A Multiphase Multiobjective Dynamic Genome-Scale Model Shows Different Redox Balancing among Yeast Species of the *Saccharomyces* Genus in Fermentation

David Henriques,a Romain Minebois,b Sebastián N. Mendoza,c Laura G. Macías,b Roberto Pérez-Torrado,b Eladio Barrio,d Bas Teusink,c Amparo Querol,b Eva Balsa-Cantoa

a(Bio)process Engineering Group, IIM-CSIC, Vigo, Spain
bSystems Biology in Yeast of Biotechnological Interest, IATA-CSIC, Paterna, Spain
cSystems Bioinformatics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
dDepartment of Genetics, University of Valencia, Burjassot, Spain

ABSTRACT Yeasts constitute over 1,500 species with great potential for biotechnology. Still, the yeast *Saccharomyces cerevisiae* dominates industrial applications, and many alternative physiological capabilities of lesser-known yeasts are not being fully exploited. While comparative genomics receives substantial attention, little is known about yeasts’ metabolic specificity in batch cultures. Here, we propose a multiphase multiobjective dynamic genome-scale model of yeast batch cultures that describes the uptake of carbon and nitrogen sources and the production of primary and secondary metabolites. The model integrates a specific metabolic reconstruction, based on the consensus Yeast8, and a kinetic model describing the time-varying culture environment. In addition, we proposed a multiphase multiobjective flux balance analysis to compute the dynamics of intracellular fluxes. We then compared the metabolism of *S. cerevisiae* and *Saccharomyces uvarum* strains in a rich medium fermentation. The model successfully explained the experimental data and brought novel insights into how cryotolerant strains achieve redox balance. The proposed model (along with the corresponding code) provides a comprehensive picture of the main steps occurring inside the cell during batch cultures and offers a systematic approach to prospect or metabolically engineering novel yeast cell factories.

IMPORTANCE Nonconventional yeast species hold the promise to provide novel metabolic routes to produce industrially relevant compounds and tolerate specific stressors, such as cold temperatures. This work validated the first multiphase multiobjective genome-scale dynamic model to describe carbon and nitrogen metabolism throughout batch fermentation. To test and illustrate its performance, we considered the comparative metabolism of three yeast strains of the *Saccharomyces* genus in rich medium fermentation. The study revealed that cryotolerant *Saccharomyces* species might use the γ-aminobutyric acid (GABA) shunt and the production of reducing equivalents as alternative routes to achieve redox balance, a novel biological insight worth being explored further. The proposed model (along with the provided code) can be applied to a wide range of batch processes started with different yeast species and media, offering a systematic and rational approach to prospect nonconventional yeast species metabolism and engineering novel cell factories.

KEYWORDS batch fermentation, *Saccharomyces* species, cryotolerant species, dynamic genome-scale models, redox balance

Yeasts have been used to produce fermented foods and beverages for millennia and are among the most frequently used microorganisms in biotechnology. *Saccharomyces cerevisiae* dominates the scene and many research efforts focus on engineering this species for...
particular applications (for examples, see references 1 to 4). Nowadays it is used to produce glycerol (5), biopharmaceutical proteins (6), or secondary metabolites, such as aromas or bio-flavors (7, 8).

However, yeasts constitute a large group of 1,500 (so far) described species, and much less attention has been paid to nonconventional yeasts. These species remain a mostly untapped resource of alternative metabolic routes for substrate use and product formation as well as tolerances to specific stressors (9, 10). To exploit these alternatives efficiently, it is essential to understand the metabolic pathways of these species. Given the complexity of the endeavor, a modeling approach becomes indispensable.

Genome-scale models (GEMs) can contextualize high-throughput data and predict genotype-environment-phenotype relationships (11, 12). While GEMs have been widely used for the study and metabolic engineering of *S. cerevisiae* strains in continuous (steady-state) fermentations (13), their use to predict batch (dynamic) fermentation is still scarce. Nevertheless, many yeast-based processes operate in batch mode.

In batch operation, cell culture follows a growth curve with the following phases: lag phase, exponential growth, growth under nutrient limitation, stationary phase, and cellular decay. Available dynamic GEMs of yeast metabolism focus on the exponential phase and explain reasonably well the measured dynamics of biomass growth, carbon source uptake, and the production of relevant primary metabolites (14–17). The development of GEMs that describe the five phases of batch processes, considering carbon and nitrogen metabolism throughout time and explaining secondary metabolism, is still required.

In this work, we derived a multiphase and multiobjective dynamic genome-scale model of batch fermentation, which accounts for carbon and nitrogen metabolism throughout time and explains secondary metabolism. The model required various refinements to succeed: (i) a novel metabolic reconstruction, based on an extension of the current consensus genome-scale model of *S. cerevisiae* (Yeast8 [18]); (ii) multiphase multiobjective implementation of a parsimonious flux balance analysis (pFBA) (19) to compute the dynamics of the intracellular fluxes; (iii) a model of protein turnover to explain nitrogen homeostasis; and (iv) a dynamic biomass equation to account for biomass composition variations throughout the process.

As a relevant case study, we considered the metabolism of *S. cerevisiae* and *Saccharomyces uvarum* strains in rich medium fermentation. Recent studies revealed that *S. uvarum* strains show interesting physiological properties. *S. uvarum* is more cryotolerant than *S. cerevisiae*, produces more glycerol and less ethanol than *S. cerevisiae* wine strains, and has different aroma profiles (20–23). In addition, traits such as its increased 2-phenylethanol (24, 25) yield could make this species a good candidate for metabolic engineering studies (26).

We applied the proposed model to investigate the origin of the phenotypic divergence between species. The model explained the experimental data successfully and revealed differences into how species achieve redox balance. Predicted intracellular fluxes led us to hypothesize that cryotolerant yeast strains can use the γ-aminobutyric acid (GABA) shunt as an alternative NADPH source and store reductive power—necessary to subdue oxidative stress under cold conditions—in lipids or other polymers. Additionally, our results are compatible with recent experimental observations showing that most carbon skeletons used to form higher alcohols (i.e., isoamyl alcohol, isobutanol, and 2-phenylethanol) are synthesized de novo.

