Received: 1 June 2022 | Accepted: 1 January 2023

DOI: 10.1111/1751-7915.14211

MICROBIAL BIOTECHNOLOGY Applied Microbiology International

RESEARCH ARTICLE

Modelling the physiological status of yeast during wine fermentation enables the prediction of secondary metabolism

Artai R. Moimenta[1,2] | **David Henriques**[1] | **Romain Minebois**[3] | **Amparo Querol**[3] | **Eva Balsa-Canto**[1]

[1]Bioprocess and Biosystems Engineering, IIM-CSIC, Vigo, Spain

[2]Applied Mathematics II, University of Vigo, Vigo, Spain

[3]Systems Biology of Yeasts of Biotechnological Interest, IATA-CSIC, Paterna, Spain

Correspondence
David Henriques and Eva Balsa-Canto, Bioprocess and Biosystems Engineering, IIM-CSIC, Vigo, Spain.
Email: davidh@iim.csic.es and ebalsa@iim.csic.es

Funding information
Ministerio de Ciencia, Innovación y Universidades, Grant/Award Number: AEI/10.13039/501100011033, CEX2021-001189-S, PID2021-126380OB-C31, PID2021-126380OB-C32 and RTI2018-093744-B-C33; Ministerio de Economía y Competitividad, Grant/Award Number: BES-2016-078202; Xunta de Galicia, Grant/Award Number: IN607B 2020/03

Abstract

Saccharomyces non-cerevisiae yeasts are gaining momentum in wine fermentation due to their potential to reduce ethanol content and achieve attractive aroma profiles. However, the design of the fermentation process for new species requires intensive experimentation. The use of mechanistic models could automate process design, yet to date, most fermentation models have focused on primary metabolism. Therefore, these models do not provide insight into the production of secondary metabolites essential for wine quality, such as aromas. In this work, we formulate a continuous model that accounts for the physiological status of yeast, that is, exponential growth, growth under nitrogen starvation and transition to stationary or decay phases. To do so, we assumed that nitrogen starvation is associated with carbohydrate accumulation and the induction of a set of transcriptional changes associated with the stationary phase. The model accurately described the dynamics of time series data for biomass and primary and secondary metabolites obtained for various yeast species in single culture fermentations. We also used the proposed model to explore different process designs, showing how the addition of nitrogen could affect the aromatic profile of wine. This study underlines the potential of incorporating yeast physiology into batch fermentation modelling and provides a new means of automating process design.

INTRODUCTION

The winemaking sector faces new challenges. In recent years, concerns about the impact of alcohol consumption on health and road safety have increased (Poznyak & Rekve, 2015). Additionally, alcohol negatively affects the perception of the complexity of wine flavours. Therefore, the wine industry is actively seeking ways to produce wines with a lower alcohol concentration, ideally, without compromising taste and consumer's acceptance or increasing production costs (Varela et al., 2015).

Yeasts play a critical role in responding to these challenges. Winemakers count on hundreds of commercially available *Saccharomyces cerevisiae* starters that differ in aroma profiles (Scott et al., 2021). However, all of them generate similar ethanol yields, translating into comparable concentrations of ethanol in wine. Alternatives may emerge from exploring yeast biodiversity (Hittinger et al., 2018; Steensels & Verstrepen, 2014).

Cold-tolerant species of the *Saccharomyces* genus, namely, *S. kudriavzevii* and *S. uvarum*, and their hybrids with *S. cerevisiae*, have shown promising physiological properties (Pérez-Torrado et al., 2018). Recent research explored

optimal growth temperatures (López-Malo et al., 2013), resistance to ethanol (Lairón-Peris et al., 2020) or aroma production (Gamero et al., 2013; Pérez et al., 2022) in fermentations led by these species. These studies show a general decrease in ethanol production and an increase in glycerol (Querol et al., 2018). Although these results reveal great potential for these species to produce wines with a lower ethanol content, the design of the fermentation process requires intensive experimentation.

The use of mechanistic models combined with exometabolome data throughout fermentation could systematize the process design. Models can not only facilitate the selection of the right species for a particular product, but also the automatic design of common interventions such as nitrogen addition to deal with slow fermentations (Bisson & Butzke, 2000). Therefore, the possibility of predicting and optimizing wine fermentation is of utmost importance.

Recently, wine fermentation modelling has received considerable attention (see the review by Miller & Block, 2020). Many authors adopt the macroscopic scale modelling approach to model the dynamics of biomass, relevant nutrients (sugars and yeast assimilable nitrogen, YAN), and products such as ethanol (Coleman et al., 2007; Cramer et al., 2002). These models often rely on the Monod equation to describe the rate of biomass growth as a function of a limiting substrate (e.g. YAN or sugars).

However, available macroscopic models do not account for the five phases observed in batch fermentation, that is, the lag phase, exponential growth, growth under nitrogen limitation, stationary phase and decay phase (Henriques & Balsa-Canto, 2021). In addition, they generally focus on primary metabolism (hexose uptake and ethanol formation), which does not necessarily overlap with the kinetics of aroma production.

Recently, Henriques et al. (2021) suggested an ad hoc approach to account for these phases in a dynamic genome-scale model. In this formulation, different fermentation phases were associated with the production or consumption of different extracellular metabolites. The limitation of this approach is that it lacks a mechanistic connection between the limiting substrate (nitrogen) and the duration of each phase.

In the present work, we formulate a continuous model that can accurately describe fermentation phases. To this end, we extend the model proposed by Henriques and Balsa-Canto (2021) assuming that nitrogen starvation is associated not only with carbohydrate accumulation but also with the induction of a set of transcriptional changes associated with the stationary phase (Rossignol et al., 2003). The final model accounts for the dynamics of yeast biomass and primary metabolites (hexoses, YAN and ethanol) and the production of secondary metabolism compounds (higher alcohols, acetate esters, etc.). The model was calibrated using experimental data obtained during natural must fermentation by four cold-tolerant *Saccharomyces* strains in individual cultures: *S. cerevisiae* (ScT73), *S. kudriavzevii* (SkCR85) and two strains of the species *S. uvarum* (SuBMV58 and SuCECT12600) (Minebois et al., 2020). The calibrated model successfully explained the data, showing its ability to automatically recover all fermentation phases while accurately capturing the dynamics of aroma production. Remarkably, the model shows relevant qualitative and quantitative differences between the various strains. From a process engineering point of view, the most interesting difference found was that glycerol production and higher alcohol were modulated, at least to a certain extent, by the duration of the adaptation period to nitrogen starvation and that this period appears to be strain dependent.

We finally used the model to explore different scenarios of YAN availability and supplementation showing how final composition varies not only with supplementation but with its timing. Although these particular results require additional experimental verification, since the model accurately describes the dynamics of fermentation with four strains of three different species (*S. cerevisiae*, *S. uvarum* and *S. kudriavzevii*), we are confident in its potential to design and control alcoholic fermentation with other species.

EXPERIMENTAL PROCEDURES

Experimental data

The experimental data were obtained in a previous study by the authors and reported in Minebois et al. (2020). Here, four yeast strains belonging to the species *S. cerevisiae*, *S. uvarum* and *S. kudriavzevii* were considered: the commercial strain T73 (Lalvin T73 from Lallemand), *S. cerevisiae* (ScT73); the commercial strain BMV58 (SuBMV58, Velluto BMV58 from Lallemand), the non-commercial strain CECT12600 and the natural isolate *S. kudriavzevii* CR85 (SkCR85). The time-course experimental data were obtained from fermentation assays performed in Merseguera grape must with added sucrose in order to reach a probable alcoholic grade of 12.5%. All data were derived from three independent biological replicates in 500 ml, temperature-controlled bioreactors set at 25°C. The data-sets include dry weight (DW), hexoses, organic acids, the main fermentative by-products, yeast assimilable nitrogen (YAN) and volatile compounds at 10 sampling times during fermentation.

Mathematical methods

The model consisted of 26 ordinary differential equations depending on 36 parameters (θ). The dry weight of biomass, carbohydrate and protein, and all metabolites are collected in a vector of state variables $x(t) \in X \subset \mathbb{R}^{n_x}$, which is the unique solution of the set of ordinary nonlinear differential equations:

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}, \theta, t) \tag{1}$$

where $\dot{\boldsymbol{x}} = \frac{d\boldsymbol{x}}{dt}$, $\boldsymbol{u} \in U \subset \mathbb{R}^{n_u}$ corresponds to external factors and $\theta \in \Theta \subset \mathbb{R}^{n_u}$ is the vector of model parameters where Θ is the feasible parameter space.

Moreover, given an experiment with n_o observables and n_s^o sampling times per observable o, $\boldsymbol{y^o} \in Y \subset \mathbb{R}^{n_s^o}$ regards the vector of n_s^o discrete time measurements as follows:

$$\boldsymbol{y^o}\left(t_s^o, \boldsymbol{u}, \theta\right) = \boldsymbol{g^o}\left(\boldsymbol{x}\left(\boldsymbol{u}, \theta, t_s^o\right), \theta, t\right); \qquad s = 1, \ \ldots, n_s^o \tag{2}$$

where t_s^o refers to the sth sampling time for the observable o. Thus every experimental (measured) data point is denoted as ym_s^o and, similarly, the corresponding model predictions are denoted as y_s^o. Note that in our case, all states included in the model except the carbohydrate and protein content in the biomass, were observed during the experiments.

