RESEARCH ARTICLE

Autofluorescence prediction model for fluorescence unmixing and age determination

Marco Eigenfeld ⦿ | **Roland Kerpes** | **Iain Whitehead** | **Thomas Becker**

Technical University of Munich, School of Life Science, Institute of Brewing and Beverage Technology, Freising, Germany

Correspondence
Roland Kerpes, Technical University of Munich, School of Life Science, Institute of Brewing and Beverage Technology, Freising, Germany.
Email: roland.kerpes@tum.de

Abstract

Background: Flow cytometry is a powerful tool for identifying and quantifying various cell markers, such as viability, vitality, and individual cell age, at single-cell stages. However, cell autofluorescence and marker fluorophore signals overlap at low fluorescence intensities. Thus, these signals must be unmixed before determining the age fraction.

Methods and Results: A comparison was made between principal component regression (PCR) and random forest (RF) to predict autofluorescence signals of *Saccharomyces pastorianus* var. *carlsbergensis* in a flow cytometer. RF provided better prediction results than the PCR and was therefore determined to be better suited for unmixing signals. In the subsequent application for unmixing the autofluorescence signal from the marker fluorophore signal, the Gaussian mixture analysis based on RF was in better agreement with the microscopy-determined replicative age distribution than the PCR-based method.

Conclusion: The proposed approach of single-laser spectral unmixing and subsequent Gaussian mixture analysis showed that the microscopy data was consistent with the unmixed fluorescence spectra. The demonstrated approach enables fast and reliable unmixing of flow cytometric spectral data using a single-laser spectral unmixing method. This analysis method enables age determination of cells in industrial processes. This age determination allows for quantifying the yeast cell's age fractions, providing a detailed view of age-related changes. Additionally, the bud scar labeling technique can be used to determine age-related changes in *Pichia pastoris* yeast for biotechnological applications or recombinant protein expression.

KEYWORDS
autofluorescence, flow cytometry, random forest, single-cell age analysis, spectral unmixing

- -

1 | INTRODUCTION

Biological entities such as cells or organisms undergo various changes during their lifetimes, which can be induced by external causes, including the physical or chemical environment. The physical influences include temperature, light, electric or magnetic fields, and gravity. Chemical effects include acidic or alkaline environments, gases, and solvents. For biotechnological analysis, the analysis of changes during a lifetime in high-throughput result in two main problems: (i) autofluorescence variance, limiting the detection of weak fluorophores resulting from overlapping signals of autofluorescence and fluorophore, and ii difficulties in single-cell age analysis using chitin staining, due to the small amounts of fluorophore binding sites on the cell surface. The internal causes of changes in organisms include cell growth, aging, and mutation.[1] Organisms respond to such changes by changing their molecular properties, such as emission signals. For example, *Saccharomyces cerevisiae* yeast cells were subjected to the accumulation of 3,3'-dipropylthiacarbocyanine iodide (dis-C_3(3)) and batch curing in a glucose (0.2%) medium.[2] A redshift in the maximum fluorescence (λ_{max}) from 569 to 582 nm was observed, in addition to the usual redshift that appears when the aqueous medium is changed from polar to nonpolar.[3] Also, the autofluorescence signal of *Escherichia coli* (*E. coli*) cells was reported to increase when they struggle to survive.[3] Thus, the molecular properties of cells are affected in various ways in response to external stimuli.

Many studies have used flow cytometry and luminescence detection results with plate readers to evaluate the fluorescence of cellular material.[4–7] Photoluminescence has also been used as a detection method in flow cytometry.[6,8–14] However, the presence of intrinsic cellular autofluorescence emissions (e.g., from the tryptophan residues of proteins in *Lactobacillus* sp. or *Saccharomyces* sp.,[15] nicotinamidadenindinukleotid (NADH),[16] or protoporphyrin IX[17]) can significantly limit the detection of fluorescent-labeled antibodies due to the use of antigenic determinants in flow cytometry. This limitation can be overcome using two principal methods. The first method uses two laser wavelengths: one for exciting the fluorescence and autofluorescence transitions and the second for exciting the autofluorescence transitions only.[8,9] The autofluorescence spectra are then subtracted on a cell-by-cell basis to filter the desired fluorescence signals.[9] The second method uses several detection filters to monochromatize the fluorescence signals.[1]

The unmixing process may be facilitated by staining the cells with autofluorescence emitters.[12] This method has been improved using microplate readers, which are sensitive to the specific lines of relaxation signals.[12–14,18–25] However, the microplate readers only estimate the mean response of the cell population, so information about slight variations around the mean is lost.[1] However, this existing method[1] has some limitations in that it only uses two channels (525 and 585 nm) for the entire signal, which may not provide sufficient signal resolution. In particular, the green fluorescent protein (GFP), which dominates the 525-nm channel, does not vanish at the 585-nm channel because of signal broadening.

Thus, Lichten et al.'s method[1] should be optimized for flow cytometry, which offers a more detailed cell description of, for example, autofluorescence and cell size.[8,10,11]

Furthermore, flow cytometry detects the entire signal using different channels, including the 525-nm GFP channel and other channels that detect signals except for the GFP broadening signals. Using these channels, the autofluorescence signals at the 525-nm channel can be predicted. Several approaches have been reported in the literature for such signal prediction data analysis, such as principal component analysis (PCA) for gene expression analysis of yeast cells,[26] principal component regression (PCR) for columnar ozone prediction,[27] and Bayesian PCA for yeast genomic changes and protein expression analyses.[28] The random forest (RF) regression method is another promising approach with high precision for automated use in taxonomic studies.[29–31] The main advantage of RF is that it can address non-linearity, and it avoids overfitting when scaling to higher dimensions.[32]

Overcoming the problem of unmixing autofluorescence signals from marker fluorescence signals is essential for correctly analyzing cell age distributions in mixed yeast populations.