**RESULTS**

**Novel metabolic reconstruction.** We updated the Yeast8 consensus genome-scale reconstruction of *S. cerevisiae* S288C (v.8.3.1) (18) to include 38 metabolites and 50 reactions to explain secondary metabolism (Table S1). Furthermore, comprehensive metabolic annotations, such as SBO (Systems Biology Ontology) terms and MetaNetX identifiers, were added to the new metabolites and reactions.

Among the metabolites added, 13 aroma compounds were included. Noticeably, we found that prior genome-scale reconstructions lacked methionol and tyrosol impeding
Dynamic Genome-Scale Modeling of Yeast Fermentation

simulated growth on methionine and tyrosine as sole nitrogen sources, which is known to be possible for several *S. cerevisiae* strains, including S288C.

MetaDraft, AuReMe, and the results from the orthology analysis were used to create strain-specific models for two wine strains, *S. cerevisiae* T73 and *S. uvarum* BMV58, and a strain, *S. uvarum* CECT12600, found in nonfermentation environments (further details can be found in Text S1). These strains are referred to as ScT73, SuBMV58, and SuCECT12600 here. The three models had 2, 3, and 2 reactions that were not in Yeast8, respectively.

**Multiphase multiobjective flux balance analysis framework.** Our results showed that batch fermentation modeling should be divided into five phases in which cellular objectives and flux constraints need to be modified: lag phase, exponential growth, growth under nitrogen limitation, stationary, and decay (Fig. 1A and B sketch the modeling approach). Their duration is imposed by the estimated parameters $t_1$, $t_2$, $t_3$, and $t_4$, illustrated in Fig. 1B.

Once inoculated, cells encounter new nutrients and undergo a temporary period of nonreplication, the lag phase, during which we assumed that ATP production is maximized. The exponential growth phase covers only the first hours until nitrogen exhaustion. In this phase, cells maximize growth. During growth under nitrogen limitation cells, still maximizing growth, accumulate carbohydrates. Thenceforward, a substantial fraction of the sugar is consumed during the stationary and decay phases by quiescent cells, which adjust their metabolism to cope with environmental fluctuations. In the last two phases, we assumed that cells maximize both ATP and protein production.

The general formulation of the FBA problem reads as follows:

$$ \text{maximize} \quad J_p $$
$$ \text{subject to} \quad S \cdot \nu = 0 $$
$$ v_{NH_4} \geq f_{NH_4}(NH_4) $$
$$ v_{AA} \geq f_{AA}(AA_i); \quad \forall i = 1, \ldots, 20 $$
$$ v_{Glx} = f_{Glx}(Glx, E) $$
$$ v_F = f_F(F, E) $$
$$ v_O_2 = f_O_2(O_2); \quad \forall j = 1, \ldots, 20 $$
$$ v_{Pj} = f_{Pj}(v_{Glx}, v_F); $$

where $J_p$ is the function to be maximized in each phase $p$, $S$ is the stoichiometric matrix, $\nu$ is the vector of fluxes in millimoles per gram (dry weight [DW]) per hour, $v_{Glx}$ and $v_F$ are the fluxes of glucose and fructose, $v_{O_2}$ is the flux of $O_2$ present only at the beginning of the fermentation, $v_{NH_4}$ is the flux of ammonium, $v_{AA}$ is the exchange rate of the amino acid $i$ (covering all 20 amino acids), and $v_{Pj}$ are the constraints associated with the $j = 1, \ldots, 20$ fermentation products considered. $Glx$, $F$, $NH_4$, $AA$, and $P_j$ correspond to the concentrations of glucose, fructose, ammonium, amino acids, and products, all expressed in millimoles per liter. Table S2 presents the specific formulation for each phase.

The uptake of glucose and fructose was modeled using Michaelis-Menten (MM)-type kinetics with competitive ethanol inhibition (14):

$$ v_{Glx} = -v_{max} \cdot \frac{Glx}{Glx + k_G} \cdot \frac{1}{1 + E/k_E} $$

where $v_{max}$ is the maximum uptake rate, $k_G$ is the MM constant, $k_E$ is the strength of ethanol inhibitory effect, and $E$ is its concentration (mmol/liter). A similar expression, $v_F$, exists for fructose ($F$). Additionally, in our case studies, the medium was supplemented with sucrose; thus, we included a mass action-type expression, characterized by the kinetic constant $k_{hydro}$, describing its hydrolysis.

A certain amount of dissolved oxygen is present in the medium and consumed during the lag phase (see Fig. S3B in Text S2). Its uptake is determined as follows:

$$ v_{O_2} \geq -k_{O_2} \cdot O_2 $$

where $v_{O_2}$ and $k_{O_2}$ are the oxygen uptake and transport rate constants and $O_2$ is the concentration of oxygen in the medium.
The uptake of ammonium was modeled by

$$v_{NH_4} = -v_{max_{NH_4}} \frac{NH_4}{NH_4 + k_{NH_4}}$$

(3)

where $NH_4$ is the extracellular concentration of ammonia (in millimoles per liter), $v_{max_{NH_4}}$ is the maximum uptake rate achieved, and $k_{NH_4}$ is the MM constant. To avoid an excessive number of parameters, amino acid transport was modeled following mass action kinetics.
where $AA_i$ is the extracellular concentration of the amino acid (in millimoles per liter) and $k_{AA_i}$ is the associated kinetic parameter.

Production of alcohols and higher alcohols, carboxylic acids, and esters follows mass action kinetics:

$$P_i = X_A \cdot v_{P_i}$$

with $X_A$ the active biomass and the flux, $v_{P_i}$, proportional to the amount of transported hexoses:

$$v_{P_i} = -k_{P_i} \cdot (v_{\text{Glx}} + v_{\text{F}})$$

where $P_i$ refers to the excreted product $i = 1, \ldots, 20$ and $k_{P_i}$ the production rates.

An exception to this was the formulation of the dynamics of acetate. This metabolite is produced during exponential growth and consumed during stationary phase following mass action kinetics.

**Model of protein turnover.** Since nitrogen sources are depleted before the stationary phase, we developed a new model of nitrogen homeostasis that considered protein turnover. The proposed model describes the combined use of the Ehrlich and de novo synthesis pathways during stationary and decay phases to guarantee optimal adaptation to perturbations in nitrogen homeostasis.