To identify the best model we used an iterative identification procedure (Vilas et al., 2018) which consisted of the following steps: (1) parameter estimation, to compute those parameter values that minimize the distance between model predictions and data; (2) practical identifiability analysis through the Fisher information matrix to compute confidence intervals for the parameters and to detect correlated parameters; (3) model reduction, to improve practical identifiability while ensuring goodness of fit. The procedure is implemented in the AMIGO2 toolbox (Balsa-Canto et al., 2016) devoted to the identification of biological systems models. The input files for simulation and parameter estimation are publicly available on the AMIGO2 website: https://sites.google.com/site/amigo2toolbox/examples.

Parameter estimation was formulated as a nonlinear optimization problem to find the unknown model parameters that minimize the distance between the predicted values of the model and the available data. The AMIGO2 toolbox (Balsa-Canto et al., 2016) was used to solve the parameter estimation problem. In particular, we selected the least squares approach. The problem is formulated as follows:

Find θ to minimize:

$$J(\theta) = \sum_{o=1}^{n_o} \sum_{s=1}^{n_s^o} \left[y_s^o(\theta) - ym_s^o\right]^2 \tag{3}$$

subject to the model equations and algebraic constraints defining the parameters feasible region $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$.

Model equations were solved using CVODES (Hindmarsh et al., 2005) while the global optimizer enhances Scatter Search (eSS, Egea et al., 2009) was selected to solve the parameter estimation problem due to its well-recognized efficiency and robustness.

Practical identifiability was analysed using the Fisher information matrix and the correlation matrix. The Fisher Information Matrix ($\mathcal{F}$) (Walter & Pronzato, 1997) reads as follows:

$$\mathcal{F} = \underbrace{E}_{y_m \mid \theta^*} \left\{ \left[\frac{\partial J(\theta)}{\partial \theta}\right] \left[\frac{\partial J(\theta)}{\partial \theta}\right]^T \right\} \tag{4}$$

where E represents the expectation for the optimal value of the parameters θ^*.

To compute the confidence intervals we used the Crammèr–Rao inequality, which establishes, under certain assumptions on the number of data and non-linearity of the model, that the covariance matrix may be approximated by the inverse of the Fisher information matrix ($C \leq \mathcal{F}$) (Walter & Pronzato, 1997). The confidence intervals on the parameters are defined as follows:

$$\pm t_{\alpha/2}^{\gamma} \sqrt{C_{ii}} \tag{5}$$

where $t_{\alpha/2}^{\gamma}$ is given by Student's t-distribution, $\gamma = N_d - \eta$ degrees of freedom and $(1-\alpha)\,100\,\%$ is the confidence interval selected, typically $95\,\%$.

From the covariance matrix it is also possible to compute the correlation matrix:

$$\mathrm{Cr}_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}} \quad i = 1, \ \ldots, n_\theta \ \ j = 1, \ \ldots, n_\theta \tag{6}$$

in such a way that two parameters (p_i, p_j) are completely uncorrelated if $\mathrm{Cr}_{ij} = 0$ or completely correlated if $\mathrm{Cr}_{ij} = \pm 1$.

Model reduction was achieved by fixing or removing parameters according to their influence in the model predictions, that is, those with the lowest sensitivity value. In addition, we reduced the number of parameters when a high correlation was detected between pairs of parameters. The quality of fit of reduced models was compared to that of the original model to guarantee a satisfactory goodness-of-fit.

RESULTS

Modelling biomass formation

Figure 1 presents a schematic representation of the model, the phases that we have considered during modelling and all the compounds incorporated into the model. To model biomass growth, we considered that once inoculated into the grape must, cells need a temporary period to adjust to the new environment (the

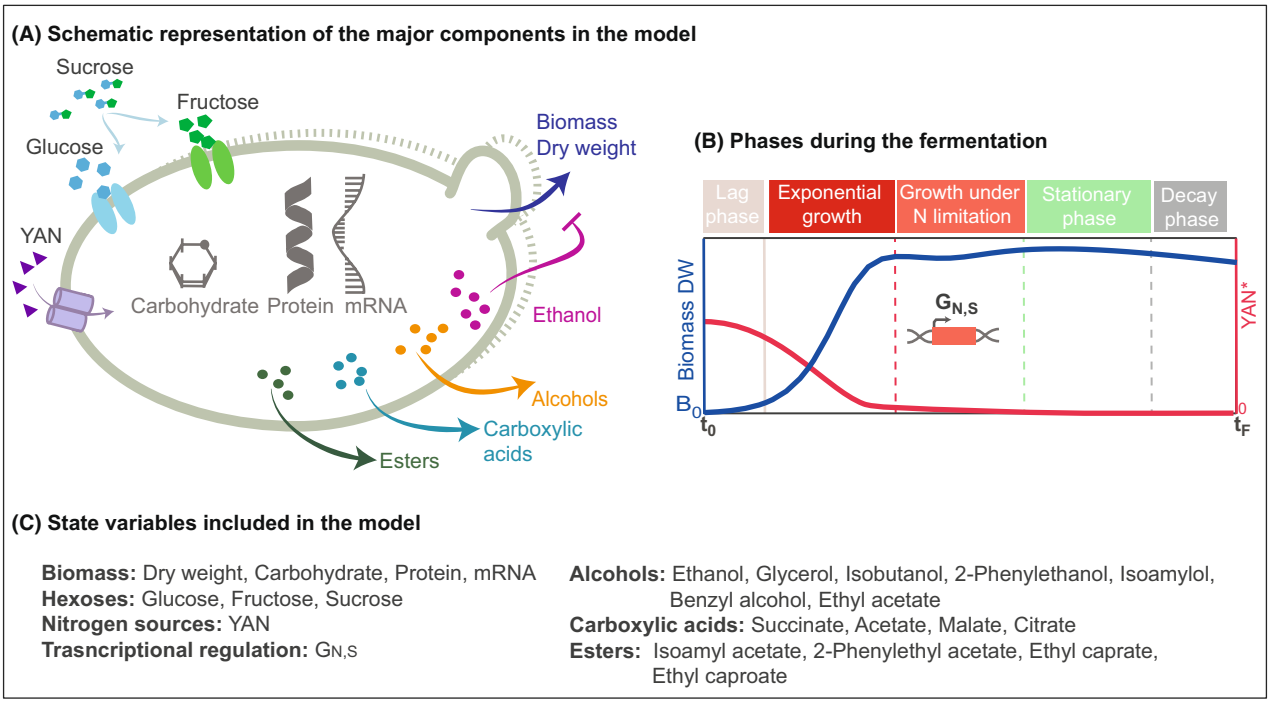

FIGURE 1 Model details. (A) shows a schematic representation of the main components included in the model; (B) shows the phases observed during fermentation: the lag phase, exponential growth, growth under nitrogen limitation, stationary phase and decay phase, which are included in the model; (C) presents all variables included in the model, including those related to biomass production and biomass composition, substrates, and products.

lag phase). Exponential growth occurs during the first hours after the lag phase, when assimilable nitrogen (YAN) is abundant. Under YAN depletion, yeasts begin to accumulate carbohydrates, producing a substantial increase in the dry weight biomass content.

The yeast growth rate was described using two different terms. The first term describes exponential growth, and this phase lasts until the assimilable nitrogen sources (YAN) are practically depleted. The second term corresponds to a secondary increase in the dry weight of the biomass as a result of the accumulation of carbohydrates, where (X) the dry weight of the biomass in (g / L), and its dynamics are as follows:

$$\frac{dX}{dt} = \text{lag}(t) \cdot \left(\mu_N + \mu_C\right) \tag{7}$$

The time-dependent function lag(t) represents the lag phase modelled here as proposed by Baranyi and Roberts (1994):

$$\text{lag}(t) = \frac{a_0}{a_0 + \left(1 - a_0\right) \cdot e^{-\mu_{maxN} \cdot t}} \tag{8}$$

where a_0 is a parameter bounded between 0 and 1 and represents the physiological state of the inoculum. If a_0 is small ($a_0 \ll 1$), then the lag phase is large. If $a_0 = 1$, then there is no lag phase.

Nitrogen sources were considered the limiting nutrient, and we used Monod kinetics to model growth associated with cell division:

$$\mu_N = X \cdot \mu_{maxN} \cdot \phi_N \tag{9}$$

where μ_{maxN} (h^{-1}) is the maximum specific growth rate and ϕ_N is the Monod equation that describes the relationship between substrate availability and growth. Here, ϕ_N is written as:

$$\phi_N = \frac{N}{N + k_N} \tag{10}$$

where N is the extracellular YAN in (g / L) and k_N (g / L) is the parameter of the Monod equation.

We also included a mechanism for carbohydrate accumulation (μ_C). This secondary growth rate (μ_C) is associated with nitrogen starvation. Therefore, μ_C reads as follows:

$$\mu_C = \phi_{N,S} \cdot X \cdot \Phi_C \cdot \left(\Theta_C - \frac{X_C}{X}\right) \tag{11}$$

where θ_C represents a target value for the ratio of carbohydrates content in biomass ($\frac{X_C}{X}$), and its impact is controlled by the value of Φ_C (h^{-1}); $\phi_{N,S}$ is a term associated with the activation of secondary growth:

$$\phi_{N,S} = 1 - \frac{N \cdot (1 + k_{sC})}{N + k_{sC}} \quad (12)$$

k_{sC} regulates the nitrogen concentration needed to induce carbohydrate accumulation and the transcriptional changes associated with nitrogen starvation.