Over the last three decades, *S. cerevisiae* has also evolved into a model organism for studying cell aging. However, aging is a very complex process, even in a eukaryote as simple as *S. cerevisiae*. In research into aging mechanisms, two approaches describing cell aging have been established: replicative cell age (RLS) and chronological cell age (CLS). However, the aging of *S. cerevisiae* is more likely to be affected by RLS than CLS.[33] Mortimer and Johnson made the first approach for studying RLS on yeast. They dissected 36 yeast cells and counted the number of bud scars, an indicator of replicative aging.[34] After this first experiment, later studies have shown that the average lifespan of yeast is relatively constant for a given strain.[35] However, the aging process negatively impacts different cell compartments like telomeres and genetic integrity,[36] and an increase in RLS is associated with decreased metabolic activity and generation time. Many studies have been performed in yeast cell separation[37] or microscopic age distribution analysis,[38] but all studies were limited by the non-representative cell numbers used for investigations. Studies according to the cell age of *Pichia pastoris* yeast, an essential industrial strain used for recombinant protein synthesis, have been performed at a lower level but are limited by similar causes. Flow cytometric analysis methods were still limited in the autofluorescence data. This signal overlap is why calibration curves were required to calculate the population's mean bud scar numbers,[39] limiting the results.

In this study, we developed an approach for spectral unmixing of fluorescence signals from autofluorescence of *S. pastorianus* yeast cells tagged with a GFP-containing protein linker for replicative age determination. The developed method can be used with flow cytometry to detect single-cell fluorescence in six channels. However, unmixing the measured signals is non-trivial. Two regression methods, PCR and RF, were compared to predict autofluorescence signal intensity in the GFP-fluorescence channel. To facilitate this comparison, samples of pure yeast cells and samples with a protein linker were measured to calibrate the single fluorescence–emission sources. This unmixing

approach allowed isolating a fluorophore signal from measurement data combining the autofluorescence signals and newly appearing GFP signals of the mixed spectrum, enabling an in detail replicative cell age analysis of yeast in high-throughput as well as representative cell numbers for the first time.

2 | MATERIALS AND METHODS

2.1 | Experiments

Details regarding the yeast strains and their cultivation are given in the Supporting Information Section S1.1. The binding mechanism of the protein to the bud scars of the yeast cells and details of the microscopy measurements are described in Sections S1.2 and S1.3, respectively. As described in Section S1.4, flow cytometry was performed on yeast cells without the protein linker to obtain the autofluorescence data. As described in Section S1.5, an adsorption isotherm was used to determine the binding parameters of His6-Sumo-sfGFP-ChBD to yeast cells.

2.2 | Analysis methods

The amount of chitin can be used as an indicator for the replicative age of yeast cells. This indicator is correlated with the number of yeast cells with a bud scar, which corresponds to the GFP-containing protein linker signal. However, yeast cells also have autofluorescence, which is linearly convoluted with the GFP fluorescence:

$$f_i = a_i + g_i, \tag{1}$$

where f_i is the fluorescence-signal intensity, a_i is the autofluorescence signal intensity, and g_i is the GFP signal intensity at wavelength i. In this study, PCR and RF regression were compared for their suitability for the prediction and subsequent unmixing of autofluorescence data.

2.2.1 | Principal component regression

Principle component analysis reduces the dimensionality of the datasets using covariance analysis between factors.[27] PCA maximizes the correlation between the original and new mutually orthogonal variables.[40] This method is suitable for analyzing highly correlated independent variables. In this study, we evaluated the use of PCA combined with subsequent regression (PCR) to reduce the size of the original dataset by excluding the prediction variable of pure yeast cells to one principal component and correlating this data with the autofluorescence data. For example, the original data of all fluorescence detectors excluding 525 nm were transformed. Then this principal component was used in a linear regression with the 525-nm autofluorescence data (code fragments are shown in Section S2.1). We then transformed the training dataset to the first principal component for

verifying the approach. Then, PCR was used to predict the autofluorescence signal at the 525-nm channel for yeast samples, which was then coupled with the protein linker to subtract this value from the raw data.

2.2.2 | Random-forest model

RF is a versatile machine-learning algorithm for classification and regression tasks, which is well suited for fitting complex datasets. The RF model can be used for regression and classification problems.[41,42] RF regression models calculate the sum of single full-grown decision trees built using classification and regression tree algorithms.[43] Each tree can be seen as a logical interpretation of an if-then criterion. Therefore, this model has a low tendency for overfitting and can be explained as a sum of weak regressions resulting in one strong regression tree. A disadvantage of the RF model is that it is a black box, i.e., it can create predictions and forecasts but usually lacks transparency.

To apply the RF model, the fluorescence signals excluding 525 nm were used for autofluorescence prediction. The code line using the *randomforest* package of R[44] is shown in Section S2.2, which implements the RF function published by Breiman et al.[45] Because the prediction variable at the 525-nm channel is not a factor, nonlinear multiple regression is used instead of classification.

2.3 | Autofluorescence analysis

This study aimed to develop a model for predicting the autofluorescence of yeast cells. The flow cytometry data consists of six parameters: front scatter (FSC), side scatter (SSC), and fluorescence signals at 525 nm (525–40), 585 nm (585–42), 610 nm (610–20), and 780 nm (780–60). We used all parameters, excluding the 525-nm channel of a reference sample (pure yeast cell), to predict the 525-nm signal. PCR and RF were used to predict and remove the overlapping autofluorescence signal to measure the pure fluorescence signal of the protein linker binding to the bud scars of yeast cells. The subtracted data were fitted to microscopy data of 11 datasets and compared to the results of Gaussian mixture analysis.

The flow cytometry data were gated via multivariate outlier analysis using the software R[46] and the aqplot function from the mvoutlier package[47,48] to remove dust particles and data points associated with properties uncharacteristic of yeast cells. The FSC and fluorescence signal at 525 nm were used for outlier detection. The fluorescence was measured as front scatter data (FSC-A) and increased with increasing cell size, as shown in Figure 1 (top). A previous study[3] revealed that the intensities of autofluorescence signals changes according to the fermentation media and conditions. The yeast cells were coupled with a protein linker that binds specifically to chitin components such as bud scars, which resulted in fluorescent bud scars and increased the fluorescence-signal intensity. The fluorescence data of yeast cells coupled with the protein linker (test data) and uncoupled yeast cells (control data) showed as overlapping, as shown in Figure 1 (bottom). The fluorescence intensities of the yeast cells with and without the

(A)