To introduce protein turnover, we simulated the degradation of the existing protein fraction inside biomass ($Prot$), into a pool of amino acids that subsequently produce new proteins. During stationary and decay phases, the lower bounds on the amino acid uptake are set as

$$v_{AA_i} \geq -\lambda \cdot Prot \cdot \alpha_{AA_i}$$

where $\lambda$ is the turnover rate, Prot is the concentration of protein, and $\alpha_{AA_i}$ is associated with the stoichiometric coefficient of the amino acid $i$ in the protein pseudoreaction. Mathematically, the dynamics of protein content reads as follows:

$$\frac{dProt}{dt} = \mu \cdot X_v \cdot Prot_{\text{content}} - Prot \cdot k_{\text{death}} + X_A \cdot v_{Prot} - \lambda \cdot Prot$$

where $\mu$ is the growth rate, Prot$_{\text{content}}$ is the fraction of protein in the newly formed biomass, $k_{\text{death}}$ (per hour) is the rate of biomass degradation during decay phase, $X_v$ is the simulated viable biomass (grams per liter), $X_A$ is the simulated active biomass (grams per liter), $v_{Prot}$ is the protein production rate (grams per gram [DW] per hour), and $\lambda$ is the protein turnover rate (per hour).

The degraded proteins ($\lambda \cdot Prot$) are converted into extracellular amino acids whose concentrations are represented by the following equations:

$$\frac{dAA_i}{dt} = X \cdot v_{AA_i} + \lambda \cdot Prot \cdot \alpha_{AA_i}$$

The rates $v_{AA_i}$ are computed by maximizing protein production ($v_{Prot}$) and ATP ($v_{\text{ATP}}$) while solving the FBA problem:

$$\text{maximize} \quad v_{Prot} \cdot \phi \cdot v_{\text{ATP}}$$

subject to

$$S \cdot v = 0$$

$$v_{AA_i} \leq \lambda \cdot Prot \cdot \alpha_{AA_i}$$

$$v_{Prot} = (v_{\text{Glx}} + v_{\text{F}}) \cdot \alpha_{Prot}$$

$$\cdots$$

where $\phi$ is estimated for each strain, $S$ is the stoichiometric matrix, $v$ is the vector of fluxes, $v_{AA_i}$ is the exchange rate of amino acid $AA_i$, $v_{P_i}$ are constraints associated with
fermentation products, and $v_{\text{Pro}}$ (r_1904) is the amount of excreted proline. The last amino acid accumulates in the extracellular medium throughout the fermentation (see Fig. S3 in Text S2), likely as a consequence of stored arginine consumption under anaerobic conditions (27). The extracellular dynamics of proline is described as follows:

$$\frac{d\text{Pro}}{dt} = X_A \cdot v_{\text{Pro}} \tag{10}$$

Depending on the kinetic constraints, amino acids can be directly incorporated into proteins or degraded to recover nitrogen for protein production. We observed that those amino acids which lacked pathways for their catabolism or elimination accumulated in the extracellular compartment during stationary phase. As an example, Fig. 1C1 shows how including a catabolic route for methionine, during stationary phase, successfully describes its dynamics, possibly indicating that this route could be active during the stationary phase.

**Dynamic biomass equation.** A static biomass equation was not able to explain nitrogen assimilation (Fig. 1C2). Therefore, we implemented the following dynamic biomass equation (15):

$$\text{Prot}_{\text{content}} = A \cdot (1 - e^{B \cdot \text{YAN}}) \tag{11}$$

where $A$ and $B$ are estimated parameters and yeast assimilable nitrogen (YAN) accounts for the ammonium and free amino acids present in the medium, excluding proline, which is not catabolized under anaerobic conditions.

Furthermore, we assumed that mRNA level was proportional to the protein content ($\text{mRNA} = \text{Prot} / \text{RNA} - \text{to} \text{Prot}$ ratio). In this framework, carbohydrates compensate for the variation in protein and mRNA content. Growth-associated ATP maintenance (GAM) was also updated to account for the polymerization costs of the different macromolecules (protein, RNA, DNA, and carbohydrates):

$$\text{GAM} = \text{GAM}_{\text{fitted}} + \text{GAM}_\text{Prot} + \text{GAM}_\text{RNA} + \text{GAM}_\text{Carbs} + \text{GAM}_\text{DNA} \tag{12}$$

where $\text{GAM}_{\text{fitted}}$ is a species- or strain-dependent parameter estimated from data and the rest are polymerization costs of the different biomass precursors (adapted from reference 18). Additionally, to represent the premature end of fermentations during the decay phase (observed in SuCET12600), we estimated the non-growth-associated maintenance (NGAM).

In addition, we discriminated between active cells (able to ferment) and viable cells (able to divide and ferment) to capture the dynamics of CFU and biomass (Fig. 1C2). The dynamics of active cell mass is represented by the equation

$$X_A = \mu \cdot X_V - X_A \cdot k_{\text{decay}} + X_A \cdot v_{\text{Prot}} - \lambda \cdot \text{Prot} \tag{13}$$

where $X_A$ is the active cell mass (grams per liter), $\mu$ is the growth rate computed with pFBA, $k_{\text{decay}}$ is the decay rate (only active during decay), $v_{\text{Prot}}$ (r_4047) is the exchange flux for protein production, and $\lambda$ is the turnover rate (both only active during stationary and decay phases). The behavior of viable cell mass differed from that of active cell mass by a decline induced by ethanol (28):

$$X_V = X_A - X_V \cdot k_{\text{death}} \cdot \frac{E^n}{E^n + k^n} \tag{14}$$

where $X_V$ (grams per liter) is the growth rate obtained with the constraint based model, $E$ is the ethanol concentration (millimoles per liter), and $k_{\text{death}}$, $n$, and $k$ are the parameters controlling susceptibility to ethanol.

The former mechanisms, coupled to parameter estimation, allowed us to predict nitrogen consumption, CFU, and biomass dynamics accurately (Fig. 1C3 to Fig. 1C6).

