protein modifications and protein-protein interactions) and metabolic levels. Therefore, understanding how metabolic flux distribution is controlled is a key requirement for increasing product biosynthesis by metabolic engineering<sup>11</sup>.

While metabolite turnover rates are difficult to determine experimentally, they can be estimated by modelling<sup>12,13</sup>. Commonly used constraint-based models (CBM) formulate metabolic networks as a stoichiometric matrix to predict intracellular fluxes through the application of experimental constraints on input and output fluxes. Depending on the network size and number of constraints, this approach, which is termed metabolic

<sup>1</sup>SPO, Univ Montpellier, INRAE, Montpellier SupAgro, Montpellier, France. <sup>2</sup>MISTEA, INRAE, Montpellier SupAgro, Montpellier, France. \*email: [sylvie.dequin@inrae.fr](mailto:sylvie.dequin@inrae.fr)

flux analysis (MFA), can be sufficient for estimating fluxes. However, in most cases, adding constraints on input and output data is not sufficient to estimate all fluxes of a network. One way to address this insufficiency is the  $^{13}\text{C}$ -MFA approach. It tracks  $^{13}\text{C}$  from labelled substrates across cellular metabolites with the aim of generating information to constrain and estimate intracellular fluxes. Another way is to apply the assumption that cellular functions of biochemical networks in a steady state are limited by physico-chemical constraints<sup>14</sup>. In this case, the flux balance analysis (FBA) approach chooses the best fitting solution through linear optimization out of a narrowed solution frame defined by the stoichiometric matrix of the CBM<sup>15</sup>. The outcome of this flux prediction depends on an applied objective function (maximization of ATP production, minimization of metabolic adjustment, or in most cases, maximization of biomass production)<sup>16</sup>.

To study how metabolic fluxes are modulated by genetic or environmental determinants, we previously used a combined  $^{13}\text{C}$ -MFA/FBA approach to estimate the intracellular fluxes of *S. cerevisiae* CCM in conditions of modified intracellular redox balance<sup>17,18</sup>. Another example of the application of FBA is the study of Quirós *et al.* (2013), which used a model developed by Vargas *et al.* (2011) to evaluate changes in yeast metabolism in high sugar must<sup>19,20</sup>. In both studies, glycolytic fluxes showed the least variation, whereas the fluxes of the pentose phosphate pathway (PPP) were highly variable.  $^{13}\text{C}$ -MFA, on the other hand, was used to study network robustness<sup>21</sup> or the effects of deletion mutants<sup>22</sup>. The latter study demonstrated interesting links between networks, e.g., a positive correlation between the PPP and biomass yield.

In recent years, studies have generated vast amounts of information about the genotypic and phenotypic diversity of *S. cerevisiae* by comparison of growth parameters in different media<sup>23–28</sup>. Furthermore, several studies extended the characterization of diversity to a greater number of phenotypic traits, including life history traits and metabolic traits, showing that origin has a broad phenotypic impact and that part of these phenotypic differences can be explained by adaptation to the ecological constraints imposed by origin<sup>29–31</sup>.

Recently, we assessed the diversity of flux distributions between *S. cerevisiae* strains from different origins<sup>32</sup>, using the constraint-based model developed by Celton *et al.* (2012) to estimate CCM flux distributions between 43 strains grown under wine fermentation conditions<sup>18</sup>. The study showed a contrasting image regarding flux variability with quasi-constancy of glycolysis and ethanol synthesis on the one hand, but large variations in other fluxes, such as the PPP and acetaldehyde production, on the other hand. In addition, the fluxes' multimodal distributions related to ecological origin revealed an association between genetic origin and flux phenotype<sup>32</sup>.

Results from flux analysis have been used for strain improvement (increased ethanol yield, optimization of metabolites production, ...) by metabolic engineering based on flux predictions<sup>33–38</sup>. Therefore, more knowledge about the impact of genomic variation on metabolic flux distributions can potentially assist with the selection or improvement of strains for diverse applications, in the food and beverage industries as in the field of biotechnology.

Quantitative trait locus (QTL) mapping, which has been applied in numerous existing studies, has become an important approach to more deeply understand the genomic complexity of *S. cerevisiae* and decipher the impact of genomic variation on yeast complex traits<sup>39</sup>. This application includes investigations of genetic determinants influencing the formation of industrially relevant traits, which have led to the discovery of allelic variants accounting for variations in these traits<sup>40–48</sup>. In all these studies, the assessed traits were straightforward to quantify. However, difficulties remain in detecting QTLs for traits with small variations or those that are more challenging to determine, such as intracellular metabolic fluxes.

The possibility of using QTL mapping to decipher genomic variations impacting metabolic profiles (rather than those affecting single metabolites) would create ways to understand the mechanisms behind metabolic flux distributions and to engineer strains with superior metabolic properties for various applications. To achieve this aim, we phenotyped 130 meiotic F2-segregants from a cross of two *S. cerevisiae* yeast strains for their production of extracellular main metabolites during the exponential phase. We modelled the intracellular fluxes of yeast CCM by applying these experimentally determined metabolite concentrations to a constraint-based model<sup>18</sup>. Subsequently, we used these estimated fluxes as phenotypic data to perform QTL mapping on metabolic flux distributions. With this approach, we were able to detect four fQTLs that influenced various metabolic fluxes. By performing reciprocal hemizygosity analysis (RHA), we confirmed the robustness of the method by validating the role of two genes, *PDB1* and *VID30*, within two fQTLs. The allelic variants of these genes show different effects on the fluxes of glycolysis, ethanol synthesis, glycerol synthesis, the tricarboxylic acid (TCA) cycle and the excretion of TCA cycle metabolites.

## Results

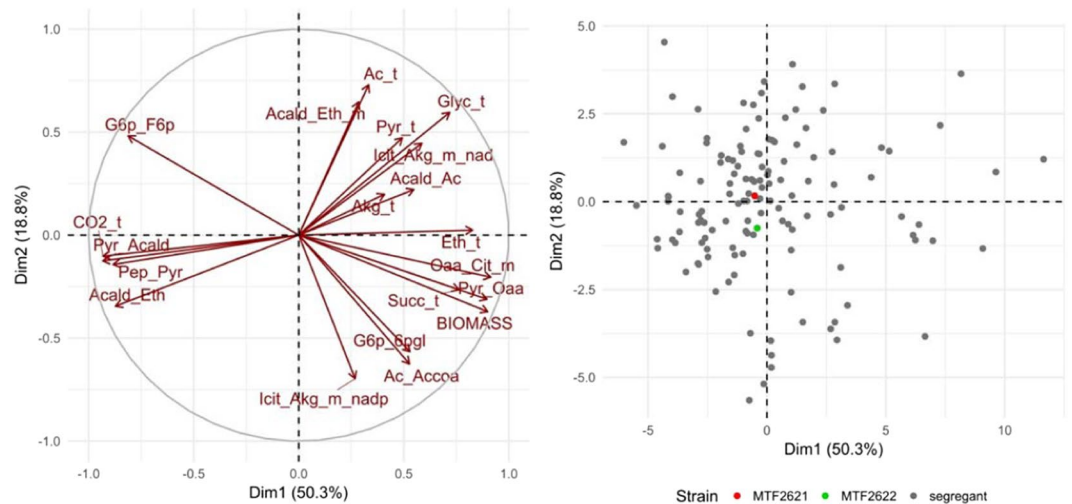
**Phenotyping of strains.** Fluxes of the CCM were predicted for all strains, using extracellular metabolite concentrations that were experimentally determined during the exponential growth stage (when cells are in a quasi-steady state). Because of the structure of the network, some fluxes are directly correlated. To assess for flux correlations, we analysed the relationships between steady-state reaction fluxes as described by Poolman *et al.* 2007<sup>49</sup> in order to produce a correlation matrix (Supplementary Fig. S1). Based on the observed strong linkage between single reactions within main metabolic pathways, representative reactions were selected for these pathways to facilitate the following analyses (Table 1).

For the production of extracellular main metabolites, the corresponding estimated excretion fluxes were chosen as representative. For main metabolic pathways, either the first flux (PPP, upper glycolysis, ethanol synthesis, TCA reductive, TCA oxidative) or the last flux of the pathway (lower glycolysis) were chosen. In the case of parallel fluxes leading to the same metabolite (ethanol synthesis, AKG synthesis), both fluxes were chosen for evaluation.

Principal component analysis (PCA) of selected flux reactions was performed to assess flux correlations and to evaluate the variation between parent and segregant strains (Fig. 1).