(B)
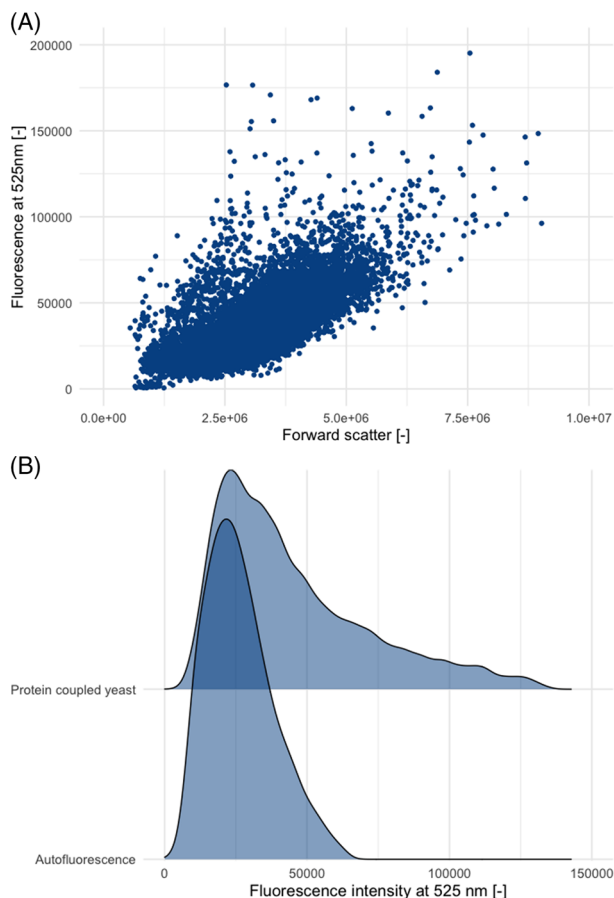
FIGURE 1 Comparison of fluorescence data at 525 nm and the cell size. Increase of autofluorescence signal intensity with size (top); comparison of yeast cells with and without protein linker, indicated by an increase in fluorescence-signal intensity (bottom)

protein linker differed significantly. Therefore, the obtained data required spectral unmixing for precise age determination of the yeast cells.

2.4 | Determination of BSN distribution

To date, no method exists that could determine BSN distributions without the use of microscopy. To address this, microscopic and fluorescence data were fitted toward each other. Microscopic BSN data were obtained as follows: Microscopy was used to count the number of visible bud scars on the yeast cell surface. Then, the Bayes theorem was used to predict the number of nonvisible bud scars based on a correlation between the cell surface and replicative age. This correlation was assumed in two ways: (i) linear transformation as shown in S3.1; and (ii) logarithmic transformation as demonstrated in S3.2.

After this step, autofluorescence-filtered data was evaluated in two ways, which were then compared against each other: (i) the microscopic data were fitted linearly to autofluorescence-cleaned data, which resulted in two coefficients; and (ii) the fluorescence data were fitted with Gaussian mixture analysis and the results were compared to

microscopic data. In this analysis, 11 independent yeast samples were used. The resulting approach enabling the determination of age distributions of yeast cells will provide more profound knowledge about the budding process in running fermentations.

2.4.1 | Spectral unmixing of fermentation data

The experiment aimed to compare the BSN distribution determined by microscopy with the fluorescence intensity determined by flow cytometry. The fluorescence data using PCR and RF methods were compared and fitted to the density distribution of the microscopy-determined BSN distribution, which had the linear form of $a_1 \times BSN + a_2$. Fluorescein isothiocyanate (FITC)-coupled wheat germ agglutinin (WGA-FITC) is the current standard for chitin labeling. It was used as a reference for comparing the density distribution of the subtracted data and the BSN distribution determined by microscopy.

2.4.2 | Gaussian mixture analysis

Gaussian mixture analysis was performed using the R and sklearn package to evaluate the microscopy-determined BSN distribution's empirical cumulative density fit plots and filtered fluorescence spectra. We assumed that cells with a similar BSN have comparable fluorescence intensities. Therefore, the whole non-normal distributed fluorescence spectra data stems from different overlapping normal distributions. Each normal distribution (Gaussian curve) represents cells of similar age. The number of Gaussian curves was varied between 1 and 15 to determine the best fit to the subtracted fluorescence spectra. For every sample, the probability of the best-fit mixture was calculated and graphically visualized. The minimum Bayesian information criterion (BIC) was used as a criteria to measure the optimum number of curves as well as the curves' center and weight. Each curve represents one age cluster of the yeast cell population. Each curve's location and weight was compared to the content of the empirical cumulative density fit curve for the microscopy-determined BSN distribution.

2.4.3 | Empirical age distribution simulation

A simulation of the age distribution under ideal conditions was performed to confirm a deviation of the age distribution from the theoretical age distribution and create an empirical distribution to compare our data to. Therefore, the generation time of yeast cells was used according to their singular cell age. Wang et al.[49] demonstrated by simulation that daughter cells have the exact duration of the G1 phase as the second-generation mother cells. Using these parameters of Wang et al.,[49] based on experimental data of Adams[50] and Yang et al.[51] for estimated cell cycle length, we simulated the age fraction distributions in the exponential growth phase of yeast cells. Due to the description of the rare event of budding in a large population,[52] as well

as the independent occurrence of budding,[53] we assumed a Poisson process, with n randomly and independent occurring events of a Poisson distribution, corresponding to the "rpois" function[54] in R. λ was taken as N divided by the corresponding cell cycle length, reported by Wang et al.[49]

The initial bud scar distribution was assumed as arithmetic distribution, as reported by Steinkraus et al.,[55] corresponding to 50% daughter cell content, 25% mother cells 1st generation, and so on.

## 2.5 | Transfer of the model using *Pichia pastoris*

Pichia pastoris cells were stained using the His6-Sumo-sfGFP-ChBD protein (S1.2) and then evaluated using flow cytometry. For comparison with the presented protein linker, WGA-FITC was used. Moreover, the cells were analyzed via microscopy, and the specific staining of the bud scars was verified (S1.3). Finally, protein affinity determination was performed using an adsorption isotherm, as described in S1.5.