**Goodness of fit of the model in the case studies.** The final model consisted of 46 ordinary differential equations depending on 66 parameters which we estimated from...
Dynamic flux ratios contributing to redox balance

**Sc T73**

**Su BMV58**

**Su CECT12600**

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Hill coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCA</td>
<td></td>
</tr>
<tr>
<td>Glycolysis</td>
<td></td>
</tr>
<tr>
<td>GABA</td>
<td></td>
</tr>
</tbody>
</table>

**Reactions consuming NADH**

- **TCA**
- **Glycolysis**
- **GABA**

**Reactions producing NADPH**

- **TCA**
- **PPP**
- **GABA**

**FIG 2** Comparative study of the fluxes through the reactions consuming and producing NADH and NADPH at the stationary phase, illustrating how strains achieve redox balance. The most significant differences between strains are found at the level of the TCA cycle, the GABA shunt, the pentose phosphate pathway, the biosynthesis of aromatic amino acids which eventually lead to production of 2-phenylethanol (PEA) and the Ehrlich pathway toward production of isoamylol. It is also important to note that *S. uvarum* diverts flux to the production of mevalonate.

Time series data for all measured external metabolites and biomass. The mean standard deviation on the parameters ranges from 2.5% for SuCECT12600 and 12.6% for SuBMV58. The reasonably low distribution on the parameters resulted in a reasonably low uncertainty associated with the model simulations (as seen in Fig. 2, 3, and 4).

The model described the dynamics of our illustrative examples successfully. The best fit to the data and the associated uncertainty, as computed by the bootstrap, are shown in Text S2 and Fig. 3 and 4. We determined the $R^2$-squared measure of goodness.
Central carbon metabolism

Sc T73
Su BMV8
Su CECT12600

Lines: model
Symbols: data

A) Glucose (g/L)
B) Fructose (g/L)

C) Acetate (g/L)
D) Succinate (g/L)
E) Ethanol (g/L)
F) 2,3-Butanediol (g/L)
G) Glycerol (g/L)

FIG 3 Redox balance in central carbon metabolism. Panels A to F show model predictions versus the experimental data extracellular metabolite concentrations associated with glycolysis and central carbon metabolism for the three strains. Panel H presents the predicted intracellular dynamic flux ratios during the stationary phase, showing how *S. uvarum* and *S. cerevisiae* strains use different redox balance strategies. These differences result in the differential production of relevant external metabolites such as acetate (C), succinate (D), ethanol (E), 2,3-butanediol (F), and glycerol (G). Explicit differences in the pathways in gray are presented in Text S3; otherwise, as indicated in the legend, width of the lines is proportional to the dynamic flux ratio.
of fit ($R^2$) for each measured variable and each strain-based fermentation. The medians of the $R^2$ values are above 0.94 for all strains.

Interested readers may find further details on the parameter estimation in Text S2. Optimal parameter values and $R^2$ values are reported in Table S3.

**Species behavior differs significantly in the stationary phase.** At the extracellular level, the most striking differences between strains occur in the production of compounds associated with the central carbon metabolism and nitrogen metabolism, in particular, in the dynamics of acetate and the yields of succinate, 2,3-butanediol, and glycerol (Fig. 3C, D, F, and G) and in the production of 2-phenylethanol and isoamyl alcohol (Fig. 4B and D).
We used the model to decipher the metabolic strategies used by the different strains that lead to such differences. Table S4 reports the dynamic flux ratios computed for the overall process and the different phases using equations 21 to 23 for those reactions in which the maximum flux ratio value over the three species is above 0.01 mmol/mmolH (where mmolH is millimoles of consumed hexose × 100). Uncertainties associated with the fluxes due to uncertainties on the parameters are also reported.

Text S3 summarizes the differences observed in flux ratios between species and phases. Again, the intracellular behavior differs significantly in the stationary phase. Thus, subsequent sections elaborate on the metabolic study of the stationary phase.

The GABA shunt as an NADPH source in cryotolerant species. Cells produced the most significant fraction of succinate during the stationary and decay phases, with a significantly higher dynamic flux ratio by SuBMV58 and SuCECT12600 (6.08 and 1.66) than ScT73 (0.42, r_2057).

During the decay phase, most succinate was produced through the tricarboxylic acid (TCA) cycle reductive branch in the two species. However, during the stationary phase, succinate production was distributed between the GABA shunt (ScT73, 0.42; SuBMV58, 2.00; SuCECT12600, 1.31; Fig. 3H, r_1023) and the reductive branch of TCA (ScT73, 0.00; SuBMV58, 4.10; and SuCECT12600, 0.35; Fig. 3H, r_1000). Remarkably, succinate production through the GABA shunt was between 3 and 4.6 times higher for S. uvarum species than for S. cerevisiae (Fig. 3H, r_0068, r_1023). This result suggested an important role for the GABA shunt in the maintenance of the cellular redox state.

Incidentally, revisiting data from reference 29, we found a considerable accumulation of GABA (93.59-fold change) by SuCECT12600 (Table S4).

Intracellular mevalonate as a reducing equivalent in cryotolerant yeast species. The three strains produced acetate during the growth phase (ScT73, 1.079; SuBMV58, 1.145; SuCECT12600, 1.422 [Table S4, r_1106]) and until the entry into the stationary phase. Afterward, while the extracellular acetate concentration remained constant in ScT73, a decrease was observed in both S. uvarum fermentations, indicating acetate consumption. As shown in Fig. 3C, our model successfully described these phenotypes.

According to the modeling constraints, the most parsimonious explanation for this observation would have been an operative glyoxylate cycle. However, based on the repression by glucose of the key enzymes of the glyoxylate cycle (i.e., ICL1 and MLS1) and previous intracellular data (29), we decided to block this cycle and explore an alternative hypothesis. As a result, the model suggested that S. uvarum strains incorporated the acetate derivative, acetyl coenzyme A (acetyl-CoA), into mevalonate (SuBMV58, 0.19, and SuCECT12600, 0.10; r_0559 and r_0103) also consuming NADPH. In addition, mevalonate is a reducing equivalent that can be further metabolized into, for example, ergosterol (r_0127 in our reconstruction) possibly acting as storage of NADPH in cryotolerant species.

S. uvarum strains also used the carnitine shuttle system to transport acetyl-CoA into the mitochondria (SuBMV58, 0.09, and SuCECT12600, 0.38; Fig. 3H, r_0254). Inside the mitochondria, acetyl-CoA was used to form isopropylmalate (Fig. 4A)—a precursor of leucine and isoamyl alcohol—or in the TCA oxidative branch toward the synthesis of 2-oxoglutarate (Fig. 3H).