We divided biomass into carbohydrates, proteins and mRNA. Their dynamics are described by the following ordinary differential equations:

$$\frac{dX_C}{dt} = \lambda_C \cdot \mu_N + \mu_C - k_D \cdot X_C \cdot \phi_{ED} \quad (13)$$

$$\frac{dX_P}{dt} = \lambda_P \cdot \mu_N - k_D \cdot X_P \cdot \phi_{ED} \quad (14)$$

$$\frac{dX_{mRNA}}{dt} = (1 - \lambda_C - \lambda_P) \cdot \mu_N - k_D \cdot X_{mRNA} \cdot \phi_{ED} \quad (15)$$

where λ_C and λ_P refer to the percentages of carbohydrates and proteins in biomass respectively; and $k_D \cdot \phi_{ED}$ represents the cellular decay due to the production of ethanol (Eth):

$$\phi_{ED} = \frac{Eth}{Eth + k_{Ed_i}} \quad (16)$$

Finally, we model the consumption of the YAN assuming that it is proportional to μ_N:

$$\frac{dN}{dt} = -(1 - \lambda_C) \cdot \frac{\mu_N}{Y_{x/N}} \quad (17)$$

where $Y_{x/N}$ (g_{DW}/g_N) is the yield of nitrogen to mRNA and protein biomass.

Modelling fermentation rate

The hexoses needed for fermentation come from sucrose, glucose and fructose. Sucrose is broken down into glucose and fructose. The transport of hexoses can be expressed as:

$$v_{Glx} = -\text{lag}(t) \cdot X_P \cdot v_{max,Glx} \cdot \frac{Glx}{Glx + ks_{Glx}} \cdot \phi_{Eth} \quad (18)$$

$$v_F = -\text{lag}(t) \cdot X_P \cdot v_{max,F} \cdot \frac{F}{F + ks_F} \cdot \phi_{Eth} \quad (19)$$

where v_{Glx} and v_F are the glucose and fructose uptake rates respectively; $v_{max,Glx/F}$ ($\frac{g/L}{h \cdot g_{DW}/L}$) is the maximum rate of hexose transport; and $ks_{Glx/F}$ (g/L) is the Michaelis–Menten constant. We assumed that hexose uptake is inhibited by the presence of ethanol (Hjersted et al., 2007):

$$\phi_{Eth} = \frac{1}{1 + \frac{Eth}{k_{E_i}}} \quad (20)$$

where k_{E_i} (g/L) refers to the strength of the inhibitory effect.

Taking these factors into account, the glucose and fructose dynamics are as follows:

$$\frac{dGlx}{dt} = v_{Glx} + \frac{1}{2}\frac{dSucrose}{dt} \quad (21)$$

$$\frac{dF}{dt} = v_F + \frac{1}{2}\frac{dSucrose}{dt} \quad (22)$$

where sucrose follows the mass action kinetics. Equations (15) and (16) describe the glucose and fructose dynamics. These equations account for the effect of sucrose degradation $\frac{dSucrose}{dt}$ and glucose/fructose transport into the cell $v_{Glx/F}$. Ethanol inhibition (ϕ_{Eth}) occurs in hexose transport, slowing the dynamics in the presence of high ethanol concentrations.

We assumed that the fermentation rate is proportional to the total protein content (X_P), in agreement with (Henriques & Balsa-Canto, 2021). We also assume that oxygen is not present during fermentation. The oxygen measured was lower than 0.088 g/L for all four species.

Modelling the production of extracellular metabolites

During alcoholic fermentation, glucose and fructose are metabolized through central carbon metabolism into various metabolites such as ethanol, acetate, glycerol or succinate. The excretion of these metabolites was assumed to be proportional to the consumption of hexoses. For example, ethanol production is as follows:

$$\frac{dEth}{dt} = -Y_{Eth} \cdot (v_{Glx} + v_F) \quad (23)$$

where Y_{Eth} (g_{Eth}/g_{hexose}) is the yield of ethanol production.

We previously observed (Henriques et al., 2021) that the production of higher alcohols began earlier than that of ethanol or glycerol. In addition, we observed that neither glycerol nor higher alcohols were produced until the end of fermentation. To accommodate these observations, we assumed that the production of a set of metabolites, such as higher alcohols (HA) and acetate esters (EST), was produced by low nitrogen as follows:

$$\frac{dHA}{dt} = -Y_{HA} \cdot (v_{Glx} + v_F) \cdot (1 - G_{N,S}) \quad (24)$$

$$\frac{dEST}{dt} = -Y_{EST} \cdot (v_{Glx} + v_F) \cdot (1 - G_{N,S}) \quad (25)$$

where Y_{HA} (g_{HA}/g_{hexose}) and Y_{EST} (g_{EST}/g_{hexose}) correspond to yields of a specific higher alcohol or ester respectively and $G_{N,S}$ represents the set of transcriptional changes induced by nitrogen starvation responsible for inhibiting the production of higher alcohols and glycerol. Specifically, we considered isoamyl alcohol, isobutanol, 2-phenyl ethanol (PEA) and benzyl alcohol belonging to the higher alcohol taxonomy and ethyl acetate, isobutyl acetate, isoamyl acetate, 2-phenylethyl acetate, ethyl caprate and ethyl caproate in the case of esters. $G_{N,S}$ intends to characterize the experimental observations by Rossignol et al. (2003). In their work, the authors observed that during fermentation of a synthetic medium that mimicked a natural must, growth arrest was caused by nitrogen exhaustion, and subsequent entry into the stationary phase triggered significant transcriptional reprogramming. To represent this transition using a limited number of parameters, we describe $G_{N,S}$ with the following ordinary differential equation:

$$\frac{dG_{N,S}}{dt} = (\phi_{N,S} - G_{N,S}) \cdot \tau_{G_{N,S}} \quad (26)$$

The production of succinate and acetate differs in fermentations led by *S. cerevisiae* and *S. uvarum*. Noticeably, none of these compounds could be described using the formulas described above for ethanol, glycerol, higher alcohols or acetate esters. In the case of succinate, its production was well described by a later start of production associated with nitrogen starvation and was not inhibited by the transcriptional changes associated with nitrogen starvation:

$$\frac{dSuccinate}{dt} = -Y_{Succ} \cdot \phi_{N,S} \cdot (v_{Glx} + v_F) \quad (27)$$

where Y_{Succ} ($g_{succinate}/g_{hexose}$) is the yield of succinate produced from hexose.

Regarding acetate, as noted in Henriques and Balsa-Canto (2021), *S. uvarum* produces acetate during growth and consumes it during the stationary phase. Therefore, we divided acetate metabolism into two separate fractions:

$$\frac{dAcetate}{dt} = v_{ACE,P} + v_{ACE,C} \quad (28)$$

where $v_{ACE,P}$ describes acetate production and $v_{ACE,C}$ its consumption. In our model, acetate production was inhibited by nitrogen starvation and represented by the following equation:

$$v_{ACE,P} = -Y_{ACE,P} \cdot (1 - \phi_{N,S}) \cdot (v_{Glx} + v_F) \quad (29)$$

where $Y_{ACE,P}$ ($g_{acetate}/g_{hexose}$) is the yield of acetate produced from hexose. In contrast, acetate consumption is activated by nitrogen starvation and is described by mass action:

$$v_{ACE,C} = -k_{ACE,C} \cdot X_P \cdot \phi_{N,S} \cdot Acetate \cdot (1 - G_{N,S}) \quad (30)$$

where $k_{ACE,C}$ ($h^{-1} \cdot g^{-1}_{protein}$) is the mass action parameter. To account for the fact that acetate was not fully consumed in the analysed experimental data, its consumption was inhibited by $G_{N,S}$.

Modelling carboxylic acid consumption

As shown in Figure S1, citrate and malate were consumed during growth. Here, we describe the time course trajectories of these carboxylic acids (CA_C) with mass action:

$$\frac{dCA_C}{dt} = -k_{CA_C} \cdot (1 - \phi_{N,S}) \cdot CA_C \quad (31)$$

where k_{CA_C} (h^{-1}) is the mass action parameter and $(1 - \phi_{N,S})$ enforces that consumption is only possible before nitrogen starvation.

Model calibration, identifiability analysis and model reduction

The original model consisted of 26 ordinary differential equations depending on 36 parameters. We calibrated the model to time series data of biomass, substrates and products obtained for fermentations led by *S. cerevisiae* T73 (ScT73), *S. kudriavzevii* CR85 (SkCR85), *S. uvarum* BMV58 (SuBMV58) and CECT12600 (SuCECT12600).

Next, we computed the Fisher information matrix and confidence intervals for the parameters using the Crammér–Rao inequality. We found that although the proposed model produced a very good adjustment to the experimental data, the confidence intervals (Table S1) were significantly wide because of practical identifiability issues (see the Supporting Information for further details). Using the information from the confidence intervals, we decided to reduce the model in such a way that the less influencing parameters were removed. The final structure of the model remained the same for all species.

Regarding the parameters associated with biomass production, we found that $k_{s,C}$ generally tended to be very small (often below 1×10^{-3}) and was associated with high uncertainty in all strains. Consequently, we set this parameter to 1×10^{-4} for all strains. From a physiological point of view, fixing this parameter at such a low value ensures that the effects associated with $k_{s,C}$ require nitrogen depletion. Furthermore, we

fixed parameters corresponding to biomass composition using the values reported by (Schulze et al., 1996) for nitrogen-limited batch fermentation (59% protein and 12% mRNA). Additionally, the parameter that accounts for cell decay due to ethanol (K_{Ed_i}) was highly correlated with the cell decay rate (K_D) (Figures S3–S6 and Table S2). Since both parameters showed large confidence intervals and were highly correlated, we decided to remove ϕ_{ED} from the model. Finally, with the exception of *S. kudriavzevii*, we found that the model could explain the data without the need to introduce a lag phase. Consequently, in the cases of *S. cerevisiae* and *S. uvarum* strains, we set a_0 to 1, which in practice means that there is no lag phase for these strains.