Flux abbreviation*	Pathway	Reaction
G6p_F6p	Upper glycolysis	$g6p[c] \rightleftharpoons f6p[c]$
Pep_Pyr	Lower glycolysis	$pep[c] + adp[c] \rightarrow pyr[c] + atp[c]$
G6p_6pgl	PPP	$g6p[c] + nadp[c] \rightleftharpoons 6pgl[c] + nadph[c]$
Pyr_Acald	Ethanol synthesis	$pyr[c] \rightarrow acald[c] + CO_2[c]$
Acald_Eth	Ethanol synthesis	$acald[c] + nadh[c] \rightarrow etoh[c] + nad[c]$
Acald_Ac	Acetate metabolism	$acald[c] + nadp[c] \rightarrow ac[c] + nadph[c]$
Ac_Acoa	Ac-CoA metabolism	$ac[c] + 2 atp[c] \rightarrow accoa[c] + 2 adp[c]$
Pyr_Oaa	TCA reductive branch	$pyr[c] + atp[c] + CO_2[c] \rightarrow oaa[c] + adp[c]$
Acald_Eth_m	Ethanol synthesis	$acald[m] + nadh[m] \rightleftharpoons etoh[m] + nad[m]$
Oaa_Cit_m	TCA oxidative branch	$accoa[m] + oaa[m] \rightarrow cit[m]$
Icit_Akg_m_nad	TCA oxidative branch	$icit[m] + nad[m] \rightarrow akg[m] + CO_2[m] + nadh[m]$
Icit_Akg_m_nadp	TCA oxidative branch	$icit[m] + nadp[m] \rightarrow akg[m] + CO_2[m] + nadph[m]$
Eth_t	Ethanol excretion	$etoh[c] \rightarrow$
Ac_t	Acetate excretion	$ac[c] \rightarrow$
Pyr_t	Pyruvate excretion	$pyr[c] \rightarrow$
Akg_t	AKG excretion	$akg[c] \rightarrow$
Succ_t	Succinate synthesis	$succ[c] \rightarrow$
Glyc_t	Glycerol synthesis	$glyc[c] \rightarrow$
CO <sub>2</sub> _t	CO <sub>2</sub> synthesis	$CO_2[c] \rightleftharpoons$
BIOMASS	Biomass formation	$3.96 g6p[c] + 0.258 r5p[c] + 0.129 e4p[c] + 0.116 g3p[c] + 0.303 3pgl[c] + 0.232 pep[c] + 0.775 oaa[c] + 1.084 pyr[m] + 0 pyr[c] + 0.176 accoa[m] + 0.252 accoa[c] + 0.106 akg[m] + 0.366 akg[c] + 0 CO_2[c] + 0.136 glu[c] + 115 atp[c] + 0.106 atp[m] + 1.499 nad[c] + 0.176 nad[m] + 0.602 nadph[m] + 5.35 nadph[c] \rightarrow 115 adp[c] + 0.106 adp[m] + 1.499 nadh[c] + 0.176 nadh[m] + 0.602 nadp[m] + 5.35 nadp[c]$

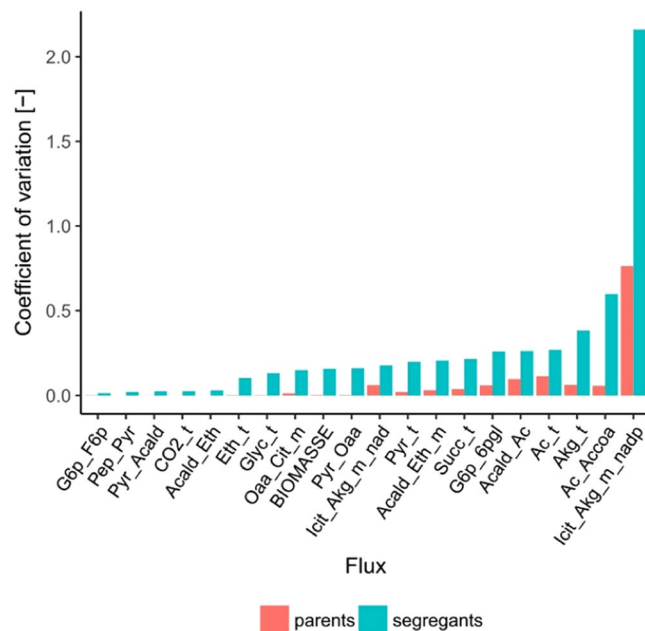
**Table 1.** Flux Selection. Selection of 20 fluxes that are representative of the main metabolic pathways. \*flux abbreviations are encoded as substrate and product connected with “\_”. For mitochondrial reactions, we added “\_m”. Extracellular transport and mitochondrial transport reactions are marked with “\_t” and “\_tm”, respectively. Metabolite abbreviations can be found in Supplementary Table S2.



**Figure 1.** Principle component analysis. PCA of selected estimated fluxes (left), and variation among parents (red and green) and segregant strains (grey) (right). Flux abbreviations are given in Supplementary Tables S2 and S3.

With the first two dimensions explaining 69.1% of trait variation, the PCA of estimated fluxes adequately depicts the variation among strains. The parent strains behave similarly and show only minor differences in their flux profiles, whereas the segregant strains are more divergent. This finding is confirmed by the visualization of trait distributions (Supplementary Fig. S2). The parental strains are located within the population of segregants for the majority of traits.

To further assess variation among strains, coefficients of variation for estimated fluxes were separately calculated for the parent and segregant strains (Fig. 2).



**Figure 2.** : Flux variation. Coefficient of variation of selected representative fluxes among the parent strains of the study (red) and among the resulting F2-segregants (blue). Flux abbreviations are given in Supplementary Tables S2 and S3.

The results demonstrate that the variation among segregant strains regarding CCM fluxes exceeds the variation found between the parent strains, for half of the traits more than ten times (Pep\_Pyr, Pyr\_Acald, Acald\_Eth, Ac\_Accoa, Pyr\_Oaa, Akg\_Succoa\_m, Eth\_t, Glyc\_t, CO2\_t, BIOMASSE). This finding confirms the conclusion drawn by PCA (Fig. 1), which is explicitly that the parent strains show higher similarities in flux distributions than the segregant population. However, differences in variation can be seen between single fluxes. To better visualize trait variation between all segregants, the distributions of fluxes were plotted around the mean value for representative reactions within the central carbon metabolic network (Fig. 3).

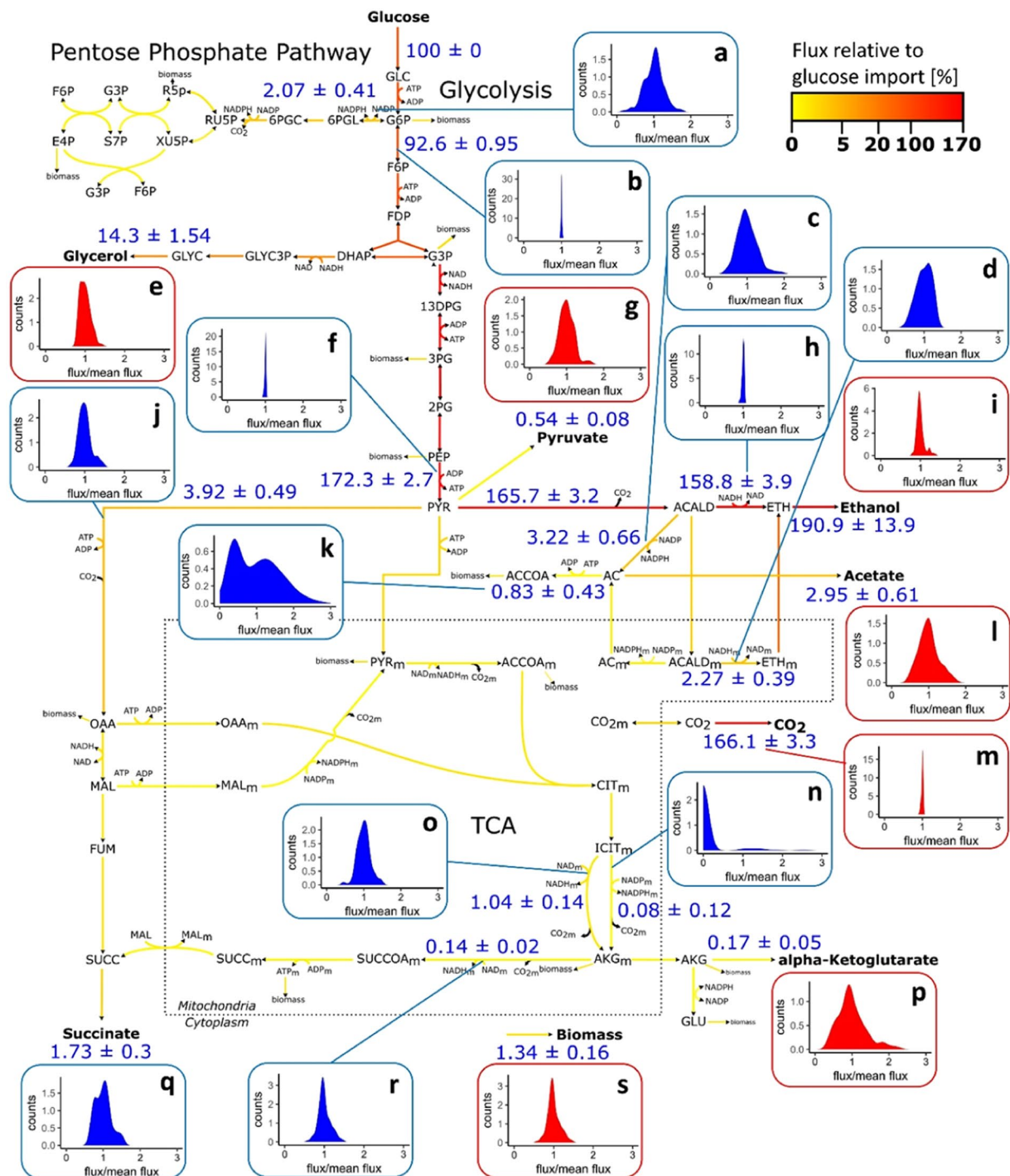
While the fluxes of glycolysis and ethanol synthesis vary by only  $\pm 2.5\%$  around the mean value, the fluxes of the PPP or metabolite production, such as glycerol or acetic acid production, diverge up to 200% around the mean. Most fluxes are normally distributed; however, few outliers can be seen, particularly for  $\alpha$ -ketoglutarate, pyruvate and ethanol excretion. The distributions of two fluxes differ from a normal distribution, namely, the synthesis of acetyl-CoA from acetate (Ac\_Accoa) and the NADP-dependent mitochondrial flux from isocitrate to AKG (Icit\_Akg\_m\_nadp) of the TCA cycle oxidative branch. In the case of Icit\_Akg\_m\_nadp, the analyses indicate that this flux is inactive in the majority of segregant strains and both parents (Supplementary Fig. S2). In the case of Ac\_Accoa, a subpopulation of segregants shows a reduced flux towards acetyl-CoA. The presence of two distinctive populations indicates the major influence of one allele on the trait.