The production of higher alcohols contributed to the redox balance. Higher-alcohol production was most prominent during the stationary phase for the three strains. Our model predicted that carbon skeletons of isoamyl alcohol, 2-phenylethanol (PEA), and isobutanol were in great part synthesized de novo from glycolytic and pentose phosphate pathway (PPP) intermediates, rather than coming from the catabolism of precursor amino acids (leucine, valine, and phenylalanine, respectively) (Fig. 4A).

S. uvarum strains produced more PEA than ScT73 (ScT73, 0.158; SuBMV58, 0.659; and SuCECT12600, 0.383; r_1590), while the opposite occurred for isoamyl alcohol.
We found that the production of PEA and isoamyl alcohol contributed substantially to the redox metabolism related to glycerol accumulation. Approximately 43%, 36%, and 27% of the glycerol produced by the ScT73, SuBMV58, and SuCECT12600 strains was attributable to NADH derived from isoamyl alcohol and PEA.

The higher production of PEA observed in S. uvarum strains occurs through the higher flux through the shikimate pathway (Fig. 4A, r = 0.089) and 2-acetyllactic acid (r = 0.099). Noticeably, S. uvarum strains also contributed to acetyl-CoA using the carnitine shuttle (r = 0.254). 2-Acetyllactic acid can further be converted to 3-methyl-2-oxobutanoate, consuming one NADPH (r = 0.096), which also showed two different fates: the production of 2-isopropylmalate (leading to isoamyl alcohol; r = 0.025, r = 0.072, and r = 0.179) or valine (leading to the synthesis of isobutanol via the Ehrlich pathway; r = 1.087, r = 0.062, and r = 0.018).

Production of PEA and isoamyl alcohol also affected NADP⁺/NADPH metabolism. While the GABA shunt and oxidative PPP were the main producers of NADPH, consumption of NADPH was attributable mostly to isoamyl alcohol and PEA. In fact, the increased isocitrate dehydrogenase flux (r = 0.659) observed in ScT73 (1.646 against 0.917 and 0.727 in the S. uvarum strains) was associated with the need for shuttling NADPH into the mitochondria (r = 0.231), utilized in 3-methyl-2-oxobutanoate synthesis (precursor of the isoamyl alcohol; see Fig. 3H and Fig. 4A). On the other hand, production of PEA without the oxidative PPP (predicted in the S. uvarum strains) resulted in excess NADP⁺. In the case of S. uvarum, given the reduced influence of oxidative PPP, NADP⁺ recycling was achieved mostly through the GABA shunt.

**DISCUSSION**

Genome-scale models have the potential to decipher how nonconventional yeast species use metabolism to produce industrially relevant products and tolerate specific stressors, such as cold temperatures. This study aimed to develop a dynamic genome-scale model to investigate the dynamics of yeasts primary and secondary metabolism in batch cultures. To generate biological hypotheses for modeling, we considered the description of the metabolism of S. cerevisiae and cryotolerant S. uvarum strains in a rich medium (grape must) fermentation.

The first question in this research was how to model all phases in the batch process: lag, exponential growth, limited nitrogen growth, stationary, and decay. Prior studies focused on the exponential growth phase and were based on available reconstructions with many missing reactions, particularly those related to secondary metabolism (14, 15, 17, 30). Also, their static nature hinders the description of the sequential nature of amino acid consumption (31).

As a first step, we needed to extend a yeast genome-scale reconstruction to account for the production of higher alcohols, carboxylic acids, or esters. We extended the Yeast8 consensus model incorporating missing reactions and metabolites. A similar curation process has been recently applied to the iMM904 reconstruction (32). The authors fitted the model to data from the literature, concluding that further curation and adaptations were necessary to successfully predict metabolism and biomass dynamics. We experienced such difficulties in our first iterations in the modeling process and introduced several new features to obtain more accurate simulations of carbon and nitrogen metabolism throughout time.

A critical aspect for an improved accuracy was the multiphase multiobjective dynamic FBA scheme. Previous works focused on ATP consumption to explain the metabolism after depletion of the limiting nutrient (15, 33, 34); we incorporated the production of protein as a cellular objective (together with protein degradation) with
accurate results. Also, we modeled protein turnover to account for the uptake of amino acids and inorganic nitrogen in the stationary phase. To the best of our knowledge, this is the first dynamic genome-scale metabolic model describing nitrogen homeostasis during the stationary phase. Finally, we introduced a dynamic biomass equation which further improved the model accuracy. This result agrees with observations in previous studies (35–37) pointing out the relevance of detailing biomass composition in a context-specific manner.

The second question in this research was to decipher the differences in the metabolism of three strains of two different species, *S. cerevisiae* and *S. uvarum*, in rich medium (grape must) fermentation. Recently, it was hypothesized that *S. cerevisiae* and *S. uvarum* species might have different redox balance strategies (38). Notably, the model confirmed this hypothesis and brought novel insights into the specific routes used by the two species.

Our predictions suggest alternative pathways for cryotolerant species to produce succinate and consume acetate. In principle, yeasts might form succinate via four main pathways, all based on the reactions of the TCA cycle (39). The selected pathway depends on the environmental conditions and strain. Our model predicted that ScT73 and SuBMW58 produced overall the most succinate via the TCA reductive branch, in agreement with previous findings (40). However, our results also suggest an important role for the TCA oxidative branch until 2-oxoglutarate for the *S. uvarum* strains during the stationary phase. This result is consistent with the recent intracellular data obtained by Minebois et al. (25), who observed a noticeable intracellular accumulation of 2-oxoglutarate in SuBMV58.

One somewhat unexpected finding of the model was the extent to which the GABA shunt would contribute to succinate formation during the stationary phase, an effect particularly evident in the case of *S. uvarum* strains. The role of this pathway in yeast is not fully understood (41–43). Bach et al. (41) observed that glutamate decarboxylase (encoded by *GAD1*) was poorly expressed when succinate was produced in *S. cerevisiae* and that the GABA shunt played a minor role in redox metabolism. In contrast, Coleman et al. (44) showed that *GAD1* expression is required for oxidative stress tolerance in *S. cerevisiae*. Similarly, Cao et al. (42) showed that *GAD1* confers resistance to a heat stress effect that might be related to NADPH production. Additionally, *GAD1* was upregulated during the stationary phase under nitrogen starvation (41, 43, 45), and López-Malo et al. (29) observed high intracellular GABA levels in the cryotolerant species *S. uvarum*.