On the basis of the identifiability analysis, we were also able to eliminate terms associated with fermentation rate. In the initial model, the maximum specific uptake rate parameters ($v_{max,Glx}$ and $v_{max,F}$) and the substrate affinity parameters ($k_{s,Glx}$ and $k_{s,F}$) were highly correlated for glucose and fructose (Figures S3–S6). To address this issue, we assumed a unique maximum uptake rate for both transport rates. Similarly, the parameter that accounts for the inhibition of transport due to ethanol (k_{E_i}) and the hexose affinity parameters ($k_{s,Glx}$ and $k_{s,F}$) were highly correlated in all strains, and we were able to eliminate ϕ_{Eth} without substantial impacts on the quality of the model.

The final model depends on 33 parameters to be estimated from the experimental data. Their optimal values and the corresponding confidence intervals are reported in Tables 1 and 2. Figures 2 and 3 present the predictions of the model for biomass and variables associated with growth and fermentation rates. Despite the extensive reduction of the model, it can still satisfactorily explain the dry weight, sugar uptake and YAN trajectories. The median R^2 was 0.95 for all observables and all strains. It should be noted that the model captured well the dynamics of glycerol, acetate and higher alcohols, such as 2-phenylethanol and isoamyl alcohol, all of which are relevant to the quality of the wine.

Differences in primary and secondary metabolism

As expected, *S. cerevisae* T73 was the highest producer of ethanol in terms of both final concentration and estimated yield ($Y_{eth} = 0.47 \cdot g_{eth} \cdot g^{-1}_{glucose}$). In contrast, as shown in Figure 2, ScT73 presented the lowest final glycerol concentration. Interestingly, for T73, the estimated yield for glycerol production ($Y_{Gly} = 0.047 \cdot g_{glycerol} \cdot g^{-1}_{glucose}$) was higher than those of SuBMV58 (0.043) and SuCECT12600 (0.043). This result is explained by the dynamics of $G_{N,S}$ and modulated by the parameter $\tau_{G_{N,S}}$, which was estimated to be between 2.3 and 7.6 times higher in ScT73 than in the other strains. Therefore, to a large extent, differences in glycerol production were influenced by how long after nitrogen depletion the cells continued to produce glycerol. Furthermore, the same pattern of early inhibition of higher alcohols by $G_{N,S}$ was also able to explain the formation of aromas such as 2-phenylethanol and

TABLE 1 Biomass and nutrients uptake-related parameters: optimal values and associated confidence intervals.

	ScT73		SuBMV58		SuCECT12600		SkCR85	
Par	**Value**	$\theta(\%)$	**Value**	$\theta(\%)$	**Value**	$\theta(\%)$	**Value**	$\theta(\%)$
Φ_C	2.82×10^{-1}	512	1.61×10^{-1}	136	1.61×10^{-2}	78.4	1.65	783
$Y_{x/N}$	13.70	12	5.47	30.3	9.01	11.49	9.53	17.6
μ_{maxN}	5.00×10^{-1}	17.50	2.43×10^{-1}	22.4	4.35×10^{-1}	21.9	1.02×10^{-1}	8.1
Θ_C	4.05×10^{-1}	23	7.90×10^{-1}	31.6	1	55.7	4.69×10^{-1}	22.9
k_N	2.90×10^{-1}	28.90	9.22×10^{-2}	54.5	2.74×10^{-1}	39.5	4.24×10^{-2}	43.7
kd	5.07×10^{-3}	19.1	8.01×10^{-3}	123	9.20×10^{-3}	40.4	6.37×10^{-3}	21.2
$v_{max,Glx}$	1.20	14.60	3.85	59.8	2.08	28.5	1.88	28.9
$k_{s,Glx}$	37.85	21.1	51.07	128	55.47	47.8	46.95	45.2
$k_{s,F}$	1.40×10^{2}	16.9	1.69×10^{2}	92.3	1.95×10^{2}	39.6	2.00×10^{2}	35.2
$k_{sucrose}$	3.38×10^{-2}	0.9	3.06×10^{-2}	0.802	3.61×10^{-2}	0.273	2.14×10^{-2}	1.23
$\tau_{G_{N,S}}$	1.12×10^{-1}	19.1	1.48×10^{-2}	52	2.45×10^{-2}	36.3	4.80×10^{-2}	26.1
a_0	1	–	1	–	1	–	2.18×10^{-2}	63.2
λ_C	0.29	–	0.29	–	0.29	–	0.29	–
λ_P	0.59	–	0.59	–	0.59	–	0.59	–
k_{sC}	1.00×10^{-4}	–	1.00×10^{-4}	–	1.00×10^{-4}	–	1.00×10^{-4}	–

Note: Confidence intervals marked with the symbol – correspond to the parameters whose values were fixed during the optimization. The parameter descriptions can be found in Section 3.

TABLE 2 Parameters related to the production and consumption of products: Optimal values and associated confidence intervals.

	ScT73		SuBMV58		SuCECT12600		SkCR85	
Par	**Value**	$\theta(\%)$	**Value**	$\theta(\%)$	**Value**	$\theta(\%)$	**Value**	$\theta(\%)$
Y_{Eth}	4.71×10^{-1}	1.16	4.32×10^{-1}	4.88	4.57×10^{-1}	2.82	4.67×10^{-1}	2.9
$Y_{Glycerol}$	4.71×10^{-2}	4.13	4.28×10^{-2}	7.16	4.29×10^{-2}	4.2	5.78×10^{-2}	5.28
Y_{Ace}	2.63×10^{-3}	14.4	4.13×10^{-3}	37.1	2.28×10^{-3}	49.8	1.39×10^{-2}	7.83
kc_{Ace}	1.04×10^{-2}	44.8	3.53×10^{-2}	51.2	6.56×10^{-2}	24.6	1.50×10^{-2}	37.7
$kc_{Citrate}$	1.09×10^{-2}	15.6	4.83×10^{-2}	38.6	1.52×10^{-2}	20	1.41×10^{-2}	21
kc_{Malate}	1.88×10^{-3}	16.5	1.72×10^{-2}	39.1	4.67×10^{-3}	19.1	7.44×10^{-3}	20.7
$Y_{Succinate}$	3.49×10^{-3}	7.39	3.48×10^{-2}	31.9	1.29×10^{-2}	19.8	7.45×10^{-3}	9.99
$Y_{Lactate}$	1.24×10^{-3}	4.01	1.13×10^{-3}	11	1.35×10^{-3}	5.19	9.37×10^{-4}	5.44
$Y_{EthylAcetate}$	1.07×10^{-3}	10.4	9.11×10^{-4}	8.15	9.53×10^{-4}	3.95	9.13×10^{-4}	9.81
$Y_{IsobutylAcetate}$	4.36×10^{-6}	17.9	2.75×10^{-6}	5.43	2.68×10^{-6}	3.02	1.62×10^{-6}	8.93
$Y_{Isobutanol}$	5.22×10^{-4}	10.8	3.89×10^{-4}	9.43	3.44×10^{-4}	3.26	3.61×10^{-4}	9.46
$Y_{IsoamylAcetate}$	1.80×10^{-4}	9.68	9.05×10^{-5}	10.2	9.54×10^{-5}	3.29	2.50×10^{-5}	9.34
$Y_{Isoamylalcohol}$	4.57×10^{-3}	8.73	2.59×10^{-3}	8.12	2.01×10^{-3}	4.42	2.12×10^{-3}	10.1
$Y_{PhenylethylAc}$	3.38×10^{-5}	8.3	9.90×10^{-5}	10.6	6.67×10^{-5}	3.53	1.63×10^{-5}	6.91
$Y_{Phenylethanol}$	1.37×10^{-3}	7.87	4.65×10^{-3}	7.66	2.64×10^{-3}	4.88	9.78×10^{-4}	12.7
$Y_{EthylCaproate}$	3.06×10^{-5}	6.62	1.08×10^{-5}	6.24	1.06×10^{-5}	3.95	1.94×10^{-5}	8.98
$Y_{EthylCaprate}$	3.82×10^{-5}	18.6	7.14×10^{-5}	6.69	5.58×10^{-5}	9.56	4.58×10^{-5}	4.48
$Y_{BenzylAlcohol}$	8.42×10^{-5}	25.6	–	–	1.01×10^{-4}	6.86	1.93×10^{-4}	4.89

Note: Parameters marked with $Y\left(g / g_{hexose}\right)$ refer to the yield production of a specific higher alcohol or ester. Parameters marked with $k\left(h^{-1} \cdot g_{protein}^{-1}\right)$ are related to the mass action kinetics. Parameter values marked with the symbol – correspond to those cases in which available data were scarce or too noisy to estimate the corresponding model parameters.

isoamyl alcohol, highlighted by the good R^2 square values obtained (between 0.86 and 0.98).

In the proposed model, succinate production was initiated at the beginning of nitrogen depletion. As shown in Figure 3, the onset of succinate flux is well captured by the model (R^2 between 0.82 and 0.94). Interestingly, the model could not accurately explain the data in the case of ScT73 (R^2 = 0.69). Given that, compared to other species, ScT73 shows much lower succinate production, possibly because its production has different causal explanations in fermentations of *S. cerevisiae* and the cold-tolerant strains analysed here.