**Genome-wide identification of QTLs influencing metabolic carbon fluxes.** In the first step, QTL mapping was performed on metabolite yields determined during the exponential phase, using a previously obtained segregant marker map<sup>46</sup>. Linkage analysis led to the detection of 8 QTLs on 5 chromosomes influencing 7 traits (Table 2), which included most of the determined metabolite production yields as well as differences in sugar uptake (expressed as G/F ratio, the ratio of glucose and fructose remaining in the medium).

The highest detected LOD score was 4.71 for QTL chr4@152.6 influencing succinate yield, meaning that almost 16% of trait variation can be explained by this locus.

In the second step, QTL mapping was performed on estimated intracellular carbon fluxes. The analysis detected 24 single results of genomic regions influencing intracellular fluxes (Supplementary Table S4). Loci with highly identical borders were grouped to fQTLs (Table 3).

A total of 4 fQTLs on chromosomes II, V, VII and VIII were detected that influence 7 traits, which includes fluxes of glycolysis/ethanol synthesis, glycerol synthesis, the TCA cycle oxidative branch, biomass formation and metabolite transport/excretion. No QTLs could be detected for the fluxes of the PPP, the TCA cycle reductive branch and glutamate cycle, although these fluxes show the most substantial variation among the segregant strains (Fig. 2). As no QTL with an influence on the synthesis of acetyl-CoA from acetate (Ac\_Accoa) was detected, the previous hypothesis from phenotyping that one locus has a major impact on this flux could not be confirmed (Fig. 3). The highest LOD score of 4.63 was found for the influence of QTL chr7@18.0 on glycolysis and ethanol synthesis, meaning that 15.7% of trait variation can be explained by this locus. The QTL region furthermore influences the most traits. In addition to the fluxes of glycolysis and ethanol synthesis, the fluxes of biomass formation, the TCA cycle oxidative branch and metabolite transport/excretion are affected.



**Figure 3.** Average flux distribution. Schematic representation of modelled metabolic network with average flux distribution among segregant strains. Flux strength is expressed as percentage of glucose input and is displayed with a colour gradient from yellow to red. Average flux values  $\pm$  standard deviation are indicated in blue for selected representative metabolic reactions, together with the variation of these fluxes around the mean that is normalized to a value of 1 (a-s). Distributions are displayed in red for reactions constrained by experimental data and in blue for modelled reactions. Metabolite abbreviations are given in Supplementary Table S2.

Combining the results obtained with both QTL mappings, 12 QTLs with a minimum peak distance of 10 cM were detected on 8 chromosomes (Fig. 4), of which 8 QTLs influence metabolite yields or the consumption of sugars during the exponential phase of fermentation, and 4 QTLs influence estimated metabolic fluxes of CCM.

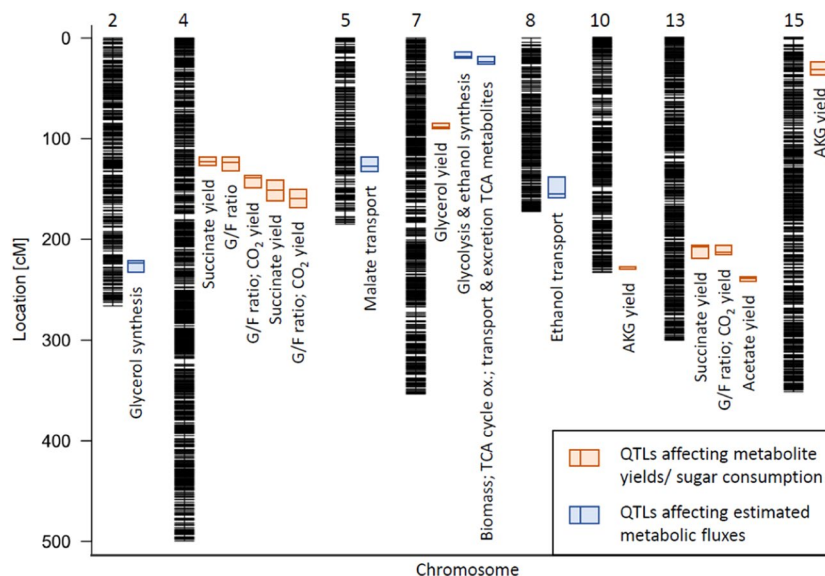
**Validation of detected fQTLs.** To evaluate the solidity of obtained fQTL mapping results, candidate genes within the 2 fQTLs with the highest LOD score were selected for validation. These genes were chosen according to their biological function related to central carbon metabolism and their distance to the QTL peak (Table 4).

Trait	QTL name	Chromosome	QTL start [bp]	QTL end [bp]	LOD
Succinate yield	chr4@122.8	IV	356071	380035	4.42
G/F ratio	chr4@125.4	IV	356071	400864	3.67
G/F ratio	chr4@139.3	IV	409433	448045	3.63
CO <sub>2</sub> yield	chr4@139.3	IV	410742	448045	4.05
Succinate yield	chr4@152.6	IV	426649	488205	4.71
G/F ratio	chr4@160.0	IV	448242	505548	3.76
CO <sub>2</sub> yield	chr4@160.0	IV	448242	505548	4.33
Glycerol yield	chr7@90.3	VII	252047	273771	3.76
AKG yield	chr10@242.3	X	717987	648141	4.38
Ethanol yield	chr13@208.9	XIII	622064	660267	4.22
G/F ratio	chr13@214.5	XIII	624189	648141	3.62
CO <sub>2</sub> yield	chr13@214.5	XIII	624189	648141	3.7
Acetate yield	chr13@237.7	XIII	710548	726277	3.84
AKG yield	chr15@31.8	XV	67745	111309	3.52

**Table 2.** Detected metabolite QTLs. List of QTLs with an influence on metabolite yields and the ratio of residual glucose to fructose concentrations (G/F ratio) during the exponential phase.

Trait	QTL name	Chromosome	QTL start [bp]	QTL end [bp]	LOD max
Glycerol synthesis	chr2@222.9	II	662795	701771	4.58
Malate transport	chr5@128.3	V	354177	400836	4.01
Glycolysis & ethanol synthesis	chr7@18.0	VII	40689	58851	4.63
Biomass	chr7@25.5	VII	52412	82449	3.73
TCA cycle oxidative branch	chr7@25.5	VII	52412	82449	4.05
Transport & excretion TCA metabolites	chr7@25.5	VII	52412	82449	4.05
Ethanol transport	chr8@155.7	VIII	443664	483121	3.45

**Table 3.** Detected fQTLs. List of QTLs with an influence on modelled metabolic fluxes.



**Figure 4.** QTL map. Plot of all detected QTLs against related chromosomes with marker positions. QTLs detected to influence metabolite yields or sugar consumption during the exponential phase (orange) and QTLs detected to influence estimated metabolic fluxes (blue) are displayed as 1-LOD support interval with the peak position as horizontal bar.