Recently, Liu et al. (46) found clear indication that the GABA shunt may be involved in supplying NADPH for lipid synthesis in the oleaginous yeast *Yarrowia lipolytica*. Also, Bach et al. (41) showed that *S. cerevisiae* can degrade GABA into succinate or γ-hydroxybutyric acid (GHB) and that GHB was used to form the polymer polyhydroxybutyrate (PHB). Also noteworthy is that PHBs are synthesized by numerous bacteria as carbon and energy storage compounds (47). PHBs are also strongly associated with bacterial cold tolerance (48), suggesting a similar function in yeast.

Another important finding is that *S. uvarum* strains consume acetate once nitrogen sources are depleted, coinciding with the extracellular accumulation of succinate. This finding was also reported by Kelly et al. (49), who showed that an *S. uvarum* yeast isolate can metabolize acetate, resulting in lower acetic acid, ethyl acetate, and acetaldehyde concentrations in wine. The model predicted that some of the acetate carbon was directed toward mevalonate, which is in line with recent experimental work by Minebois et al. (25). According to Bach et al. (41), a route for acetyl-CoA incorporation into PHB polyester through 3-hydroxybutyrate-CoA seems plausible. However, this hypothesis is not taken into account by the genome-scale reconstructions.

The fact that López-Malo et al. (29) found high intracellular GABA and GHB in cryotolerant strains grown in synthetic must (without GABA) at low temperature, and the flux predicted in the present work, suggests that *S. uvarum* stores lipids or polyesters (i.e., PHBs) as reducing equivalents to withstand oxidative stress induced by low
temperatures. The role of the GABA shunt and the production of reducing equivalents in the metabolism of cryotolerant species may be plausible routes worth exploring.

Our model also predicted that the carbon skeletons of higher alcohols (e.g., isobutanol and isoamyl alcohol) were mainly synthesized de novo rather than from the incorporation and catabolism of amino acids (e.g., leucine and valine). This result agrees with the findings of Crépin et al. (50), who explored the fate of the carbon backbones of aroma-related exogenous amino acids using 13C isotopic tracer experiments. Similarly, our results indicate that 2-phenylethanol was mostly synthesized de novo. We hypothesize that a positive contribution in glycerol content may also explain why the production of 2-phenylethanol and isoamyl alcohol is a conserved evolutionary trait in yeasts.

Noticeably, most of NADPH consumption was associated with higher-alcohol synthesis. The increased flux through cytosolic isocitrate dehydrogenase associated with isoamyl alcohol is compatible with reports associating high expression of IDP2 with nitrogen deficiency (51) and stationary phase (52). Furthermore, the prediction that during the stationary phase, PEA is produced through the chorismate synthesis pathway (downstream of the nonoxidative PPP), thus regenerating some NADPH associated with the conversion of succinic semialdehyde into succinate, provides a counterintuitive rationale for understanding the correlation between the observed differences in higher-alcohol and succinate production between S. cerevisiae and S. uvarum.

Interestingly, the model predicts that most other higher alcohols (tyrosol, methionol, etc.) accumulate in small amounts due to perturbations in the amino acid pool. These results confirm the hypothesis raised by Shopska et al. (53), who suggested that the two schemes to produce higher alcohols—the Ehrlich pathway and de novo synthesis—are not in contradiction but two extremes of a common mechanism. Incidentally, Yuan et al. (54) showed that the assembled leucine biosynthetic pathway coupled with the Ehrlich degradation pathway results in high-level production of isoamyl alcohol.

Our bootstrap-based identifiability analysis of the proposed model and the fact that model predictions are consistent with numerous previous findings lead us to conclude that the present model (along with the provided code) can simulate yeast metabolism in batch culture in a general chemically characterized medium; the only requirement would be to update the metabolic reconstruction if required for the specific yeast species. The model can also be used to explore and engineer novel metabolic pathways toward specific bioproducts.

MATERIALS AND METHODS

Yeast strains. In this study, three yeast strains belonging to S. cerevisiae and S. uvarum were used: the commercial strain T73 (Lalvin T73 from Lallemand Montreal, Canada), originally isolated from wine in Alicante, Spain (55), was selected as our wine S. cerevisiae (ScT73) representative; the commercial strain BMV58 (SubMV58, Velluto BMV58 from Lallemand Montreal, Canada), originally isolated from wine in Utiel-Requena (Spain), and the noncommercial CECT12600 strain, isolated from a nonfermentative environment (SuCECT12600, Alicante, Spain), represented S. uvarum.

Fermentation experiments. Fermentation assays were performed with grape must obtained from Merseguera white grapes, collected in the 2015 vintage in Titaguas (Spain) and stored in several small frozen volumes (4 liters at 20°C). Before its use, the must was clarified by sedimentation for 24 h at 4°C and sterilized by adding dimethyl dicarbonate at 1 ml liter−1. All fermentations were performed in 3 independent biological replicates in 500-ml controlled bioreactors (MiniBio; Applikon, The Netherlands), and effluent carbon dioxide level (INNOVA 1316 multigas monitors; LumaSense Technologies). Data were integrated into the BioExpert software tools (Applikon, The Netherlands). The fermentation was complete when a constant sugar content was reached as measured by high-performance liquid chromatography (HPLC).

Sampling and quantification of extracellular metabolites. Extracellular metabolites, including sugars, organic acids, main fermentative by-products, and yeast assimilable nitrogen (YAN), were determined at 10 sampling times during the fermentation. Residual sugars (glucose and fructose), organic
Metabolite concentrations were quantified by calibration graphs \((R^2 \text{ value} > 0.99)\) of the previously obtained standards from a linear curve fit of the peak areas using 10 standard mixtures.