3 | RESULTS

3.1 | Prediction of Autofluorescence

In order to measure how accurately the fluorescence intensity at 525 nm ($\hat{\beta}_{525}$), which corresponds to the autofluorescence of the yeast cells, could be predicted, a training set of 13 sets (yeast samples inoculated on different days) of 3 yeast cell samples without the protein linker were prepared. In each set of yeast samples, the three samples were inoculated at the same time, prepared, and handled under similar conditions as described in S1.2. The samples were then measured using flow cytometry as described in Section S1.4 and FSC as well as fluorescence signals at 585, 610, and 780 nm were used to predict the fluorescence signal at 525 nm. For each sample RF and PCR were calibrated. Based on this calibration, the coefficient of determination was recorded for the prediction of the fluorescence signal at 525 nm in the sample, as well as the coefficient of determination for the prediction of the other two samples in the set. These will be referred to as the in-sample coefficient of determination (R^2_{IS}) and the out-of-sample coefficient of determination (R^2_{OOS}), respectively.

The PCR was generated for each sample using the two principal components, which generally accounted for most of the total variance within the data. When using PCR, the mean R^2_{IS} for the samples was 0.775 while the mean R^2_{OOS} for the samples was 0.538. Each RF was optimized on the in-sample data using k-fold cross-validation, with k being 3. RF provided a mean R^2_{IS} of 0.95 for the samples and a mean R^2_{OOS} of 0.803. When comparing the results, RF performed significantly better than the PCR both in the sample ($p < 0.01$) and out of the sample ($p < 0.01$). An example of the fit of the prediction can be seen in Figure 2 for both RF (top) and PCR (bottom). A comparison of the subtracted fluorescence spectra of RF (left) and PCR (right) can be found in the Supplementary Section in Figure S1. Therefore, for future autofluorescence prediction, the RF model is used.
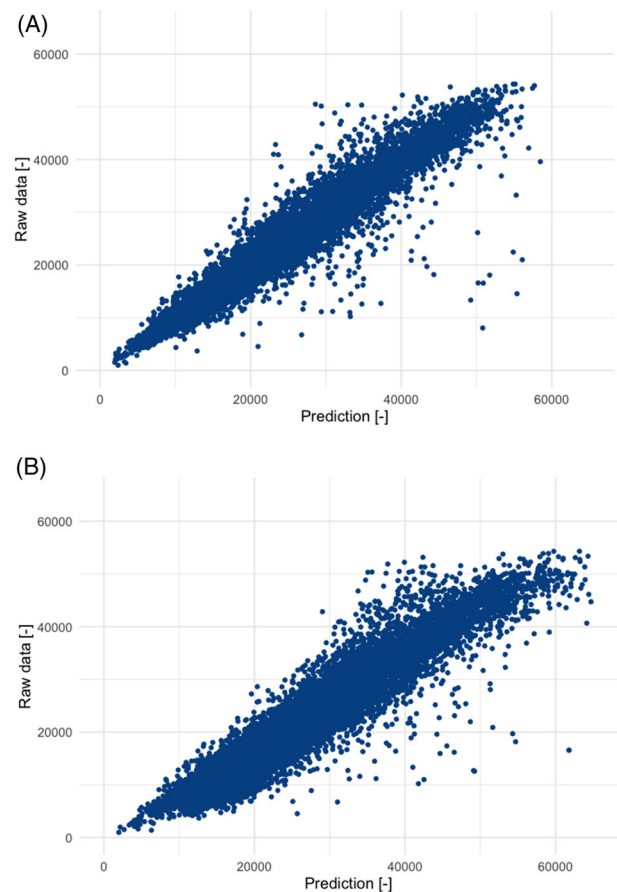

FIGURE 2 Comparison of predicted and raw data for random-forest model (upper) and principal component regression (lower)

3.2 | Model fitting of the microscopy-estimated BSN distribution

After unmixing fluorescence intensities and autofluorescence, resulting in the fluorescence of fluorophores (His6-Sumo-sfGFP-ChBD) on yeast cell bud scars, the question of interpretation of the fluorescence intensity remains. Therefore, a fitting of unmixed fluorescence intensities with the BSN of yeast cells is necessary.

In a previous study,[56] we established a method for bud scar visualization using a fluorescence protein and flow cytometric determination of the corresponding fluorescence-signal intensities. Furthermore, we demonstrated a correlation between these fluorescence intensities and the microscopy-determined median BSN. Median BSN distribution was determined utilizing a relationship between the BSN distribution and cell numbers determined by Powell et al.[57] However, due to variances in cell morphology in yeast populations, the regression parameters determined by Powell et al. cannot be used for every yeast strain. This method does not always result in a perfect prediction of the absolute BSN. For example, it is unrealistic to obtain a total BSN of 28 based on zero visible bud scars, as calculated by the Bayes theorem using the correlation explained earlier. An adjustment of the regression parameters was necessary to obtain valid data for the
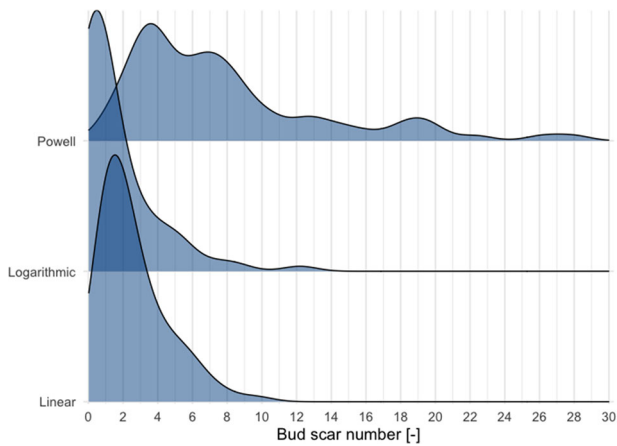
FIGURE 3 Comparison of absolute bud scar numbers predicted by different methods: Powell et al. (top), logarithmic transformation (middle), and linear transformation (bottom)

subsequent comparison of the flow cytometric fluorescence distribution with microscopic data.

To correct the existing parameters of Powell et al.'s method to the given data, we considered linear and logarithmic transformations to obtain the correlation between the number of visible bud scars and yeast cell surface area. The approach to each transformation and the resulting coefficients can be found in S3.1 and S3.2.