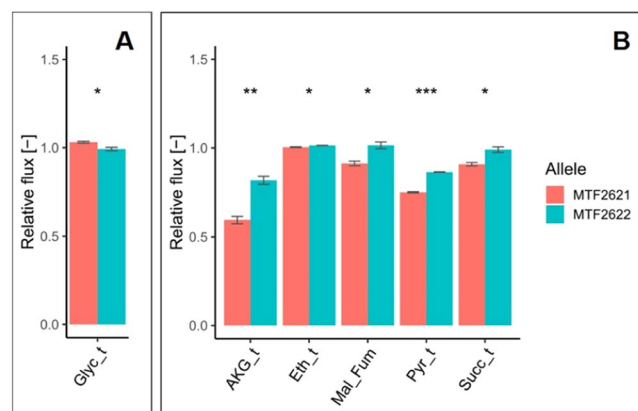
The impact of these genes and their allelic variants was evaluated by RHA. In QTL chr2@222.9, detected to influence glycerol synthesis, *PDB1* with a distance of 22.5 kb to the QTL peak was assessed, and a significant influence of the allelic variants on the trait was detected. The allelic variants of the gene differ in three non-synonymous single-nucleotide polymorphisms (SNPs; Table 5). The MTF2621 allele of *PDB1* increases glycerol synthesis fluxes by 5% (Fig. 5).

QTL name	Trait	Evaluated genes	Different impact of allele on trait as MTF2621/MTF2622 [factor]
chr2@222.9	glycerol synthesis	<i>PDB1</i>	1.05* glycerol synthesis
chr7@25.5	biomass glycolysis & ethanol synthesis TCA metabolite transport & excretion TCA cycle oxidative branch	<i>HAP2</i>	no effect
		<i>VID30</i>	0.99* glycolysis & ethanol synthesis 0.92* – 0.73** TCA metabolite excretion (0.90* TCA cycle reductive branch)

**Table 4.** Verification of fQTLs. Differences caused by selected allelic gene variants regarding the influenced traits were detected by RHA and are given as the ratio of phenotype MTF2621 to phenotype MTF2622. (p-value: \*  $\leq 0.05$ , \*\*  $\leq 0.01$ ).

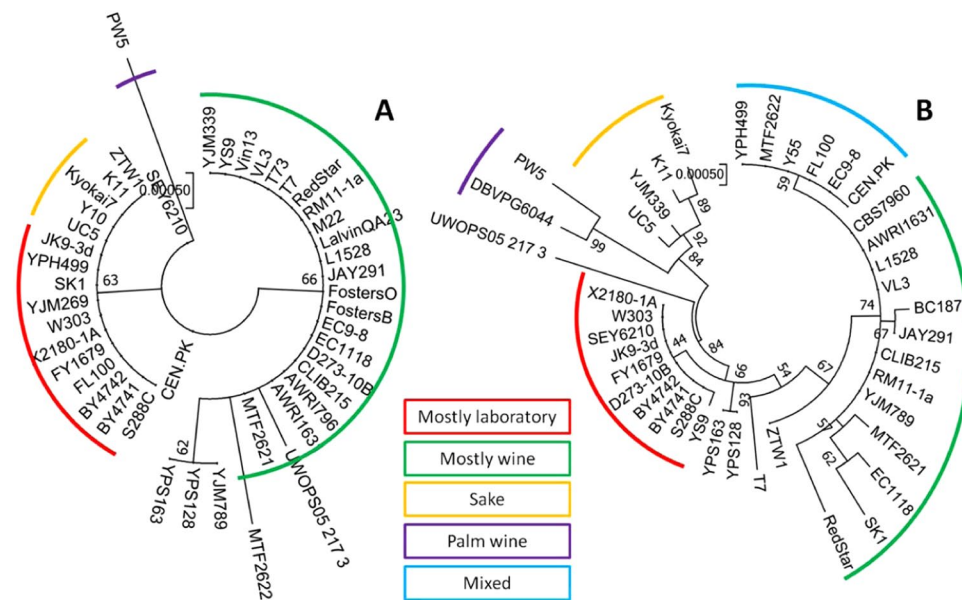
Gene	Length in AA	AA position	S288C	MTF2621	MTF2622
<i>PDB1</i>	366	14	A	A	T
		26	—	A	—
		289	V	V	I
<i>VID30</i>	958	37	H	Y	H
		672	E	E	G
		882	I	V	I

**Table 5.** Polymorphisms between target gene alleles. Differences in amino acid (AA) sequences of expressed validated gene variants caused by non-synonymous SNPs between the parent strains. Comparison of the strains MTF2621 and MTF2622 to *S. cerevisiae* type strain S288C.



**Figure 5.** Allelic effect of *PDB1* (A) and *VID30* (B) on different estimated fluxes of yeast CCM. Fluxes are expressed in relation to the corresponding flux determined in the undeleted heterozygote. (Glyc\_t: glycerol excretion; AKG\_t: AKG excretion; Eth\_t: ethanol excretion; Mal\_Fum: malate to fumarate flux; Pyr\_t: pyruvate excretion; Succ\_t: succinate excretion; p-value: \*  $\leq 0.05$ , \*\*  $\leq 0.01$ , \*\*\*  $\leq 0.001$ ).

In region chrVII:40,689..82,449 detected to influence glycolysis, ethanol synthesis, biomass formation, TCA cycle fluxes and transport/excretion of TCA cycle metabolites, two genes were selected for validation, *HAP2* with a distance of 23.0 kb to the QTL peak, and *VID30* with a distance of 4.0 kb to the QTL peak. While the variants of *HAP2* did not show significant differences regarding estimated metabolic fluxes, the contribution of *VID30* to the detected phenotype variations could be validated (Table 4). The allelic variants of *VID30* differ in three non-synonymous SNPs (Table 5). One SNP lies in the 1000-bp upstream region of the gene at position –318; however, no predicted transcription factor binding site is affected. No SNP was detected in the terminator region. The MTF2621 allele of *Vid30* was found to decrease the fluxes of glycolysis and ethanol synthesis in the parental hemizygote by 1% (Fig. 5). The total variation regarding ethanol synthesis is only 6% among the segregant strains (Fig. 3). Although differences between the hemizygotes can not be directly compared to variation between the haploid segregant cells, the determined decrease by 1% caused by the alleles of *Vid30* indicates a considerable contribution to total variation. To further evaluate the significance of the determined effect of *Vid30* on glycolysis and ethanol synthesis, the accordance of modelled and experimental determined sugar uptake was assessed for all evaluated candidate genes (Supplementary Fig. S3). The mean deviation of the divergence between modelled and measured sugar uptake was 1.2% for triplicate measurements. Therefore, the meaningfulness of the 1% difference in glycolysis/ethanol synthesis found significant for the alleles of *VID30* has to be questioned.



**Figure 6.** Phylogenetic analysis of target genes. Phylogenetic trees of target genes *PDB1* (A) and *VID30* (B) made from variant sequences of strains with different origins. Gene sequences were obtained from the *Saccharomyces* Genome Database (<https://www.yeastgenome.org>). Maximum likelihood trees were constructed by the bootstrap method with 200 replications using MEGA v7.0.26<sup>83</sup>.

In addition to ethanol synthesis, the MTF2621 allele of *Vid30* was found to decrease the excretion of pyruvate,  $\alpha$ -ketoglutarate and succinate by up to 27%. A significant influence of the alleles on the fluxes of the TCA cycle oxidative branch could not be detected. However, the fluxes of the reductive branch of the TCA cycle are significantly affected, with the MTF2621 allele leading to a decrease in fluxes from malate to succinate by 10%.

**Variation of *PDB1* and *VID30* alleles among the *S. cerevisiae* population.** To visualize the natural variation of validated target genes within the *S. cerevisiae* population and to potentially link the variants to strain origins, phylogenetic trees were drawn using publicly available *PDB1* and *VID30* gene sequences (Fig. 6 and Supplementary Table S5).

Regarding *PDB1*, allelic variants from yeast strains of different origins do not show much nucleotide variation. Two main clusters can be seen, one consisting of mostly laboratory strains and the other consisting of mostly wine strains. The allelic variants of the parental strains in this study are comparatively close. In contrast, the phylogenetic tree of *VID30* variants displays more variation between strains of different origins. A probable explanation for the larger variation is a larger gene size. Several clusters can be distinguished, including a laboratory strain cluster, a cluster consisting of African and sake strains and a wine strain cluster with a sub cluster of mixed strains. The parental variants are more separated. While the MTF2621 allele of *VID30* is similar to the allele of strain EC1118, a genotypically close wine x flor strain, the MTF2622 allele is located within the mixed cluster.