Determination of yeast assimilable nitrogen in the form of amino acids and ammonia was carried out following the same protocol as described previously \((56)\). A volume of supernatant was removed from the fermentor, and amino acids and ammonia were separated with an ultraperformance liquid chromatograph (UPLC; Dionex Ultimate 3000). Thermo Fisher Scientific, Waltham, MA) equipped with a Kinetex 2.6-μm C18 100A column (Phenomenex, Torrance, CA) and Accurel C18, 10-μm 4.6-mm, 2.6-μm Defender guards (Thermo Fisher Scientific, Waltham, MA). For derivatization, 400 μl of the sample was mixed with 430 μl of borate buffer (1 M, pH 10.2), 300 μl of absolute methanol, and 12 μl of diethyl ethoxymethylenemalonate (DEEMM) and ultrasonicated for 30 min at 20°C. The ultrasonicated sample was incubated up at 80°C for 2 h to allow the complete degradation of excess DEEMM. Once the derivatization finished, the sample was filtered with a 0.22-μm filter before injection. The target compounds in the sample were then identified and quantified according to the retention times, UV-Vis spectral characteristics, and calibration curves \((R^2 \text{ value} > 0.99)\) of the derivatives of the corresponding standards. Amino acid standard (reference no. AAS18), asparagine, and glutamine purchased from Sigma-Aldrich were used for calibration.

**Higher alcohols and esters.** We also determined the concentrations of higher alcohols and esters for each sampling time. Volatile-compound extraction and gas chromatography (GC) were performed following the protocol described in reference \(71\). Extraction was performed using headspace solid-phase-microextraction sampling (SPME) with polydimethylsiloxane (PDMS) fibers \((Supelco; Sigma-Aldrich, Barcelona, Spain)\). Aroma compounds were separated by GC in a Thermo TRACE GC Ultra chromatograph \((Thermo Fisher Scientific, Waltham, MA)\) equipped with a flame ionization detector (FID), using an HP-INNOWAX 30-m by 0.25-mm capillary column coated with a 0.25-mm layer of cross-linked polyethylene glycol \((Agilent Technologies, CA)\). Helium was the carrier gas used \((flow, 1 \text{ ml min}^{-1})\). The oven temperature program was as follows: 5 min at 60°C, 5°C min \(^{-1}\) to 190°C, 20°C min \(^{-1}\) to 250°C, and 2 min at 250°C. The detector temperature was 280°C, and the injector temperature was 220°C under splitless conditions. The internal standard was 2-heptanone \((0.05\% \text{ wt/vol})\). Volatile compounds were identified by the retention time for reference compounds. The quantification of the volatile compounds was determined using the calibration graphs of the corresponding standard volatile compounds.

**Physiological and biomass parameters.** Physiological and biomass parameters, including OD\text{\textsubscript{600}}, dry weight (DW), CFU, and average cell diameter (ACD), were determined at each sample time, provided that the cell sample was sufficient to perform the corresponding measure. DW determination was performed by centrifuging 2 ml of the fresh sample placed in a preweighed Eppendorf tube in a MiniSpin centrifuge \((Eppendorf, Spain)\) at maximum speed \((13,200 \text{ rpm})\) for 3 min. After centrifugation, the supernatant was carefully removed, and the pellet was washed with 70% \((\text{vol/vol})\) ethanol and centrifuged under the same conditions. After washing, the aqueous supernatant was removed carefully and the tube placed in a 65°C oven for 72 h. DW was finally obtained by measuring the mass weight difference of the tube with a BP121S analytical balance \((Sartorius, Goettingen, Germany)\). OD\text{\textsubscript{600}} was measured at each sampling time using a diluted volume of sample and a BioPhotometer spectrophotometer \((Eppendorf, Germany)\). CFU were determined using 100 to 200 μl of a diluted volume of samples plated in YPD solid medium \((2\% \text{ glucose}, 2\% \text{ agar}, 0.5\% \text{ peptone}, 0.5\% \text{ yeast extract})\) and incubated for 2 days at 25°C. The resulting colonies were counted with a Comecta S.A. colony counter. Only plates with a CFU count between 30 and 300 were used to calculate the CFU of the original sample. For ACD determination, a volume of cell sample was diluted into a phosphate-buffered saline solution and cell diameter measured using a Scepter handheld automated cell counter equipped with a 40-μm sensor \((Millipore, Billerica, MA)\).

**Orthology analysis and genome-scale metabolic reconstruction.** Genomes of \textit{Sc}T73 \((57)\) and \textit{Su}BMV58 and \textit{Su}CCT21600 \((PRJNA471597)\) were used for subsequent analyses.

**Flux balance analysis.** Flux balance analysis \((FBA)\) \((60, 61)\) is a modeling framework based on knowledge of stoichiometry and mass/charge balances. The framework relies on the pseudo steady-state assumption that fluxes can only change slowly on the time scale considered. A stoichiometric matrix is constructed that describes the complete transformation of metabolites into and out of the intracellular compartment. The FBA model is then solved subject to constraints that ensure that the cell can grow and maintain its size. The resulting solutions provide information about how the cell allocates its available resources to meet its metabolic and biomass needs.
state assumption (no intracellular accumulation of metabolites occurs). This is captured by the well-known expression

\[ S \cdot v = 0 \]  

(15)

where \( S \) is stoichiometric matrix of \( n \) metabolites by \( m \) reactions and \( v \) is a vector of metabolic fluxes. The number of unknown fluxes is higher than the number of equations, and thus, the system is underdetermined. Still, it is possible to find a unique solution under the assumption that cell metabolism evolves to pursue a predetermined goal which is defined as the maximization (or minimization) of a certain objective function \( J \):

\[
\begin{align*}
\text{max} \quad & J \\
\text{s.t.} : \quad & S \cdot v = 0 \\
& \text{LB} < v < \text{UB}
\end{align*}
\]

(16, 17, 18, 19)

where \( \text{LB} \) and \( \text{UB} \) correspond to the lower and upper bounds on the estimated fluxes. Examples of objective functions \( J \) include growth rate, ATP, and the negative of nutrient consumption, etc.

Typically, multiple optimal solutions exist for a given FBA problem. In parsimonious FBA (pFBA), the result is the most parsimonious of optimal solutions, i.e., the solution that achieves the specific objective with the minimal use of gene products and the minimization of the total flux load (62).