The model was also able to capture the dynamics of acetate production and consumption. SkCR85 produced and consumed the largest amounts of acetate. In contrast, ScT73 produced a similar amount of acetate compared to SuBMV58 or SuCECT12600 but consumed only a very small fraction at the beginning of nitrogen depletion, which is in line with the small estimate for $k_{ACE,C}$ ($1.04x10^{-2} \cdot h^{-1} \cdot g_{DW}^{-1}$) and the previously mentioned rapid onset of $G_{N,S}$.

Regarding the fit to the data for higher alcohols and esters, the median R^2 values over all measured compounds were 0.93 for ScT73, 0.92 for SuBMW58, 0.95 for SuCECT12600 and 0.91 for SkCR85. The fits to the data can be seen in Figure 2 and Figures S1 and S2. The yields are well identified from the data with confidence intervals $\leq 25\%$ for most compounds (Table 2). It should be noted that their values are generally quite low, with orders of magnitude ranging from 10^{-6} and 10^{-3}. When comparing the performance of strains in terms of these metabolic products, ScT73 produces the highest amounts of isobutanol, isoamyl alcohol, isobutyl acetate, isoamyl acetate, and ethyl caproate; SuBMV58 produces the highest amount of phenylethanol, phenylethyl acetate and ethyl caprate.

Exploring the modulation of aroma production with nitrogen supplementation

We then used the model to predict the role of nitrogen supplementation in the production of relevant metabolites. We considered two different scenarios: first, in which we modified the initial concentration of YAN (the results are shown in Figure 4) and second,

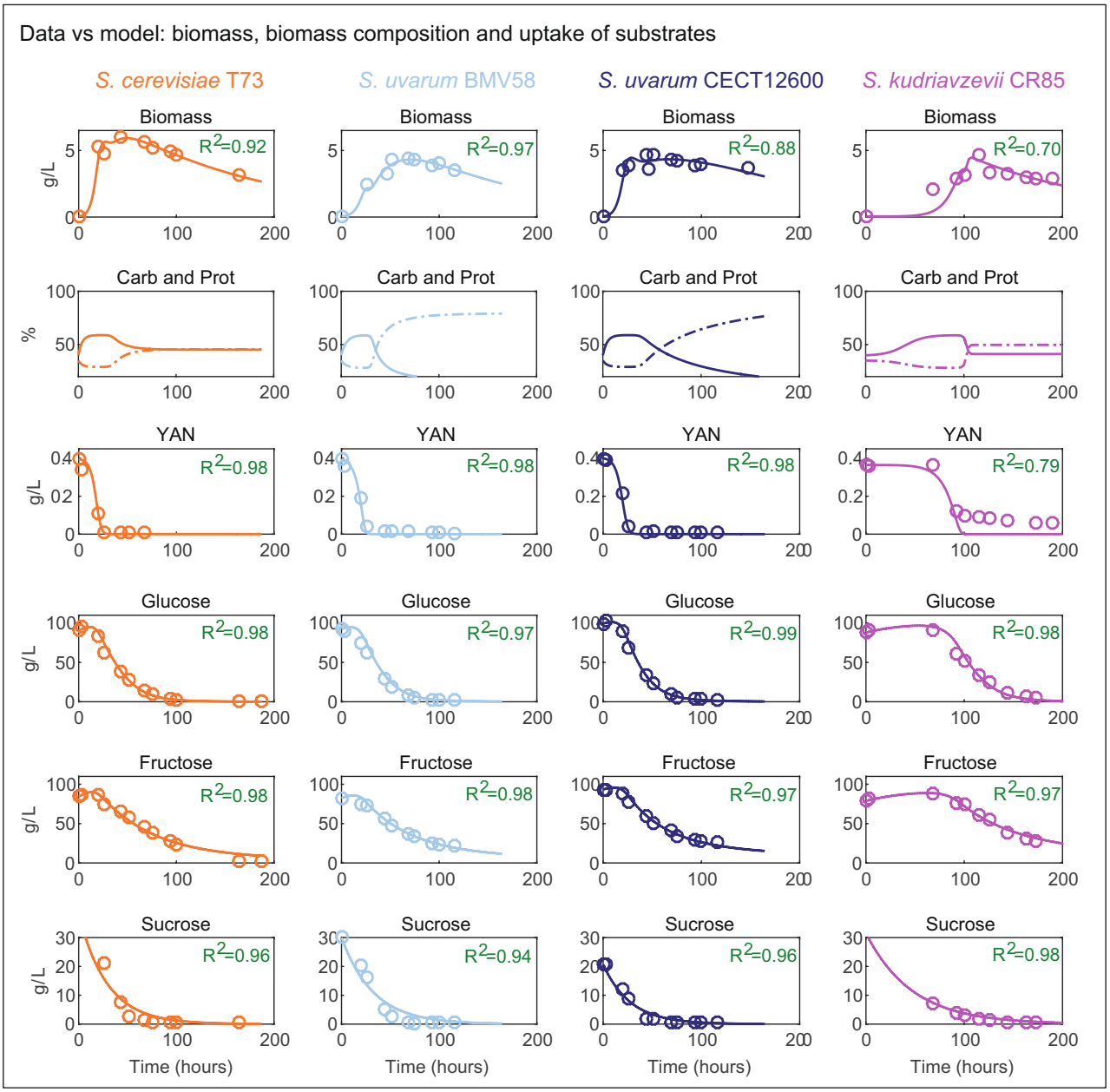

FIGURE 2 Fit of the model to growth-related time course data for batch fermentations led by four strains of three species of the *Saccharomyces* genus. Figures show the dynamics of biomass, YAN, glucose, fructose and sucrose. The figures also show the percentage of carbohydrates (dashed line) and proteins (solid line) in the biomass, as predicted by the model.

in which YAN was supplemented later during fermentation (the results are shown in Figure 5). In general, the supplementation of YAN accelerates the dynamics of growth, being this effect more evident for *S. cerevisiae* and *S. uvarum* strains. Early supplementation further enhances this effect (Figure S11). We estimated the uncertainty associated to model predictions using the uncertainty associated to the experimental error (Figures S7–S10).

Figure 4 shows a screening comparison between all species by the time at which fermentation ended (residual glucose is 2 g/L). As expected, the final concentration of the metabolites was directly correlated with the initial nitrogen concentration, with the exception of the production of acetate, succinate, and carboxylic acids. In these cases, in which the metabolites are consumed at the beginning of nitrogen depletion, low nitrogen supplementation produced higher amounts of the metabolite at the end of the process. For example, as shown in Figure 4, ScT73 produced 0.221 ± 0.042 g/L of acetate when using 200 mg, while a final concentration of 0.190 ± 0.036 g/L was achieved for high supplementation (500 mg). This pattern is also followed in the production of citrate and malate for

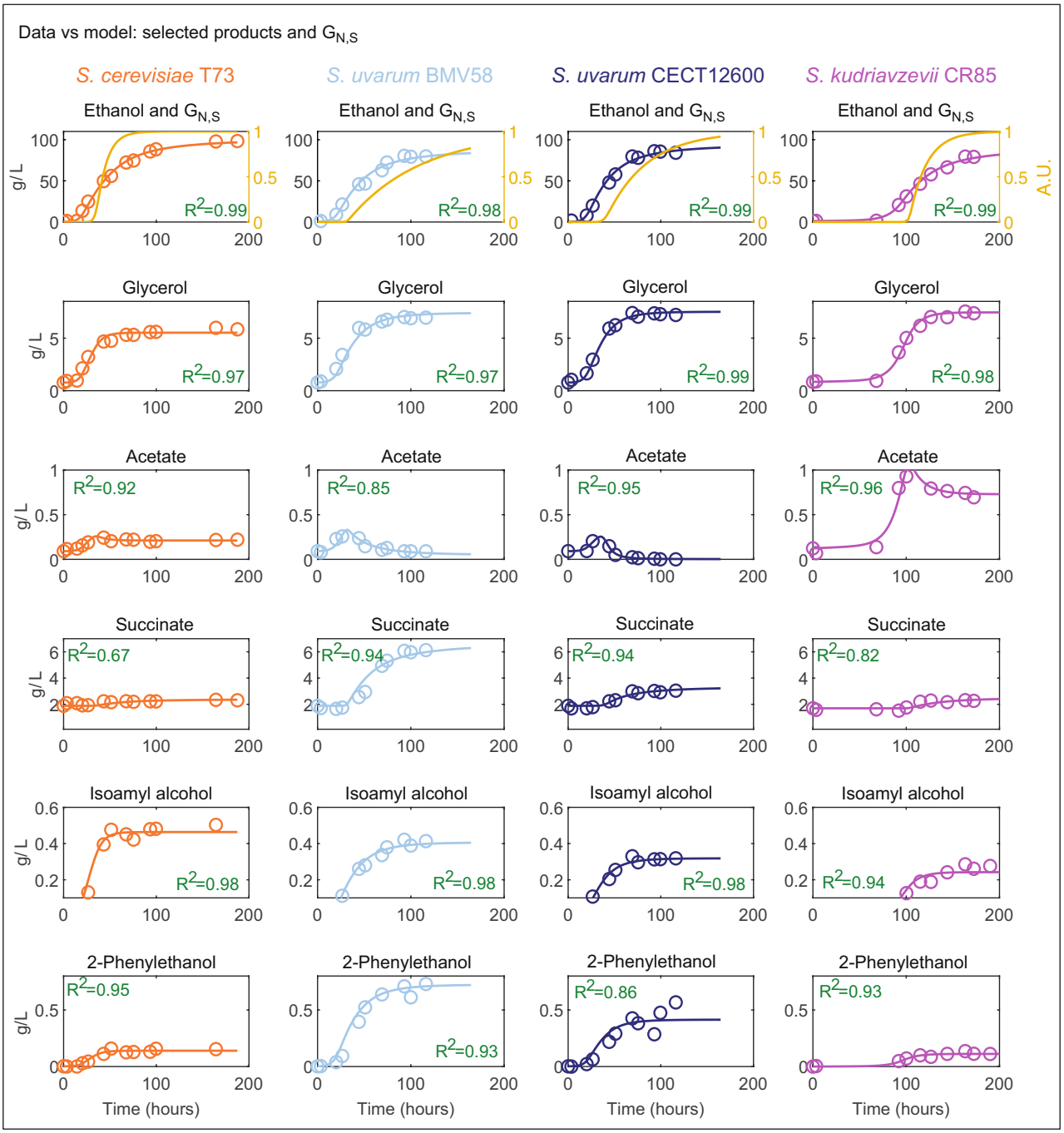

FIGURE 3 Fit of the model to fermentation time course data for batch fermentations led by four strains of three species of the *Saccharomyces* genus. Figures show the dynamics of the production of ethanol, glycerol, acetate, succinate, isoamyl alcohol and 2-phenylethanol.