Figure 3 compares the predicted BSN distribution based on the parameters reported by Powell et al. with that obtained after linear or logarithmic transformation. Using either transformation to correct the distribution resulted in a large number of daughter cells. The literature suggests that a yeast cell population has a theoretical daughter cell content of 50%.[55] In experiments, a daughter cell content slightly higher than this theoretical value was determined,[38] which was attributed to differences in cell cycle durations with increasing replicative aging.[49] Considering the cell cycle length, a simulation was performed in Section 3.4, confirming the distribution calculated by the logarithmic transformation.

3.3 | Verification of microscopic data by the Gaussian mixture model

In order to determine the BSN distribution from the fluorescence intensities a direct linear correlation of the fluorescence intensities and distributions determined through microscopy was investigated (Figure S2 and S3). This approach proved to be unsuitable as the relationship between fluorescence and BSN varied from batch to batch, see Section S3.3. To account for this, an alternate method was developed using Gaussian mixture models to analyze the unmixed fluorescence intensity distribution and validate the analysis through the Bayesian estimated BSN distribution from Section 3.2 determined through microscopy. The Gaussian mixture model was used to analyze unmixed fluorescence intensities of Section 3.1, where the fluorescence spectra were analyzed independent of each sample's fluorescence intensities.
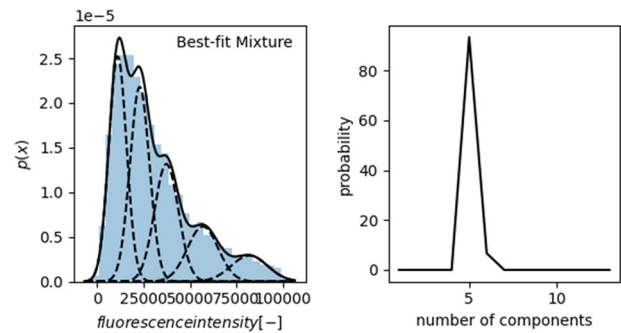

FIGURE 4 Gaussian mixture analysis of unmixed fluorescence spectra. Fluorescence spectra with the best-fitting number of Gaussian peaks (left). Probability that a defined number of peaks, that best explains the spectrum (right)

We assumed that each fluorescence spectra is a mixture of Gaussian bell-shaped curves, each curve indicating cells with similar BSN. This method has the advantage that different yeast strains with varying amounts of chitin can be analyzed because the Gaussian mixture only analyzes the density distributions, not fluorescence intensities. Therefore, the distribution of the bud scars resulting from the logarithmic transformation of Sections 3.2 and S3.2 was verified through the analysis of the subtracted fluorescence data in the absence of the microscopic data. Thus, we developed the following hypothesis: the overall spectrum is a mixture of single Gaussian peaks, and each Gaussian peak explains an age cluster of yeast cells with a similar BSN. Because different yeast populations have different maximum BSNs, we did not fix the number of peaks. We applied the Gaussian mixture analysis to the spectrum to verify this hypothesis. First, the spectrum was analyzed to obtain the optimal number of Gaussian peaks and the corresponding probabilities, as shown in Figure 4 (right). The optimal number of peaks was then used as an overlay for the fluorescence density, as shown in Figure 4 (left). The overall spectrum was best explained using five Gaussian peaks. Therefore, the maximum BSN taken into account by the model was 5. Due to uncertainties in the autofluorescence prediction, a small peak with negative fluorescence intensity was obtained after subtraction. A good correlation was observed between the microscopy-determined number of cells per age cluster and the clusters according to the Gaussian mixture analysis. A similar peaks weight was obtained when the same analysis was performed using WGA-FITC as the staining component. Table 1 shows the weights of different peaks. Notably, the coverage of the data with WGA-FITC showed more peaks than the data with the protein linker. Finally, the autofluorescence was predicted for data subtraction using both the RF and PCR methods. The Gaussian mixture analysis of the subtracted data confirmed the presence of five peaks (Figure S4 and S5). Overall, the results obtained from PCR and RF models were similar and the prediction using both methods are in good agreement with the data from the microscopy measurements.[38,58]

Yeast cells can undergo more than five divisions. Assuming that the age distribution corresponds to the theoretical distribution described by Steinkraus et al.,[55] the daughter cell content was 50%, where 25% of the yeast cells have one bud scar, 12.5% of the cells have two bud scars,

TABLE 1 The weights of each Gaussian peak and microscopy-determined clusters for comparison

Peak	1	2	3	4	5	6
Peak weight RF, corresponding to age fractions	0.553	0.148	0.099	0.062	0.036	-
Peak weight PCR, corresponding to age fractions	0.577	0.206	0.133	0.084	-	-
Peak weight WGA and RF, corresponding to age fractions	0.482	0.166	0.132	0.098	0.067	0.0054
Microscopy-determined age fractions obtained using logarithmic transformation	0.517	0.198	0.121	0.072	0.0479	0.030
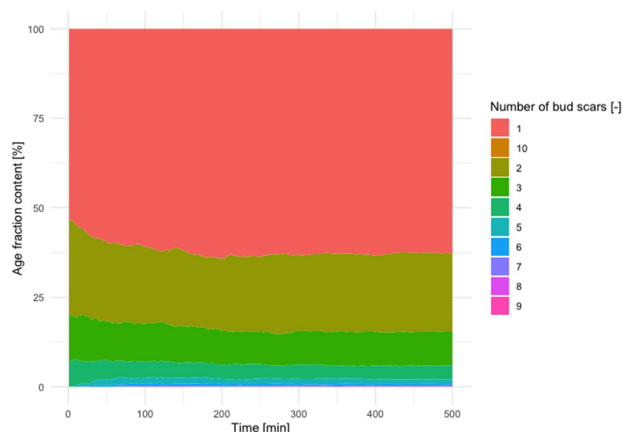
FIGURE 5 Simulation of bud scar number distribution over time. The assumption of the G1 duration parameter was taken from Wang et al. (2017).[49]

and so on. In this case, only a small number of cells (< 3.125%) have more than five bud scars and were therefore considered negligible. To compare our obtained results with an empirical model, we performed a simulation using the cell cycle length of different aged yeast cells.