## Discussion

Many *S. cerevisiae* traits of interest for applications in traditional food industries or industrial biotechnologies are dependent on the distribution of carbon fluxes within yeast CCM. We recently showed the pathway-dependent variability of flux distributions between *S. cerevisiae* strains, which was linked to the strain origin for some fluxes<sup>32</sup>. These findings suggest the existence of a stock of genetic resources that can help us understand the genetic basis of flux distributions and identify relevant targets for yeast strain improvement. In recent years, powerful methods, such as QTL mapping, have been developed to link phenotypic and genomic variations. Our objective was to assess the potential of QTL mapping to detect genomic regions influencing metabolic fluxes (fQTLs).

To this end, we used a population of 130 F2-segregants obtained from a cross of two wine yeast strains. For these strains, intracellular carbon fluxes were estimated by applying concentrations of extracellular metabolites during the exponential growth phase to a constraint-based stoichiometric model of yeast CCM. Analysis of flux deviations within the population of segregants indicated a positive correlation between the fluxes of the PPP, the TCA cycle oxidative branch and biomass formation, whereas these fluxes were negatively correlated to the fluxes of upper glycolysis. Negative correlations were furthermore found between the fluxes of lower glycolysis/ethanol synthesis and glycerol formation (Fig. 1). These observations are consistent with previous studies<sup>32,50</sup>.

Although the parent strains did not display large variation for most modelled fluxes, substantial variation among the segregants was observed (Fig. 2). For some fluxes, for example, the PPP or main metabolic synthesis fluxes, the variation among segregants reached the variation among *S. cerevisiae* strains from different ecological origins as determined by Nidelet *et al.* 2016<sup>32</sup>. This finding emphasizes the complex nature of intracellular flux distributions and indicates a rich genomic resource for metabolic profile optimization.



With 8 detected QTLs, the number of regions influencing metabolite yields during the exponential phase was higher than the detected 4 QTLs influencing modelled metabolic fluxes. However, the regions from both linkage analyses differ from each other (Fig. 4), demonstrating that the modelling step was crucial for fQTL detection. While QTL mapping of metabolite yields naturally considers a single information, the modelling step preceding fQTL mapping represents the integration of multiple information.

As reported in the literature and seen in our study, variances of certain intracellular fluxes, such as glycolysis, are generally low. As result, increased statistical power is needed for a more thorough determination of the impact of genomic variation on these fluxes. This determination could be achieved by increasing the number of segregants or by performing strategies of multiple QTL mapping, which has the potential to find QTLs with minor contributions.

All QTLs were compared to 8 loci detected during our previous study to influence extracellular metabolite production yields after 80% of the fermentation using the same parent strains<sup>46</sup>. Only QTL chr7@18.0, which influences the fluxes of glycolysis, ethanol synthesis, biomass production, the TCA cycle and transport/excretion of TCA cycle metabolites, was detected in our previous study to influence pyruvate production yield after 80% of the fermentation. This finding indicates that the impact of QTL chr7@18.0 on flux distributions has a long-lasting effect on metabolite formation that can still be detected at the end of fermentation. However, considering the concentrations of extracellular metabolites during the exponential phase and after 80% of the fermentation, all detected QTL regions differ from each other, and there are no common QTLs affecting metabolite production during both phases of fermentation. This finding indicates that different genomic regions could control metabolite production in the growth phase and stationary phase, which actually corresponds to very different physiologic states of yeast during fermentation.

Assessment of the two detected fQTLs with the highest LOD scores by RHA revealed the implication of *PDB1* and *VID30* in the evaluated traits. *PDB1*, shown to influence the fluxes of glycerol synthesis (Table 4, Fig. 5), encodes the beta subunit of pyruvate dehydrogenase (PDH), which is part of the large multienzyme PDH complex<sup>51</sup>. The PDH complex, which includes the other components dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase, converts pyruvate into acetyl-CoA<sup>52</sup>. The allelic variants of *PDB1* differ in three non-synonymous SNPs (Table 5), one of which, SNP V289I, lies in the pyruvate-ferredoxin oxidoreductase domain II of the protein. A possible explanation for the impact of Pdb1 variants on glycerol fluxes is that the MTF2621 allele of Pdb1 shows an increased conversion rate of pyruvate to acetyl-CoA, which leads to a higher formation of the redox cofactor NADH. The resulting cofactor excess is then compensated through increased glycerol synthesis, which maintains redox balance by NADH consumption<sup>53</sup>. However, no significant difference in the estimated flux from pyruvate to acetyl-CoA (Pyr\_Accoa\_m) was detected by RHA for the alleles of *PDB1*.

The second validated gene, *VID30*, possibly influences the fluxes of glycolysis/ethanol synthesis and has an impact on the TCA cycle reductive branch and the excretion of TCA cycle metabolites (Table 4, Fig. 5). Two functions of Vid30 could potentially account for these observed differences: the regulation of genes involved in glutamate/glutamine synthesis and the degradation of various metabolic enzymes. The expression of *VID30* is repressed by ammonia and upregulated in response to low ammonia levels, a characteristic limitation during fermentation in grape must, which was the case in this study. Vid30 regulates various nitrogen catabolic genes, including *GDH1*, *GDH2*, *GDH3*, *GLN1* and *GLT1*. These genes express enzymes involved in the synthesis (and interconversion) of glutamate and glutamine from AKG and ammonia, therefore explaining the role of Vid30 in central carbon metabolism since AKG is part of the fluxes of the TCA cycle oxidative branch. Gdh1, Gdh3 and Gln1 catalyse reactions from AKG to glutamine<sup>54–56</sup>, whereas Gdh2 catalyses the conversion of glutamate to AKG<sup>57</sup>. Glt1 synthesizes glutamate from either AKG or glutamine<sup>58</sup>. In low ammonia environments, Vid30 behaves as a positive regulator of *GDH1*, *GDH3* and *GLT1*, which increases the flux from AKG to glutamate<sup>59</sup>. Since decreased AKG production was detected for the MTF2621 allele of Vid30 (Fig. 5), we suggest that this variant stimulates increased flux from AKG to glutamate through the positive regulation of *GDH1*, *GDH3* and *GLT1*.

Another potential role of Vid30 in central carbon metabolism is its regulation of various metabolic enzymes through degradation. When glucose-starved yeast is again transferred to glucose-rich medium, e.g., during inoculation, metabolism increases the expression of glycolytic enzymes and simultaneously inactivates gluconeogenic enzymes through catabolite inactivation. Vid30 possesses two functions in this process. It acts as a subunit of the glucose induced degradation (GID) protein complex that performs the ubiquitination of enzymes, which leads to their proteasome dependent inactivation<sup>60–63</sup>. Furthermore, Vid30 plays an important role in the formation of vesicles of the vacuole import and degradation pathway<sup>64</sup>, which carries out the degradation of enzymes expressed under the growth on non-fermentable carbon sources<sup>65–67</sup>. Regulation performed in this manner includes the turnover of hexose transporters Hxt3 and Hxt7<sup>68,69</sup>. In addition, various enzymes are regulated through degradation by Vid30 that catalyse gluconeogenesis reactions, such as fructose-1,6-bisphosphatase, cytosolic malate dehydrogenase, isocitrate lyase and phosphoenolpyruvate carboxykinase<sup>70–73</sup>. The reactions catalysed by these enzymes strongly affect the fluxes of glycolysis and the TCA cycle.

The allelic variants of *VID30* differ in three non-synonymous SNPs (Table 5), one of which, SNP V882I, lies in the CTLH/CRA domain of the protein, a protein-protein interaction domain also found in other components of the GID complex. We propose that the SNPs in the MTF2621 variant of Vid30 influence the protein's ability to inactivate hexose transporters and gluconeogenesis enzymes by degradation, hypothetically by an altered affinity to other components of the GID complex. This hypothesis is supported by the observed influence of allelic variants on the fluxes of the TCA cycle reductive branch (Fig. 5), as cytosolic malate dehydrogenase, which catalyses the reaction from malate to oxaloacetate, is among the enzymes inactivated by Vid30<sup>65</sup>. Furthermore, the reported SNPs could affect the role of Vid30 in the regulation of enzymes involved in the synthesis of glutamate from AKG. This hypothesis is supported by the detected significant influence of Vid30 alleles on AKG formation (Fig. 5). On the other hand, a significant difference between the evaluated alleles in the flux from AKG to glutamate could not be detected by RHA. The difference in AKG formation could also be explained by the role of Vid30 in the degradation of isocitrate lyase. The enzyme catalyses the reaction from isocitrate to succinate, which could influence AKG synthesis.

## Conclusion

In this study, we prove the feasibility of using modelled phenotypic data to detect regions in the genome influencing the distribution of carbon fluxes within central carbon metabolism (fQTLs) in *S. cerevisiae*. We used concentrations of extracellular main metabolites during the exponential phase of yeast fermentation to estimate intracellular fluxes, applying a constraint-based model. This strategy allowed the integration of otherwise independent quantifiable traits and resulted in the detection of 4 fQTLs with an influence on 4 main metabolic pathways and various metabolite transport and excretion fluxes. These QTLs could not have been found by linkage analysis considering extracellular metabolite concentrations alone, demonstrating the need for the modelling step.