SuBMV58 and SuCECT12600. In contrast, no significant differences were observed in acetate production in SkCR85 when changing the initial nitrogen conditions (0.711 ± 0.075 g/L for 500 mg, -0.709 ± 0.075 g/L for 400 mg, -0.715 ± 0.076 g/L for 300 mg and -0.723 ± 0.077 g/L for 200 mg).

The model predicts that ScT73 will produce wines with more ethanol but with lower amounts of glycerol than the other strains under all conditions tested. SuBMV58 will produce less ethanol than the wild strain SuCECT12600. SkCR85 would produce wine containing between $9.95 \pm 0.94\%$ and $9.98 \pm 0.94\%$ of alcohol and would be the most glycerol-producing strain (9.072 ± 0.420 g/L except for nitrogen-limited media). Interestingly, the available YAN at the beginning of fermentation does not significantly affect the final amount of ethanol but greatly improves the efficiency of the process. The results revealed that we can obtain similar

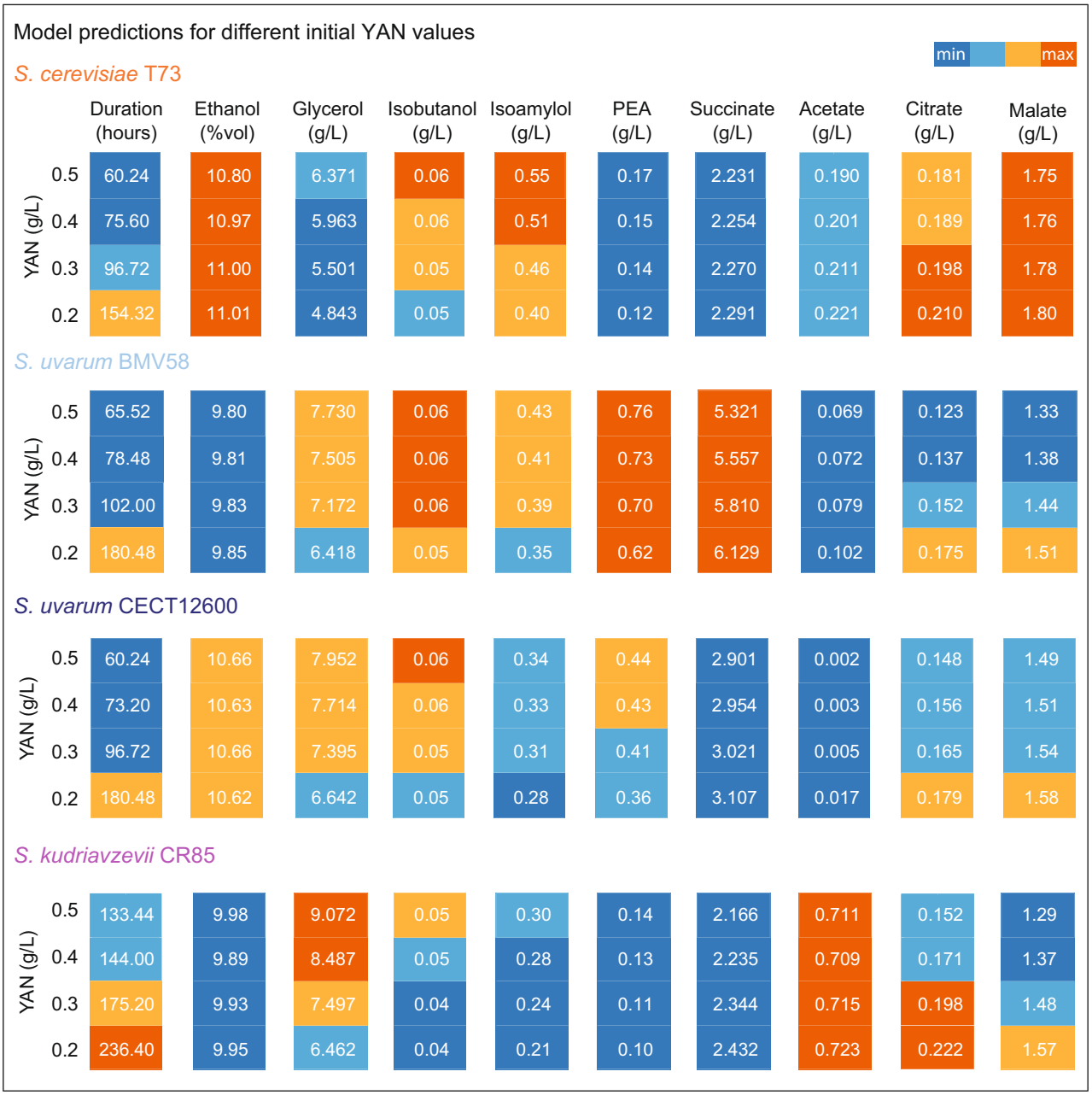

FIGURE 4 Predictions of the model for various initial nitrogen conditions. The figure shows the duration of fermentation in hours, as well as the final production of the relevant compounds. Four different initial values were considered that ranged from 200 to 500 mg of YAN. Dark blue, corresponds to the minimum production, while orange, corresponds to the maximum production. The same scale was used for all strains to facilitate the comparison.

amounts of ethanol but reduce the duration of fermentation between a minimum of 10% and a maximum of 47% by increasing the initial amount of YAN by 100 mg.

Simulations show similar behaviour in isobutanol production for all species and under all tested conditions. In contrast, there are differences in the production of 2-phenylethanol and isoamyl alcohol. SuBMV58 produced the maximum amount of 2-phenylethanol (0.76 ± 0.15 g/L for 500 mg of YAN) followed by SuCECT12600 with a maximum of (0.44 ± 0.07 g/L for 500 mg of YAN). ScT73 and SkCR85 produced similar amounts of 2-phenylethanol, with ScT73 being slightly more efficient, while ScT73 produced more isoamyl alcohol than the other species under all tested conditions (ScT73: 0.55 ± 0.15 g/L; SuBMV58 0.43 ± 0.10 g/L; SuCECT1600: 0.34 ± 0.04 g/L; SkCR85: 0.30 ± 0.09 g/L for 500 mg of YAN).

We have also considered the scenario in which the initial amount of YAN is the one found in the natural must (approximately 300 mg) and we supplemented the medium with 200 mg later in the process. We compared the additions in the exponential phase and during the