3.4 | Empirical age distribution simulation

The results obtained from the analysis in Section 3.3 contrast the classical assumption; that is, age fractions are constant in yeast cell populations. It is assumed that under optimal conditions, the daughter cell content is 50%, the mother cell content is 25%, the second-generation content is 12.5%, and so on.[59] A deviation of this assumption can be explained by the difference in the G1 phase, depending on cell age. To test, if the age distributions resulting from Section 3.3 are realistic, we have performed a simulation depending on the age-related differences in G1 phase.

Assuming that cells are in a division state instead of a chronological age phase (stationary phase), Figure 5 illustrates the development of the cell age fraction in the exponential growth phase without the impact of stress conditions. The figure shows that the daughter cell content is > 50%, which is realistic. In the equilibrium phase, a daughter cell content of 62% could be calculated. A prolonged G1 phase is explained by the volume of the daughter cell, which is much smaller than that of mother cells. Thus, daughter cells require more time to enlarge to the required cell volume for division[60] and the daughter cell content increases to > 50%.
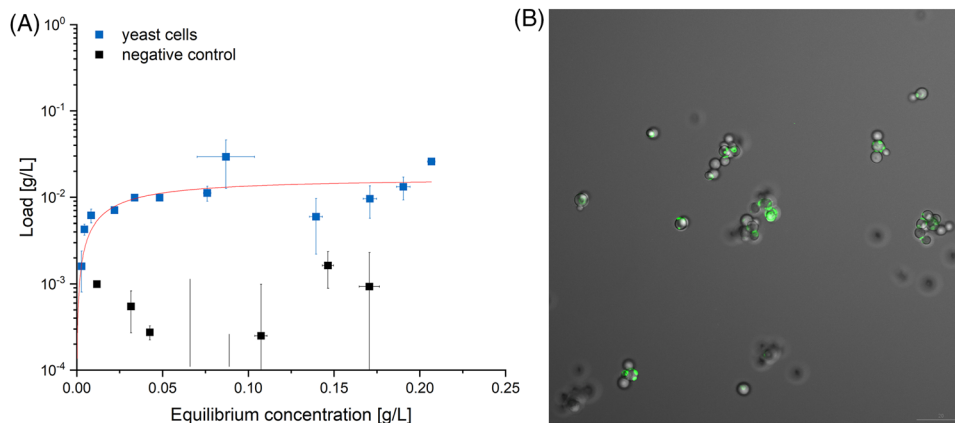
3.5 | Transfer toward *Pichia pastoris*

Pichia pastoris yeast cells are of high interest in the pharmaceutical industry due to the recombinant synthesis of proteins. But less is known about the impact of the singular cell age on cell growth, protein synthesis, or nutrient uptake. Assuming that *P. pastoris* cells have a similar heterogeneity as brewing or bakers' yeast, cell cycle duration and nutrient uptake highly depend on the singular cell age. Based on Krainer et al.,[61] who first determined chitin compounds on *P. pastoris* cell wall, similar staining of bud scars with His6-Sumo-sfGFP-ChBD should be possible.

3.5.1 | Flow cytometry results

First, *P. pastoris* yeast cells were labeled using the His6-Sumo-sfGFP-ChBD protein linker, followed by flow cytometry measurements. For comparison, yeast cells were labeled using WGA-FITC as the commercial standard for chitin labeling and as already published by Krainer et al.[61] In Figure S6, a comparison of the FSC signals of the three samples is given, indicating no visual differences in the particle size and, therefore, cell size due to the binding of protein on the bud scars. The mean FSC value was 1,666,116 for the reference sample, 1,612,097 for the test sample, and 1,456,880 for the WGA-FITC-labeled yeast. However, the intensity of the fluorescence signals at 525 nm varied significantly, as determined by the Kruskal–Wallis test and a subsequent posthoc Dunn test. The adjusted p-value was < 0.0001 for all comparative tests. Further, His6-Sumo-sfGFP-ChBD-labeled yeast cells resulted in fluorescence intensity between 98,400 and 330,123 (1st and 3rd quartile, respectively). In contrast, the WGA-FITC-labeled yeast cells had a lower fluorescence intensity of 23,405 and 84,476 (1st and 3rd quartile, respectively), indicating a lower affinity toward chitin.

3.5.2 | Evaluation of the specific binding parameters

To test this hypothesis of a lower affinity, adsorption isotherms were performed to determine the binding parameters of the His6-Sumo-sfGFP-ChBD protein linker toward *P. pastoris* cells (Figure 6). Compared to *P. pastoris* cells, *E. coli* cells lack chitin components on the cell surface and are therefore a negative control. From this figure, the value of q_{max}, indicating the maximum protein load per cell surface, was determined to be $3.417 \pm 0.738 \times 10^6$ molecules μm^{-2}. In our previous study[56] we determined a maximum load of $2.45 \pm 0.11 \times 10^6$ molecules

FIGURE 6 Left: Adsorption isotherm indicating the specific binding of the protein linker toward Pichia pastoris cells; $N = 3$; Right: Overlay of the transmission and fluorescence channel of Pichia pastoris yeast cells labeled with the His6-Sumo-sfGFP-ChBD protein linker

μm^{-2} for *Saccharomyces pastorianus ssp. carlsbergensis*, indicating a 40% higher protein load per cell for *Pichia pastoris* yeast. This higher protein load indicated a higher protein binding efficiency and therefore more binding positions by a higher chitin content or better steric arrangement of the proteins on the cell surface. *E. coli* cells showed no significant increase of the protein load by increasing equilibrium concentration, indicating no specific protein binding. The dissociation constant of His6-Sumo-sfGFP-ChBD to *P. pastoris* was determined to be 537.95 ± 261.46 nM indicating a lower affinity of WGA-FITC, which was previously determined to be $26.66 \pm 12.08 \mu M$.[56] The microscopic images of labeled *Pichia pastoris* cells are shown in Figure 6b, indicating a specific protein binding due to the ring structure of the staining on the cell surface.

3.5.3 | Age determination of the Pichia pastoris cells

After successfully demonstrating the specific binding of His6-Sumo-sfGFP-ChBD to the cell surface of *P. pastoris* cells, in particular to the bud scars, the unmixing procedure described in this work were applied to determine the age distribution of stained *P. pastoris* cells. The subtracted fluorescence spectra of the *P. pastoris* yeast cells (Figure S7) were fitted with Gaussian curves, resulting in the following age fractions: $34.545 \pm 2.827\%$ (daughter cells), $30.853 \pm 2.524\%$ (mother cells, 1st generation), $15.424 \pm 4.559\%$ (mother cells, 2nd generation), $8.939 \pm 1.446\%$ (mother cells, 3rd generation), and $6.282 \pm 1.288\%$ (mother cells, 4th generation) (Figure S7).