The solidity of our approach was further confirmed by the validation of two target genes within the identified fQTLs, *PDB1* and *VID30*. The allelic variants of *PDB1* account for variation in the glycerol synthesis flux, which we hypothesize to be caused by redox imbalances as a result of altered pyruvate conversion. The variants of *VID30* impact the fluxes of glycolysis, ethanol synthesis and the TCA cycle, which we propose to be caused by differences in the regulation of enzymes catalysing glutamate formation or differences in the catabolite-induced degradation of enzymes involved in sugar uptake, gluconeogenesis and the TCA cycle. Compared to strains of other origins, the parental variants of the evaluated target genes are comparatively close. The characterization of more distant variants and an evaluation of their influence on intracellular flux distributions will disclose genetic resources that bear further potential to shape the metabolic profile of strains.

In summary, our findings of fQTLs and allelic variants impacting metabolic fluxes increase our knowledge of the links between genomic variation and yeast metabolic properties and provide a proof-of-concept for the applicability of QTL mapping on modelled metabolic fluxes. This result offers exciting opportunities for uncovering superior allelic variants impacting these traits, which could be used to improve strains for manifold purposes, e.g., the production of biofuels or other bulk and fine chemicals.

## Materials and Methods

**Media.** Yeast strains were cultured at 28 °C in yeast extract peptone dextrose (YPD) media, containing 10 g/L yeast extract, 20 g/L peptone and 20 g/L glucose. Solid YPD media contained 1.5% agar. Selective YPD media contained 200 µg/mL geneticin (G418), 200 µg/mL nourseothricin (clonNAT) or 200 µg/mL hygromycin B.

Fermentations were carried out in synthetic grape must (SM200) as described by Bely *et al.* 1990<sup>74</sup>. The medium contains glucose and fructose at a concentration of 100 g/L each, and assimilable nitrogen in the form of ammonium and free amino acids at a concentration of 200 mg/L.

**Strain generation.** This study used a set of 130 F2-segregants generated in a former study to map QTLs for main and volatile metabolite production by *S. cerevisiae* during alcoholic fermentation<sup>46</sup>. The parent strains of this segregant population are strain MTF2621 (haploid spore of 4CAR1 [ $\Delta HO::Neo^r$ ]) and MTF2622 (haploid spore of T73 [ $\Delta HO::Nat^r$ ]). Strain T73 belongs to the phylogenetic clade of wine strains, whereas strain 4CAR1 belongs to the group of champagne strains, which originated through crossings between strains of the wine clade and the flor clade<sup>75</sup>.

**Genotyping of strains.** The segregant strain marker map used for linkage analysis in this study was generated during a previous study in our working group<sup>46</sup>. Parent and segregant strains were genotyped by whole genome sequencing using Illumina technology, and a global set of 18155 biallelic variant positions was obtained. This set was reduced to a dataset of 3727 SNP markers with a minimum spacing of 2.0 kb between SNPs. Four segregant strains with the most ambiguous markers were excluded to increase the meaningfulness of the analysis. One strain was excluded because of its close genomic proximity to another segregant. This exclusion left a population of 125 F2-segregants for statistical analyses. Information available in the *Saccharomyces* Genome Database (<https://www.yeastgenome.org>) was used to associate SNPs with annotated protein domains. The effect of detected SNPs on putative transcription factor binding sites was analysed using YEASTRACT (release 2017)<sup>76</sup>.

**Phenotyping of strains.** Segregant strains were phenotyped in duplicate with the parent strains as a control. Sterilized 300-mL glassware mini fermenters were filled with 280 mL of SM200 (synthetic must with 200 mg/L of assimilable nitrogen) and closed with an air lock. The fermenters were inoculated with overnight yeast cultures to a cell density of  $1 \times 10^6$  cells/mL, weighed and left at 24 °C under stirring (300 rpm). Fermentations were sampled during the exponential phase (when 10 g/L of CO<sub>2</sub> was produced), which was determined by weighting the fermenters regularly. For each fermentation, dry biomass was determined in duplicate by filtering 10 mL of cell suspension through a nitrocellulose membrane with a porosity of 0.45 µm (Millipore, France) and known dry weight. The membrane was rinsed twice with 10 mL of distilled water, dried for 48 h and weighed to determine the dry biomass of the sample. The concentrations of unfermented sugars (glucose and fructose) and extracellular carbon metabolites (ethanol, glycerol, acetic acid, succinic acid, pyruvic acid, and  $\alpha$ -ketoglutaric acid) in the fermentation medium were determined by high-performance liquid chromatography (HPLC). The flow rate of the device (HPLC 1290 Infinity, Agilent Technologies, USA) was set to 0.6 mL/min (0.005 N H<sub>2</sub>SO<sub>4</sub>). Samples were separated by a pre-column and an ion-exclusion column (Phenomenex REZEX™ ROA-Organic Acid H+(8%)), which was thermostatically controlled at 60 °C. Compounds were detected using a refractometer in combination with a UV spectrometer at 210 nm. Chromatograms were processed with Agilent EZChrom software.

**Modelling of metabolic fluxes.** To predict intracellular metabolic fluxes, we used DynamoYeast, a previously developed model of yeast central carbon metabolism<sup>18</sup>. This model has been chosen for several reasons. First, it had been designed specifically to predict fluxes in oenological fermentation. It covers *S. cerevisiae* central carbon metabolism with the specification of anaerobic metabolism, e.g., TCA fluxes are composed of an oxidative and reductive branch and therefore do not form a cycle. Second, the biomass reaction function of the model had been calibrated with wine yeast strain EC1118, which is phylogenetically closer to the parent strains used

in this study than the widely used laboratory reference strain<sup>18</sup>. Third, it had been used and validated in two previous projects in very similar conditions<sup>18,32</sup>. Last, the choice of a comparatively small model (favored over a genome-scale model) is consistent with other studies in literature that use large scale models to examine qualitative phenotypes<sup>13</sup>, whereas small-scale models are used for quantitative modelling<sup>20–22,77</sup>.

The DynamoYeast model covers 68 reactions (Supplementary Table S3) and 61 metabolites (Supplementary Table S2) of central carbon metabolism, and distinguishes three compartments, the mitochondria, cytoplasm and extracellular excretion. Extracellular metabolite concentrations (in mmol/L) and dry mass weight (in g/L) obtained by the phenotyping of segregant strains during the middle of the exponential phase were used to constrain the model. The error margin for the metabolite reaction rate boundaries of the model was set to  $\pm 2.5\%$ . Flux distribution throughout the metabolic network was obtained for each segregant by mass balance analysis with a minimization of glucose input as the objective function. As consequence, our method directly computes mass distribution, unlike other constraint-based methods, e.g., metabolic flux analysis (Vallino and Stephanopoulos, 1993<sup>78</sup>), that derive mass data to obtain flux distributions (see Celton *et al.* 2012<sup>18</sup> for details about the methodology). For computing mass distribution, we follow the assumption that all fluxes - along with the biomass composition - are constant during the exponential phase. The estimation of mass distribution allows for the expression of fluxes as yields instead of reaction rates.

Fructose was treated as glucose in the modelling approach as this did not impact flux predictions. Finally, estimated fluxes were normalized to the predicted sugar uptake in order to enable the comparison of relative flux distributions between strains. Due to block effects concerning the determination of succinate concentrations, succinate excretion fluxes were constrained by setting a fixed range that corresponded to the maximum succinate excretion flux variation between *S. cerevisiae* strains determined in a previous study by our working group<sup>32</sup>.

All predictions were performed using the programming language R v3.2.3 with the R/sybil v2.0.0 and R/sybilSBML v2.0.11 libraries<sup>79</sup>.

**QTL mapping.** QTL mapping of modelled metabolic fluxes was carried out by single QTL mapping (interval mapping method) using the genotype and phenotype information of segregant strains. Linkage analysis was performed using the programming language R v3.2.3 (<https://www.r-project.org>) with the R/ql v1.40–8 and R/eql v1.1–7 libraries<sup>80</sup>. Two different phenotype models were tested, the normal model using Haley-Knott regression and non-parametric analysis. The statistical analyses resulted in logarithm of odds (LOD) scores for each marker and pseudo-marker every 2.5 cM. Individual LOD score thresholds for a false discovery rate of 0.05 were determined with 1000 permutations. If the same locus was detected with both phenotype models, the results with the higher LOD score were selected. An interval estimate of the location of each QTL was obtained as the region in which the LOD score is within 1 unit of the peak LOD score (1-LOD support interval). QTL mapping results for single metabolic reactions were considered to be a common fQTL if their peaks were less than 10 cM apart.