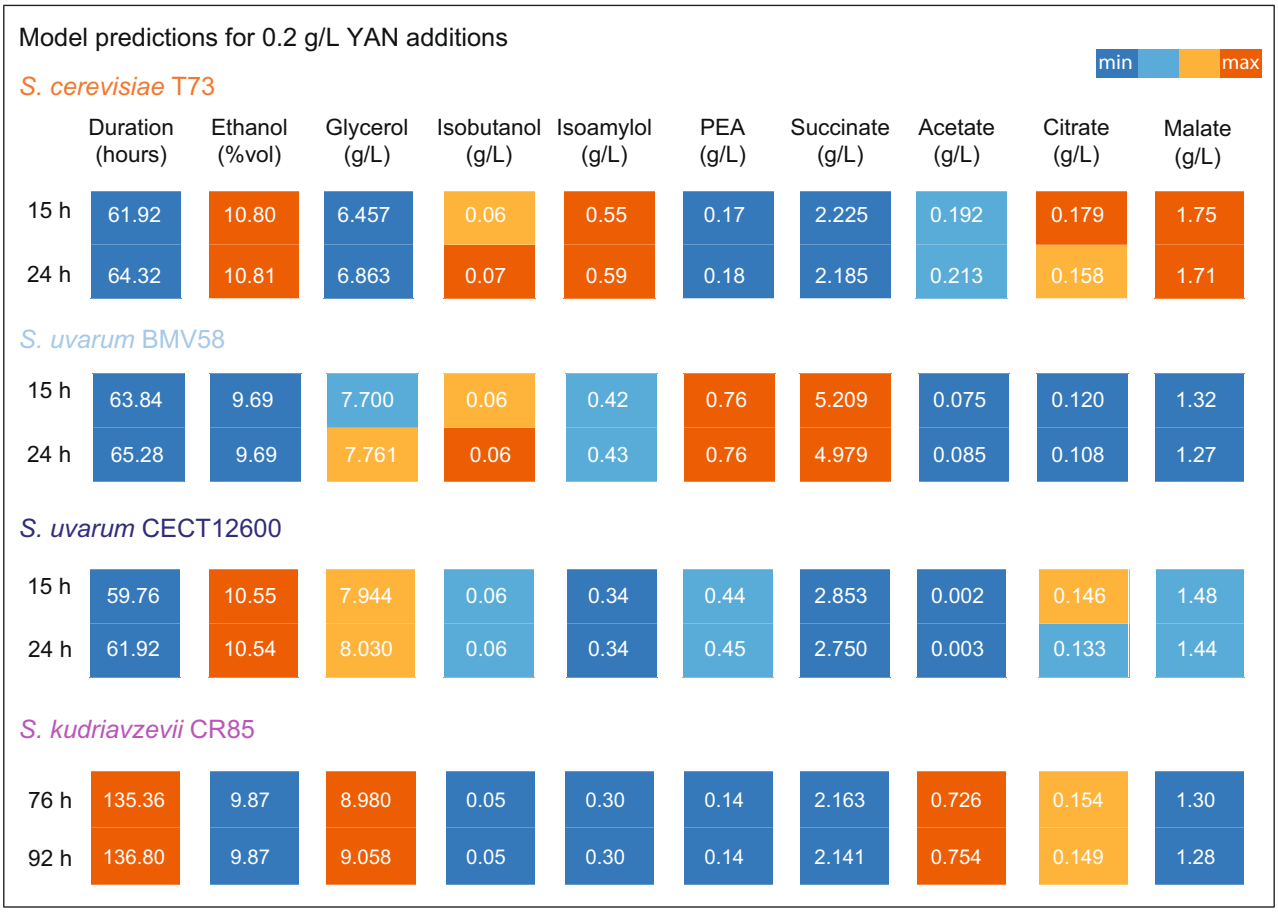

FIGURE 5 Predictions of the model for various nitrogen supplementations. The figure shows the duration of fermentation in hours, as well as the final production of the relevant compounds. The figure shows the time in hours at which the YAN was added: during the exponential growth phase (15 h for ScT73, SuBMV58 and SuCECT12600, and 76 h for SkCR85) and during the transition to YAN exhaustion (24 h for ScT73, SuBMV58 and SuCECT12600, and 92 h for SkCR85). The initial YAN corresponded to 0.307 g/L. Dark blue corresponds to the minimum production, while orange corresponds to the maximum production compared between different species.

transition to YAN exhaustion. The results reveal that supplementation later in the process tends to prolong the duration of fermentation, up to 10% in the case of ScT73. The glycerol yield was also slightly larger (Figure S11). Higher differences were observed in the final amounts of succinate, around the 4% increase for the *S. uvarum* strain when supplemented in the exponential phase. In contrast, supplementation in the exponential phase decreased the final amount of acetate, up to 50% in the case of SuCECT12600. The final malate and citrate concentrations were also favoured by an early supplementation.

Comparing the results obtained with the results previously obtained for initial supplementation (Figure 4, 0.5 g/L), we observe that the final amount of ethanol and succinate tends to increase when supplementing at the beginning of fermentation.

DISCUSSION

To the best of our knowledge, this is the first work to propose a kinetic model to predict secondary metabolism of *Saccharomyces* yeast in wine. Here, we developed a model based on experimental data, previous modelling efforts, and evidence from the literature. More specifically, we combined a previous model of biomass formation (Henriques & Balsa-Canto, 2021) with a mechanism that simulates the impact of transcriptional reprogramming (Rossignol et al., 2003) on metabolic fluxes in nitrogen-starved cells, in particular on the synthesis of fermentation byproducts such as glycerol and aromas.

The introduction of this mechanism ($G_{N,S}$) is based on the earlier finding that, unlike ethanol, there was no higher alcohol production during the decay phase (Henriques et al., 2021). Until recently, the most prevalent view was that higher alcohol production in fermented beverages was mainly associated with the catabolism of exogenous amino acids. However, this is in contradiction to previous results showing higher alcohol formation without nitrogen sources (Stevens, 1961). Recently, Crépin et al. (2017) added new information on this question, showing that a significant fraction of the higher alcohols and ester produced during alcoholic

fermentation come from the catabolism of amino acids synthesized de novo, that is, sugar metabolism.

Regarding biomass formation, we divided growth into two components (μ_N and μ_C). The first is associated with the availability of the limiting substrate (YAN) and the second with the accumulation of carbohydrates when nitrogen becomes scarce. Although the formation of new protein and mRNA is strictly dependent on YAN (required for the synthesis of amino acids), storage carbohydrates, such as glycogen, can be transiently formed from extracellular hexoses in the absence of nitrogen (Schulze et al., 1996; Varela et al., 2004). Thus, during early growth ($\mu_N \gg \mu_C$), biomass composition remains relatively stable, with mRNA and protein covering a considerable fraction. In contrast, when low nitrogen is available, growth is characterized by a reduced contribution of μ_N and increased μ_C. Although this strategy adds complexity to the model, it is important to consider the fact that biomass can increase after nitrogen depletion to accurately describe the experimental data. Although this model has been show to nicely capture biomass composition data in two independent data sets (Henriques & Balsa-Canto, 2021; Schulze et al., 1996; Varela et al., 2004), a possible contribution from the mobilization of arginine storages (Crépin et al., 2014), in some experimental conditions, cannot be ruled out.

We performed a parameter identifiability analysis and found that the model has wide confidence intervals, particularly in the parameters associated with biomass formation and fermentation rate. To address this problem, we simplified the model. The finding that the parameters associated with hexose transport ($V_{max,H}$ and $k_{s,H}$) are highly correlated is relevant because it probably applies previous modelling efforts. Pizarro et al. (2007) used parameter values obtained for hexose transporters in vitro and information on the concentrations of multiple hexose transporters. Therefore, this alternative would only be applicable to *S. cerevisiae*, for which detailed kinetic studies have been performed on the transporters. Although poor identifiability could limit the predictive capabilities of inferred models (Chis et al., 2016; Lira-Parada et al., 2021), in the future, these issues could be addressed with a careful experimental design (Balsa-Canto et al., 2008), maximizing the information content of the parameters.

Important differences were found between the species analysed. For example, the dynamics of $G_{N,S}$ were fast in *S. cerevisiae* and comparatively slower in *S. kudriavzevii* and especially in *S. uvarum*. These differences are interesting since, if confirmed, they could be exploited to tune the aromatic profiles of wines.

Recently, Godillot et al. (2022) investigated the aroma composition of wines with varying initial nitrogen (amino acids with diphosphate ammonia), added nitrogen and the timing of nitrogen addition. This recent dataset offers the possibility to validate or invalidate some of the modelling assumptions described in this work. In this respect, consistent with the observations of Godillot et al. (2022) our model predicts that the addition of nitrogen results in an increase in the final amount of PEA. However, the same authors did not observe the effect on the production of isoamyl alcohol that we predicted with our model. Furthermore, the authors noted a negative linear effect on the initial nitrogen, which is in contrast to our model predictions. A possible explanation for this finding could be the fact that the model, despite the good fit in terms of the R^2 score, is not perfectly synchronized with the isoamyl alcohol production data. In fact, for ScT73 and SkCR85, isoamyl alcohol production appears to begin early compared to the data, while the inverse pattern can be observed for *S. uvarum* strains (Figure 3). Taking into account the results of amino acid consumption obtained by Crépin et al. (2012), it seems likely that these patterns in higher alcohol behaviour are caused by the activation of genes under nitrogen catabolite repression by the SPS system (Ljungdahl, 2009), leading to the spillover of excess higher alcohols in the absence of exogenous amino acids.

CONCLUSIONS

In this work we propose a continuous model that incorporates the physiological status of yeast to improve the description of the fermentation process and, in particular, of glycerol and higher alcohol formation. The strength of our model is that it is able to accurately capture the multiple phases of fermentation by introducing a simple representation of regulatory mechanisms. Including the role of preferential nitrogen sources (ammonia, glutamate and alanine) and relating them to the activation of genes leading to amino acid degradation genes should motivate future research on the detailed modelling of the dynamics of aromagenesis.

We expect that the model could also be applied to non-Saccharomyces yeasts by recalibrating the parameters involved with a similar data setting. This opens new opportunities for the design of novel single culture fermentation processes.

AUTHOR CONTRIBUTIONS

Artai R. Moimenta: Formal analysis (equal); investigation (equal); visualization (equal); writing – original draft (equal). **David Henriques:** Formal analysis (equal); investigation (equal); supervision (equal); visualization (equal); writing – original draft (equal). **Romain Minebois:** Data curation (equal); investigation (equal); writing – review and editing (equal). **Amparo Querol:** Conceptualization (equal); data curation (equal); funding acquisition (equal); investigation (equal); writing – review and editing (equal). **Eva Balsa-Canto:**

Conceptualization (lead); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (lead); supervision (equal); visualization (equal); writing – original draft (equal); writing – review and editing (lead).

ACKNOWLEDGEMENTS

This work has received funding from MCIU/AEI/FEDER, UE grant references: RTI2018-093744-B-C31, RTI2018-093744-B-C33; MCIN/AEI/10.13039/501100011033 and NextGenerationEU/PRTR grant reference: PLEC2021-007827 and Xunta de Galicia (IN607B 2020/03). RM was supported by an FPI grant from the Ministerio de Economía y Competitividad, Spain (ref. BES-2016-078202).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

ORCID

Artai R. Moimenta https://orcid.org/0000-0002-5609-2317
David Henriques https://orcid.org/0000-0002-9477-292X
Romain Minebois https://orcid.org/0000-0001-6959-1572
Amparo Querol https://orcid.org/0000-0002-6478-6845
Eva Balsa-Canto https://orcid.org/0000-0002-1978-2626

REFERENCES

Balsa-Canto, E., Alonso, A.A. & Banga, J.R. (2008) Computational procedures for optimal experimental design in biological systems. *IET Systems Biology*, 2(4), 163–172.