**Parameter estimation.** The aim of parameter estimation is to compute the unknown parameters—growth-related constants and kinetic parameters—that minimize some measure of the distance between the data and the model predictions. The maximum likelihood principle yields an appropriate measure of such distance (63):

\[
J_{\text{mc}}(\theta) = \sum_{k=1}^{n_{\text{exp}}} \sum_{j=1}^{n_{\text{obs}}} \sum_{i=1}^{n_{\text{st}}} \left( \frac{y_{jkm}(\theta) - y_{ijk}^\text{obs}}{\sigma_{ijk}} \right)^2
\]

(20)

where \( n_{\text{exp}}, n_{\text{obs}}, \) and \( n_{\text{st}} \) are, respectively, the number of experiments, observables (measured quantities), and sampling times, while \( \sigma_{ijk} \) represents the standard deviation of the measured data as obtained from the experimental replicates; \( y_{ijk}^\text{obs} \) represents each of the measured quantities, \( x^\text{m} \) and \( C^\text{m} \) in our case, and \( y(\theta) \) corresponds to model predicted values, \( \hat{x} \) and \( \hat{C} \). Observation functions were included for CFU and OD\text{600} in order to scale viable cell mass (\( \hat{x} \)) and active cell mass (\( \hat{C} \)), respectively.

Parameters are estimated by solving a nonlinear optimization problem where the aim is to find the unknown parameter values \( \theta \) to minimize \( J_{\text{mc}}(\theta) \), subject to the system dynamics—the model—and parameter bounds (64).

**Uncertainty analysis.** In practice, the value of the parameters \( \theta \) compatible with noisy experimental data is not unique; i.e., parameters are affected by some uncertainty (64). The consequence of significant parametric uncertainty is that it may impact the accuracy of model predictions.

To account for model uncertainty, we used an ensemble approach. To derive the ensemble, we apply the bootstrap smoothing technique, also known as bootstrap aggregation (the bagging method) (65, 66). The bagging method is a well-established and effective ensemble model/model averaging device that reduces the variability of unstable estimators or classifiers (66). The underlying idea is to consider a family of models with different parameter values \( \theta = [\theta_1, \ldots, \theta_N]^T \) compatible with the data \( y^\text{m} \), when using the model to predict untested experimental setups. The matrix of parameter values \( \hat{\theta} \) consistent with the data is obtained using \( N \) realizations of the data obtained by bootstrap (67). Each data realization has the same size as the complete data set, but it is constructed by sampling uniformly from all replicates (3 biological replicates per sampling time). Within each iteration, each replicate has an approximate chance of 37% of being left out, while others might appear several times. The family of solutions, \( \theta \), is then used to make \( N \) predictions (dynamic simulations) about a given experimental scenario. The median of the simulated trajectories regards the model prediction, while the distribution of the individual solutions at a given sampling time provides a measure of the uncertainty of the model.

**Analysis of dynamic metabolic fluxes.** We selected the most relevant metabolic pathways using a flux ratio, which provides a measure of the net flux over time during growth and stationary phases. In particular, we computed the integral of each flux multiplied by the biomass (millimoles per hour) over time and normalized its value with the accumulated flux of consumed hexoses (glucose and fructose):

\[
S_{\text{ic}} = 100 \times \frac{\int_0^{t_f} v_{i}(t) \cdot \text{DW}(t) \, dt}{\int_0^{t_f} v_{\text{glc}}(t) \cdot \text{DW}(t) + \int_0^{t_f} v_{\text{fru}}(t) \cdot \text{DW}(t)}
\]

(21)
where $S_{iS}$ corresponds to the score of the flux $i$ during growth, $S_{iD}$ corresponds to the score during the stationary phase, $S_{iD}$ corresponds to the score during the decay phase, $v_i(t)$ (millimoles per hour per gram dry weight) is the flux under scrutiny, $v_{G}(t)$ (millimoles per hour per gram dry weight) is the flux of glucose, $v_{F}(t)$ (millimoles per hour per gram dry weight) is the flux of fructose, and $DW$ is the predicted dry weight biomass (grams). Results correspond to millimoles of produced compound per millimole of consumed hexose $\times 100$ (mmol/100). Score values indicate the overall impact of each reaction in the net oxidation or reduction of electron carriers during the given phase of the fermentation.

**Numerical tools.** To automate the modeling pipeline, we used the AMIGO2 toolbox (68). To solve the dynamic flux balance analysis (dFBA) problem, we used a variable-step, variable-order Adams-Bashforth-Moulton method to solve the system of ordinary differential equations that describe the dynamics of the extracellular metabolites. At each time step, the pFBA problem was solved using the Bashforth-Moulton method to solve the system of ordinary differential equations that describe the dynamics of the extracellular metabolites. The global optimizer Enhanced Scatter Search (eSS (70)) was used to find the optimal parameter values in reasonable computational time.

The ensemble model generation procedure is computationally intensive. However, since each parameter estimation instance in the ensemble is an entirely independent task, we were able to solve this problem in less than a day using 60 CPU cores on a Linux cluster. These tasks were automated with the help of bash scripts and the Open Grid Scheduler. All the scripts necessary to reproduce the results are distributed (https://sites.google.com/site/amigo2toolbox/examples).

**SUPPLEMENTAL MATERIAL**

Supplemental material is available online only.

- **TEXT S1**, PDF file, 0.3 MB.
- **TEXT S2**, PDF file, 2.7 MB.
- **TEXT S3**, PDF file, 1.1 MB.
- **TABLE S1**, XLSX file, 0.01 MB.
- **TABLE S2**, XLSX file, 0.1 MB.
- **TABLE S3**, XLSX file, 0.02 MB.
- **TABLE S4**, XLSX file, 0.2 MB.

**ACKNOWLEDGMENTS**

This project has received funding from MCIU/AEI/FEDER, UE (grant references RTI2018-093744-B-C31, RTI2018-093744-B-C32, RTI2018-093744-B-C33, and PID2019-104113RB-I00) and Xunta de Galicia (IN607B 2020/03). R.M. was supported by an FPI grant from the Ministerio de Economía y Competitividad, Spain (reference BES-2016-078202). S.N.M. acknowledges funding from CONICYT Becas Chile grant 72180373. S.N.M. and B.T. acknowledge support from YogurtDesign, EraCoBioTech grant 053.80.733.

D.H. implemented the model. R.M. performed the experiments. S.N.M. performed the reconstruction of strain-specific genome-scale models. L.G.M. assembled and annotated the genome. D.H., S.N.M., and L.G.M. participated in the gap filling process of metabolic functions in the model and in the development of the biomass equations. E.B.-C., R.P.-T., A.Q., B.T., and E.B. coordinated the work and design of the study/pipeline. D.H., R.M., S.N.M., L.G.M., and E.B.-C. wrote the original manuscript. All authors read, edited, and approved the final paper.

**REFERENCES**