**Reciprocal hemizyosity analysis.** Molecular dissection of the detected QTLs was performed by RHA<sup>81,82</sup>. Target genes in QTLs were chosen according to a biological function associated with central carbon metabolism and the gene's proximity to the determined QTL peak. The gene sequences were deleted in both parent strains by homologous recombination with a disruption cassette containing the hygromycin B resistance gene (*hph<sup>r</sup>*). The disruption cassettes were amplified by polymerase chain reaction (PCR) of the plasmid pAG32 with the primers del\_(GENE)\_fw and del\_(GENE)\_rv (Supplementary Table S1). Positive integration was selected by plating the transformed cells on YPD-agar plates containing hygromycin B. Correct gene deletion was verified by PCR using primer test\_(GENE)\_fw, which binds upstream of the deleted gene, and primer Hygro\_rv, which binds within the deletion cassette. Subsequently, deleted parent strains were mated with the opposite undeleted parent to form a heterozygote that is hemizygous for the target gene. Hemizygous constructions were phenotyped in triplicate. The significance of the influence of allelic gene variants on modelled metabolic fluxes was evaluated by student's t-test. If a variant's impact on several fluxes was tested, p-values were not adjusted for multiple comparisons.

## Data availability

Genome sequencing data generated during the current study is available from NCBI under bioproject number PRJNA433287; SNP data, marker map and phenotypic data set are available from the INRA Dataverse: <https://data.inra.fr/dataset.xhtml?persistentId=doi:10.15454/C1F8MO>.

Received: 12 May 2019; Accepted: 21 December 2019;

Published online: 07 February 2020

## References

- Cavaliere, D., McGovern, P. E., Hartl, D. L., Mortimer, R. & Polsinelli, M. Evidence for *S. cerevisiae* fermentation in ancient wine. *J. Mol. Evol.* **57**, S226–S232 (2003).
- Nielsen, J., Larsson, C., van Maris, A. & Pronk, J. T. Metabolic engineering of yeast for production of fuels and chemicals. *Curr. Opin. Biotechnol.* **24**, 398–404 (2013).
- Mattanovich, D., Gasser, B., Egermeier, M., Marx, H. & Sauer, M. Industrial Microorganisms: *Saccharomyces cerevisiae* and other Yeasts. *Ind. Biotechnol. Microorg.* **2**, 673–686 (2017).
- Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. *Physiology of the bacterial cell: a molecular approach*. **20**, (Sinauer Associates Sunderland, MA, 1990).
- Ma, H.-W. & Zeng, A.-P. The connectivity structure, giant strong component and centrality of metabolic networks. *Bioinformatics* **19**, 1423–1430 (2003).
- Bechthold, I., Bretz, K., Kabasci, S., Kopitzky, R. & Springer, A. Succinic acid: a new platform chemical for biobased polymers from renewable resources. *Chem. Eng. Technol.* **31**, 647–654 (2008).
- Ahn, J. H., Jang, Y.-S. & Lee, S. Y. Production of succinic acid by metabolically engineered microorganisms. *Curr. Opin. Biotechnol.* **42**, 54–66 (2016).

8. Chen, Y., Siewers, V. & Nielsen, J. Profiling of cytosolic and peroxisomal acetyl-CoA metabolism in *Saccharomyces cerevisiae*. *PLoS One* **7**, e42475 (2012).
9. Pfleger, B. F., Gossing, M. & Nielsen, J. Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metab. Eng.* **29**, 1–11 (2015).
10. Nielsen, J. It is all about metabolic fluxes. *J. Bacteriol.* **185**, 7031–7035 (2003).
11. Julleson, D., David, F., Pfleger, B. & Nielsen, J. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnol. Adv.* **33**, 1395–1402 (2015).
12. Van Gulik, W. M. & Heijnen, J. J. A metabolic network stoichiometry analysis of microbial growth and product formation. *Biotechnol. Bioeng.* **48**, 681–698 (1995).
13. Österlund, T., Nookaew, I. & Nielsen, J. Fifteen years of large scale metabolic modeling of yeast: Developments and impacts. *Biotechnol. Adv.* **30**, 979–988 (2012).
14. Palsson, B. The challenges of in silico biology. *Nat. Biotechnol.* **18**, 1147 (2000).
15. Varma, A. & Palsson, B. O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **60**, 3724–3731 (1994).
16. García Sánchez, C. E., Sáez, T. & Gonzalo, R. Comparison and analysis of objective functions in flux balance analysis. *Biotechnol. Prog.* **30**, 985–991 (2014).
17. Celton, M. *et al.* A comparative transcriptomic, fluxomic and metabolomic analysis of the response of *Saccharomyces cerevisiae* to increases in NADPH oxidation. *BMC Genomics* **13**, 317 (2012).
18. 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.* **14**, 366–379 (2012).
19. Vargas, F. A., Pizarro, F., Pérez-Correa, J. R. & Agosin, E. Expanding a dynamic flux balance model of yeast fermentation to genome-scale. *BMC Syst. Biol.* **5**, 75 (2011).
20. Quirós, M. *et al.* Metabolic flux analysis during the exponential growth phase of *Saccharomyces cerevisiae* in wine fermentations. *PLoS One* **8**, e71909 (2013).
21. Blank, L. M., Lehmebeck, F. & Sauer, U. Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts. *FEMS Yeast Res.* **5**, 545–558 (2005).
22. Velagapudi, V. R., 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.* **132**, 395–404 (2007).
23. Strobe, P. K. *et al.* The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Res.* **25**, 762–774 (2015).
24. 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.* **102**, 73–83 (2005).
25. Fay, J. C. & Benavides, J. A. Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. *PLoS Genet.* **1**, e5 (2005).
26. Legras, J., Merdinoglu, D., Cornuet, J. & Karst, F. Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol. Ecol.* **16**, 2091–2102 (2007).
27. Liti, G. *et al.* Population genomics of domestic and wild yeasts. *Nature* **458**, 337 (2009).
28. Warringer, J. *et al.* Trait variation in yeast is defined by population history. *PLoS Genet.* **7**, e1002111 (2011).
29. Legras, J.-L. *et al.* Adaptation of *S. cerevisiae* to fermented food environments reveals remarkable genome plasticity and the footprints of domestication. *Mol. Biol. Evol.* **35**, 1712–1727 (2018).
30. Spor, A. *et al.* Niche-driven evolution of metabolic and life-history strategies in natural and domesticated populations of *Saccharomyces cerevisiae*. *BMC Evol. Biol.* **9**, 296 (2009).
31. Camarasa, C., Sanchez, I., Brial, P., Bigey, F. & Dequin, S. Phenotypic landscape of *Saccharomyces cerevisiae* during wine fermentation: evidence for origin-dependent metabolic traits. *PLoS One* **6**, e25147 (2011).
32. Nidelet, T., Brial, P., Camarasa, C. & Dequin, S. Diversity of flux distribution in central carbon metabolism of *S. cerevisiae* strains from diverse environments. *Microb. Cell Fact.* **15**, 58 (2016).
33. Patil, K. R. & Nielsen, J. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Natl. Acad. Sci. USA* **102**, 2685–2689 (2005).
34. Agren, R., Otero, J. M. & Nielsen, J. Genome-scale modeling enables metabolic engineering of *Saccharomyces cerevisiae* for succinic acid production. *J. Ind. Microbiol. Biotechnol.* **40**, 735–747 (2013).
35. Bro, C., Regenber, B., Förster, J. & Nielsen, J. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab. Eng.* **8**, 102–111 (2006).
36. Burgard, A. P. & Maranas, C. D. Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnol. Bioeng.* **82**, 670–677 (2003).
37. Otero, J. M. *et al.* Industrial systems biology of *Saccharomyces cerevisiae* enables novel succinic acid cell factory. *PLoS One* **8**, e54144 (2013).
38. Bundy, J. G. *et al.* Evaluation of predicted network modules in yeast metabolism using NMR-based metabolite profiling. *Genome Res.* **17**, 510–519 (2007).
39. Swinnen, S., Thevelein, J. M. & Nevoigt, E. Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **12**, 215–227 (2012).
40. Hubmann, G. *et al.* Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering. *Metab. Eng.* **17**, 68–81 (2013).
41. Hu, X. H. *et al.* Genetic dissection of ethanol tolerance in the budding yeast *Saccharomyces cerevisiae*. *Genetics* **175**, 1479–1487 (2007).
42. Hubmann, G. *et al.* Identification of multiple interacting alleles conferring low glycerol and high ethanol yield in *Saccharomyces cerevisiae* ethanolic fermentation. *Biotechnol. Biofuels* **6**, 87 (2013).
43. Pais, T. M. *et al.* Comparative Polygenic Analysis of Maximal Ethanol Accumulation Capacity and Tolerance to High Ethanol Levels of Cell Proliferation in Yeast. *PLoS Genet.* **9** (2013).
44. Swinnen, S., Ho, P.-W., Klein, M. & Nevoigt, E. Genetic determinants for enhanced glycerol growth of *Saccharomyces cerevisiae*. *Metab. Eng.* **36**, 68–79 (2016).
45. Steyer, D. *et al.* QTL mapping of the production of wine aroma compounds by yeast. *BMC Genomics* **13**, 573 (2012).
46. Eder, M. *et al.* QTL mapping of volatile compound production in *Saccharomyces cerevisiae* during alcoholic fermentation. *BMC Genomics* **19**, 166 (2018).
47. 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.* **14**, 68 (2015).
48. Brice, C., Sanchez, I., Bigey, F., Legras, J.-L. & Blondin, B. A genetic approach of wine yeast fermentation capacity in nitrogen-starvation reveals the key role of nitrogen signaling. *BMC Genomics* **15**, 495 (2014).
49. Poolman, M. G., Sebu, C., Pidcock, M. K. & Fell, D. A. Modular decomposition of metabolic systems via null-space analysis. *J. Theor. Biol.* **249**, 691–705 (2007).
50. Heyland, J., Fu, J. & Blank, L. M. Correlation between TCA cycle flux and glucose uptake rate during respiro-fermentative growth of *Saccharomyces cerevisiae*. *Microbiology* **155**, 3827–3837 (2009).