Balsa-Canto, E., Henriques, D., Gabor, A. & Banga, J.R. (2016) Amigo2, a toolbox for dynamic modeling, optimization and control in systems biology. *Bioinformatics*, 32(21), 3357–3359.

Baranyi, J. & Roberts, T.A. (1994) A dynamic approach to predicting bacterial growth in food. *International Journal of Food Microbiology*, 23, 277–294.

Bisson, L.F. & Butzke, C.E. (2000) Diagnosis and rectification of stuck and sluggish fermentations. *American Journal of Enology and Viticulture*, 51(2), 168–177.

Chis, O., Villaverde, A.F., Banga, J.R. & Balsa-Canto, E. (2016) On the relationship between sloppiness and identifiability. *Mathematical Biosciences*, 282, 147–161.

Coleman, M.C., Fish, R. & Block, D.E. (2007) Temperature-dependent kinetic model for nitrogen-limited wine fermentation. *Applied and Environmental Microbiology*, 73(18), 5875–5884.

Cramer, A.C., Vlassides, S. & Block, D.E. (2002) Kinetic model for nitrogen-limited wine fermentations. *Biotechnology and Bioengineering*, 77(1), 49–60.

Crépin, L., Nidelet, T., Sanchez, I., Dequin, S. & Camarasa, C. (2012) Sequential use of nitrogen compounds by *Saccharomyces cerevisiae* during wine fermentation: a model based on kinetic and regulation characteristics of nitrogen permeases. *Applied and Environmental Microbiology*, 78(22), 8102–8111.

Crépin, L., Sanchez, I., Nidelet, T., Dequin, S. & Camarasa, C. (2014) Efficient ammonium uptake and mobilization of vacuolar arginine by *Saccharomyces cerevisiae* wine strains during wine fermentation. *Microbial Cell Factories*, 13(1), 1–13.

Crépin, L., Truong, N.M., Bloem, A., Sanchez, I., Dequin, S. & Camarasa, C. (2017) Management of multiple nitrogen sources during wine fermentation by *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 83(5), e02617–e02616.

Egea, J.A., Balsa-Canto, E., Garcia, M.G. & Banga, J.R. (2009) Dynamic optimization of nonlinear processes with an enhanced scatter search method. *Industrial & Engineering Chemistry Research*, 48(9), 4388–4401.

Gamero, A., Tronchoni, J., Querol, A. & Belloch, C. (2013) Production of aroma compounds by cryotolerant *Saccharomyces* species and hybrids at low and moderate fermentation temperatures. *Journal of Applied Microbiology*, 114(5), 1405–1414.

Godillot, J., Sanchez, I., Perez, M., Picou, C., Galeote, V., Sablayrolles, J. et al. (2022) The timing of nitrogen addition impacts yeast genes expression and the production of aroma compounds during wine fermentation. *Front Microbiology*, 13. https://doi.org/10.3389/fmicb.2022.829786

Henriques, D. & Balsa-Canto, E. (2021) The Monod model is insufficient to explain biomass growth in nitrogen-limited yeast fermentation. *Applied and Environmental Microbiology*, 87(20), e01084–e01021.

Henriques, D., Minebois, R., Mendoza, S.N., Macías, L.G., Pérez-Torrado, R., Barrio, E. et al. (2021) A multiphase multiobjective dynamic genome-scale model shows different redox balancing among yeast species of the *Saccharomyces* genus in fermentation. *mSystems*, 6(4), e0026021.

Hindmarsh, A.C., Brown, P.N., Grant, K.E., Lee, S.L., Serban, R., Shumaker, D.E. et al. (2005) SUNDIALS: suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, 31(3), 363–396.

Hittinger, C.T., Steele, J.L. & Ryder, D.S. (2018) Diverse yeasts for diverse fermented beverages and foods. *Current Opinion in Biotechnology*, 49, 199–206.

Hjersted, J., Henson, M. & Mahadevan, R. (2007) Genome-scale analysis of *Saccharomyces cerevisiae* metabolism and ethanol production in fed-batch culture. *Biotechnology and Bioengineering*, 97, 1190–1204.

Lairón-Peris, M., Pérez-Través, L., Muñiz-Calvo, S., Guillamón, J.M., Heras, J.M., Barrio, E. et al. (2020) Differential contribution of the parental genomes to a *S. cerevisiae* × *S. uvarum* hybrid, inferred by phenomic, genomic, and transcriptomic analyses, at different industrial stress conditions. *Frontiers in Bioengineering and Biotechnology*, 8, 129.

Lira-Parada, P.A., Pettersen, E., Biegler, L.T. & Bar, N. (2021) Implications of dimensional analysis in bioreactor models: parameter estimation and identifiability. *Chemical Engineering Journal*, 417, 129220.

Ljungdahl, P.O. (2009) Amino-acid-induced signalling via the SPS-sensing pathway in yeast. *Biochemical Society Transactions*, 37(Pt 1), 242–247.

López-Malo, M., Querol, A. & Guillamón, J.M. (2013) Metabolomic comparison of *Saccharomyces cerevisiae* and the cryotolerant species *S. bayanus* var. *uvarum* and *S. kudriavzevii* during wine fermentation at low temperature. *PLoS One*, 8(3), 1–14.

Miller, K.V. & Block, D.E. (2020) A review of wine fermentation process modeling. *Journal of Food Engineering*, 273, 109783.

Minebois, R., Pérez-Torrado, R. & Querol, A. (2020) A time course metabolism comparison among *Saccharomyces cerevisiae*, *S. uvarum* and *S. kudriavzevii* species in wine fermentation. *Food Microbiology*, 90, 103484.

Pérez, D., Denat, M., Minebois, R., Heras, J.M., Guillamón, J.M., Ferreira, V. et al. (2022) Modulation of aroma and chemical composition of Albariño semi-synthetic wines by non-wine saccharomyces yeasts and bottle aging. *Food Microbiology*, 104, 103981.

Pérez-Torrado, R., Barrio, E. & Querol, A. (2018) Alternative yeasts for winemaking: *Saccharomyces* non-*cerevisiae* and its hybrids. *Critical Reviews in Food Science and Nutrition*, 58(11), 1780–1790.

Pizarro, F., Vargas, F.A. & Agosin, E. (2007) A systems biology perspective of wine fermentations. *Yeast*, 24(11), 977–991.

Poznyak, V. & Rekve, D. (2015) *Global status report on alcohol and health 2018*. Who, P450. Available from: https://www.who.int/substance_abuse/publications/global_alcohol_report/gsr_2018/en/ [Accessed 22nd January 2023].

Querol, A., Pérez-Torrado, R., Alonso-Del-Real, J., Minebois, R., Stribny, J., Oliveira, B.M. et al. (2018) New trends in the uses of yeasts in oenology. *Advances in Food and Nutrition Research*, 85, 177–210.

Rossignol, T., Dulau, L., Julien, A. & Blondin, B. (2003) Genome-wide monitoring of wine yeast gene expression during alcoholic fermentation. *Yeast*, 20(16), 1369–1385.

Schulze, U., Lidén, G. & Villadsen, J. (1996) Dynamics of ammonia uptake in nitrogen limited anaerobic cultures of *Saccharomyces cerevisiae*. *Journal of Biotechnology*, 46(1), 33–42.

Scott, W.T., van Mastrigt, O., Block, D.E., Notebaart, R.A., Smid, E.J. & Cuomo, C.A. (2021) Nitrogenous compound utilization and production of volatile organic compounds among commercial wine yeasts highlight strain-specific metabolic diversity. *Microbiology Spectrum*, 9(1), e00485–e00421.

Steensels, J. & Verstrepen, K.J. (2014) Taming wild yeast: potential of conventional and nonconventional yeasts in industrial fermentations. *Annual Review of Microbiology*, 68(1), 61–80.

Stevens, R. (1961) Formation of phenethyl alcohol and tyrosol during fermentation of a synthetic medium lacking amino-acids. *Nature*, 191(4791), 913–914.

Varela, C., Dry, P.R., Kutyna, D.R., Francis, H.P.A.I.L., Curtin, C.D. & Chambers, P.J. (2015) Strategies for reducing alcohol concentration in wine. *Australian Journal of Grape and Wine Research*, 21, 670–679.

Varela, C., Pizarro, F. & Agosin, E. (2004) Biomass content governs fermentation rate in nitrogen-deficient wine musts. *Applied and Environmental Microbiology*, 70(6), 3392–3400.

Vilas, C., Arias-Méndez, A., Garcia, M.R., Alonso, A.A. & Balsa-Canto, E. (2018) Towards predictive food process models: a protocol for parameter estimation. *Critical Reviews in Food Science and Nutrition*, 58(3), 436–449.

Walter, E. & Pronzato, L. (1997) *Identification of parametric models from experimental data*. Masson, IA: Springer.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Moimenta, A.R., Henriques, D., Minebois, R., Querol, A. & Balsa-Canto, E. (2023) Modelling the physiological status of yeast during wine fermentation enables the prediction of secondary metabolism. *Microbial Biotechnology*, 00, 1–15. Available from: https://doi.org/10.1111/1751-7915.14211