51. Miran, S. G., Lawson, J. E. & Reed, L. J. Characterization of PDH beta 1, the structural gene for the pyruvate dehydrogenase beta subunit from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **90**, 1252–1256 (1993).
52. Pronk, J. T., Yde Steensma, H. & van Dijken, J. P. Pyruvate metabolism in *Saccharomyces cerevisiae*. *Yeast* **12**, 1607–1633 (1996).
53. van Dijken, J. P. & Scheffers, W. A. Redox balances in the metabolism of sugars by yeasts. *FEMS Microbiol. Lett.* **32**, 199–224 (1986).
54. Avendaño, A., Deluna, A., Olivera, H., Valenzuela, L. & Gonzalez, A. GDH3 encodes a glutamate dehydrogenase isozyme, a previously unrecognized route for glutamate biosynthesis in *Saccharomyces cerevisiae*. *J. Bacteriol.* **179**, 5594–5597 (1997).
55. Moye, W. S., Amuro, N., Rao, J. K. & Zalkin, H. Nucleotide sequence of yeast GDH1 encoding nicotinamide adenine dinucleotide phosphate-dependent glutamate dehydrogenase. *J. Biol. Chem.* **260**, 8502–8508 (1985).
56. Mitchell, A. P. & Magasanik, B. Biochemical and physiological aspects of glutamine synthetase inactivation in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **259**, 12054–12062 (1984).
57. Miller, S. M. & Magasanik, B. Role of NAD-linked glutamate dehydrogenase in nitrogen metabolism in *Saccharomyces cerevisiae*. *J. Bacteriol.* **172**, 4927–4935 (1990).
58. Filetici, P., Martegani, M. P., Valenzuela, L., González, A. & Ballario, P. Sequence of the GLT1 gene from *Saccharomyces cerevisiae* reveals the domain structure of yeast glutamate synthase. *Yeast* **12**, 1359–1366 (1996).
59. van der Merwe, G. K., Cooper, T. G. & van Vuuren, H. J. J. Ammonia regulates VID30 expression and Vid30p function shifts nitrogen metabolism toward glutamate formation especially when *Saccharomyces cerevisiae* is grown in low concentrations of ammonia. *J. Biol. Chem.* **276**, 28659–28666 (2001).
60. Regelmann, J. *et al.* Catabolite degradation of fructose-1, 6-bisphosphatase in the yeast *Saccharomyces cerevisiae*: a genome-wide screen identifies eight novel GID genes and indicates the existence of two degradation pathways. *Mol. Biol. Cell* **14**, 1652–1663 (2003).
61. Hämmerle, M. *et al.* Proteins of newly isolated mutants and the amino-terminal proline are essential for ubiquitin-proteasome-catalyzed catabolite degradation of fructose-1, 6-bisphosphatase of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 25000–25005 (1998).
62. Santt, O. *et al.* The yeast GID complex, a novel ubiquitin ligase (E3) involved in the regulation of carbohydrate metabolism. *Mol. Biol. Cell* **19**, 3323–3333 (2008).
63. Menssen, R. *et al.* Exploring the topology of the Gid complex, the E3 ubiquitin ligase involved in catabolite-induced degradation of gluconeogenic enzymes. *J. Biol. Chem.* **287**, 25602–25614 (2012).
64. Alibhoy, A. A., Giardina, B. J., Dunton, D. D. & Chiang, H.-L. Vid30 is required for the association of Vid vesicles and actin patches in the vacuole import and degradation pathway. *Autophagy* **8**, 29–46 (2012).
65. Hung, G.-C., Brown, C. R., Wolfe, A. B., Liu, J. & Chiang, H.-L. Degradation of the gluconeogenic enzymes fructose-1, 6-bisphosphatase and malate dehydrogenase is mediated by distinct proteolytic pathways and signaling events. *J. Biol. Chem.* **279**, 49138–49150 (2004).
66. Shieh, H.-L. & Chiang, H.-L. *In vitro* reconstitution of glucose-induced targeting of fructose-1, 6-bisphosphatase into the vacuole in semi-intact yeast cells. *J. Biol. Chem.* **273**, 3381–3387 (1998).
67. Huang, P.-H. & Chiang, H.-L. Identification of novel vesicles in the cytosol to vacuole protein degradation pathway. *J. Cell Biol.* **136**, 803–810 (1997).
68. Snowdon, C., Hlynialuk, C. & Van Der Merwe, G. Components of the Vid30c are needed for the rapamycin-induced degradation of the high-affinity hexose transporter Hxt7p in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **8**, 204–216 (2007).
69. Snowdon, C. & Van der Merwe, G. Regulation of Hxt3 and Hxt7 turnover converges on the Vid30 complex and requires inactivation of the Ras/cAMP/PKA pathway in *Saccharomyces cerevisiae*. *PLoS One* **7**, e50458 (2012).
70. Carlson, M. Glucose repression in yeast. *Curr. Opin. Microbiol.* **2**, 202–207 (1999).
71. Holzer, H. Proteolytic catabolite inactivation in *Saccharomyces cerevisiae*. *Revis. sobre Biol. Cel. RBC* **21**, 305–319 (1988).
72. Gancedo, J. M. Yeast carbon catabolite repression. *Microbiol. Mol. Biol. Rev.* **62**, 334–361 (1998).
73. López-Boado, Y. S., Herrero, P., Gascon, S. & Moreno, F. Catabolite inactivation of isocitrate lyase from *Saccharomyces cerevisiae*. *Arch. Microbiol.* **147**, 231–234 (1987).
74. Bely, M., Sablayrolles, J.-M. & Barre, P. Automatic detection of assimilable nitrogen deficiencies during alcoholic fermentation in oenological conditions. *J. Ferment. Bioeng.* **70**, 246–252 (1990).
75. Coi, A. L. *et al.* Genomic signatures of adaptation to wine biological ageing conditions in biofilm-forming flor yeasts. *Mol. Ecol.* **26**, 2150–2166 (2017).
76. Teixeira, M. C. *et al.* The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **42**, D161–D166 (2013).
77. Gombert, A. K., dos Santos, M. M., Christensen, B. & Nielsen, J. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J. Bacteriol.* **183**, 1441–1451 (2001).
78. Vallino, J. J. & Stephanopoulos, G. Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnol. Bioeng.* **41**, 633–646 (1993).
79. Gelius-Dietrich, G., Desouki, A. A., Fritzemeier, C. J. & Lercher, M. J. sybil—Efficient constraint-based modelling in R. *BMC Syst. Biol.* **7**, 125 (2013).
80. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889 (2003).
81. Steinmetz, L. M. *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**, 326–330 (2002).
82. Warringer, J., Liti, G. & Blomberg, A. Yeast reciprocal hemizygosity to confirm the causality of a quantitative trait loci-associated gene. *Cold Spring Harb. Protoc.* **2017**, pdb-prot089078 (2017).
83. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).

## Acknowledgements

We are grateful to Vincent Fromion for his contribution to the revision of the manuscript. The research leading to the generation of these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007–2013/ under REA grant agreement no. 606795, project acronym: YEASTCELL. The funding body was not involved in the design of the study, in the collection, analysis and interpretation of data or in writing the manuscript.

## Author contributions

M.E. conducted the experiments, contributed to metabolic modelling and linkage analysis, and wrote the manuscript. T.N. established and performed the metabolic modelling. I.S. established and performed the linkage analysis. C.C. contributed to the phenotyping of strains. J.L.L. analysed the genotyping data. S.D. conceived and planned the study, with contribution of C.C. and J.L.L. All authors revised the manuscript and read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-57857-3>.

**Correspondence** and requests for materials should be addressed to S.D.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020