4 | DISCUSSION

The conventional spectral unmixing is expensive, as it requires the use of two lasers: one for exciting the fluorescent and autofluorescence transitions and the second for exciting the autofluorescence transitions. To overcome this limitation, we developed a spectral unmixing method that uses only one laser and different detector signals. To the best of our knowledge, this is the first use of spectral unmixing of flow

cytometry data to analyze single-cell fluorophores on the surface of yeast cells. Other studies have focused on analyzing color deconvolution and autofluorescence in images,[62,63] flow cytometry images,[64] and microscopy.[65] Furthermore, no study that additionally considers the connection between the fluorescence intensities of the stained bud scars and the replicative cell age has used the total fluorescence intensity for data analysis.

We found that PCR gave less precise predictions than RF for spectral unmixing. This supports the previous results[66] that the eigenvector weights from PCR may have no relationship with the predicted variable. In this case, negligible variables become more important than the actual explanatory variables, which degrades the prediction performance. We compared the BSN distributions estimated using the Bayes theorem following different assumptions. The parameters determined by Powell et al. were unsuitable for estimating the absolute BSN. In contrast, the logarithmic transformation gave the most realistic parameters based on the resulting age distribution and the determination of a precise amount of daughter cells, each indicated by a fluorescent birth scar. . A result of a daughter cell content of $> 50\%$ is possible under optimal conditions.[59] as shown by the simulation of age distribution in an exponential growth phase, using age-dependent cell cycle length. Due to these assumptions and the fact that daughter cells need more time till the first budding cycle, justified by the necessary volume increase, the cell division of the daughter cells forms the bottleneck of the cell age distribution of yeast populations. According to Narziß et al.[59] and Steinkraus et al.,[55] a yeast cell can have a maximum BSN of 25 because of the reduced surface area available for budding. Using microscopy, Bühligen et al.[38] showed that a proportion of 4.3% to 7.75% cells had five or more bud scars, similar to our determined content and those in other studies.[58] Hagiwara et al. observed a discrepancy in the theoretical bud scar distribution.[67] They have noticed that nearly no viable cells with six or more bud scars could be detected by microscopy.[67] In further studies, the maximum BSN of yeast cells varies depending on the strain[68,69] and culture media,[67] making it difficult to compare data from the industrial yeast strain used with laboratory strains. Chaudhari et al. focused on fluorescence-activated cell sorting of YPD cultivated yeast cells. They demonstrated that this population has a rare content

of yeast cells with three or more bud scars (12.5%) in the stationary phase.[69] After FACS using WGA-Alexa 488 as bud scar staining fluorophore, the purity of yeast cells with more than one scar is around 65–70%.[69] These limitations of flow cytometry were also reported by Kong et al.[70] Therefore, more recent studies, for example, Kong et al.[70] have determined similar age fraction contents as determined in this study.

To compare the empirical density plots estimated by the direct analysis of the subtracted fluorescence spectra, the Gaussian mixture analysis was performed. The weights of each Gaussian peak were analyzed and compared with the age contents determined by the empirical density plots. The observed deviations were attributed to the stochastic model used for the Gaussian mixture analysis. Furthermore, the fitting of peaks with higher fluorescence intensity by the Gaussian mixture analysis was better than lower intensity peaks because there were fewer overlays at higher intensities than at lower intensities. Both approaches, RF and PCR, obtained peak areas that indicated age clusters similar to the microscopy-determined age fractions. The presented models' limitation is that a reference sample is necessary for each yeast sample. A transfer of the model trained on a specific yeast sample to a different strain of yeast cells showed low regression values. These results confirm that the developed model successfully unmixed autofluorescence and fluorescence signals of the protein linker, enabling fast and reliable age-fraction calculations. Overall, these results confirm the suitability of the logarithmic transformation using the Bayes theorem for BSN estimation, considering the age-dependent increase in cell surface area. Furthermore, the Gaussian mixture analysis gave results similar to that of the Bayes theorem, as confirmed by the simulation.

The transfer of the proposed model to another yeast species like *Pichia pastoris* demonstrated its applicability. We could show that chitin is also present in the bud scars of *P. pastoris* yeast cells and that the used protein linker His6-Sumo-sfGFP-ChBD stains these specific components. Furthermore, we demonstrated that chitin-binding domains have a higher protein load per cell for *P. pastoris* than for *S. pastorianus ssp. Carlsbergensis*, resulting in higher fluorescence intensities and increased detection. Compared to *Saccharomyces* yeast, this higher protein load indicates higher chitin concentrations in the bud scars and therefore higher protein binding capacity. The age-fraction analysis of *P. pastoris* cells currently has some limitations. The first limitation is the aggregation behavior of *P. pastoris* cells, as shown by our microscopic visualization and observed in a previous study.[71] Therefore, the daughter cell content determined by the Gaussian mixture analysis would be higher if *P. pastoris* cells were present in the single-cell stage. This limitation can be overcome by optimizing the staining method; for example, by changing the buffer system or pH. Nevertheless, our results confirm the availability of chitin in *P. pastoris* yeast cells. An in-depth analysis of the age-related protein synthesis in *P. pastoris* cells will be enabled using the chitin-binding protein linker combined with the proposed method for fluorescence unmixing and bud scar determination.

In summary, this approach of unmixing fluorescence intensities combined with Gaussian mixture analysis could be used in further research to elucidate the relationship between the age-fraction content and cell growth or metabolite formation in *Saccharomyces* and *P. pastoris* yeast. This spectral unmixing of autofluorescence data from fluorescence spectra is necessary to prevent false positive data in studies using weak fluorophores or antibodies/proteins detecting small antigen amounts on surfaces. Furthermore, in-depth studies of separation efficiencies are possible without microscopic approaches by detecting BSNs on the whole cell surface by flow cytometry instead of microscopic visible parts.

CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

The raw datasets generated during the current study are available on the flowrepository webpage (https://flowrepository.org/id/FR-FCM-Z5QX).

AUTHOR CONTRIBUTIONS

Marco Eigenfeld: Conceptualization; Data curation; Methodology; Validation; Visualization; Writing – original draft; Writing – review & editing. Roland Kerpes: Conceptualization; Project administration; Writing – review & editing. Iain Whitehead: Conceptualization; Data curation; Validation; Writing – review & editing. Thomas Becker: Conceptualization; Methodology; Project administration; Writing – review & editing

ORCID

Marco Eigenfeld https://orcid.org/0000-0001-8026-4468

REFERENCES

1. Lichten, C. A., White, R., Clark, I. B. N., & Swain, P. S. (2014). Unmixing of fluorescence spectra to resolve quantitative time-series measurements of gene expression in plate readers. *Bmc Biotechnology [Electronic Resource]*, *14*, 11.
2. Gášková, D., Brodská, B., Heřman, P., Večeř, J., Malínský, J., Sigler, K., Benada, O., & Plásek, J. (1998). Fluorescent probing of membrane potential in walled cells: diS-C3(3) assay in Saccharomyces cerevisiae. *Yeast*, *14*, 1189–1197.
3. Surre, J., Saint-Ruf, C., Collin, V., Orenga, S., Ramjeet, M., & Matic, I. (2018). Strong increase in the autofluorescence of cells signals struggle for survival. *Science Reports*, *8*, 12088.
4. Drbohlavova, J., Adam, V., Kizek, R., & Hubalek, J. (2009). Quantum dots – characterization, preparation and usage in biological systems. *International Journal of Molecular Sciences*, *10*, 656–673.
5. Liu, Z., Tabakman, S. M., Chen, Z., & Dai, H. (2009). Preparation of carbon nanotube bioconjugates for biomedical applications. *Nature Protocols*, *4*, 1372–1382.

6. Mahtab, F., Yu, Y., Jacky, W., Liu, J. C., Zhang, B., Lu, P., Zhang, X., & Tang, B. Z. (2011). Fabrication of silica nanoparticles with both efficient fluorescence and strong magnetization and exploration of their biological applications. *Advanced Functional Materials*, *21*, 1733–1740.

7. Kumar, S., & Koh, J. (2012). Physiochemical, optical and biological activity of chitosan-chromone derivative for biomedical applications. *International Journal of Molecular Sciences*, *13*, 6102–6116.

8. Roederer, M., & Murphy, R. F. (1986). Cell-by-cell autofluorescence correction for low signal-to-noise systems: Application to epidermal growth factor endocytosis by 3T3 fibroblasts. *Cytometry*, *7*, 558–565.

9. Steinkamp, J. A., & Stewart, C. C. (1986). Dual-laser, differential fluorescence correction method for reducing cellular background autofluorescence. *Cytometry*, *7*, 566–574.

10. Erickson, M. G., Alseikhan, B. A., Peterson, B. Z., & Yue, D. T. (2001). Preassociation of calmodulin with voltage-gated Ca(2+) channels revealed by FRET in single living cells. *Neuron*, *31*, 973–985.

11. Garini, Y., Young, I. T., & McNamara, G. (2006). Spectral imaging: Principles and applications. *Cytometry, Part A*, *69A*, 735–747.

12. Mihalcescu, I., Van-Melle, G. M., Chelli, B., Pinel, C., & Ravanat, J. (2015). J.-L., Green autofluorescence, a double edged monitoring tool for bacterial growth and activity in micro-plates. *Physical Biology*, *12*, 066016.

13. Pilizota, T., & Yang, Y.-T. (2018). "Do It Yourself" microbial cultivation techniques for synthetic and systems biology: Cheap, fun, and flexible. *Frontier Microbiol*, *9*.

14. Zhang, R., Chouket, R., Plamont, M.-A., Kelemen, Z., Espagne, A., Tebo, A. G., Gautier, A., Gissot, L., Faure, J. D., Jullien, L., Croquette, V., & Le Saux, T. (2018). Macroscale fluorescence imaging against autofluorescence under ambient light. *Light: Science & Applications*, *7*, 97.

15. Bhatta, H., Goldys, E. M., & Learmonth, R. P. (2006). Use of fluorescence spectroscopy to differentiate yeast and bacterial cells. *Applied Microbiology and Biotechnology*, *71*, 121–126.

16. Olsen, L. F., Andersen, A. Z., Lunding, A., Brasen, J. C., & Poulsen, A. K. (2009). Regulation of glycolytic oscillations by mitochondrial and plasma membrane H+-ATPases. *Biophysical Journal*, *96*, 3850–3861.

17. Joniová, J., Gerelli, E., Zellweger, M., & Wagnières, G. (2016). Optimization and characterization of the endogenous production of protoporphyrin IX in a yeast model. *Journal of Biomedial Optics*, *21*, 125008.

18. Warringer, J., & Blomberg, A. (2003). Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in Saccharomyces cerevisiae. *Yeast*, *20*, 53–67.

19. Sheff, M. A., & Thorn, K. S. (2004). Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae. *Yeast*, *21*, 661–670.

20. de Jong, H., Ranquet, C., Ropers, D., Pinel, C., & Geiselmann, J. (2010). Experimental and computational validation of models of fluorescent and luminescent reporter genes in bacteria. *Bmc Systems Biology*, *4*, 55.

21. Patterson, A. P., Booth, S. A., & Saba, R. (2014). The emerging use of in vivo optical imaging in the study of neurodegenerative diseases. *BioMed Research International*, *2014*, 401306–401306.

22. Zulkower, V., Page, M., Ropers, D., Geiselmann, J., & de Jong, H. (2015). Robust reconstruction of gene expression profiles from reporter gene data using linear inversion. *Bioinformatics*, *31*, i71-i79.

23. Wandrey, G., Wurzel, J., Hoffmann, K., Ladner, T., Büchs, J., Meinel, L., & Lühmann, T. (2016). Probing unnatural amino acid integration into enhanced green fluorescent protein by genetic code expansion with a high-throughput screening platform. *Journal of Biological Engineering*, *10*.

24. Crane, M. M., Sands, B., Battaglia, C., Johnson, B., Meinel, L., & Lühmann, T. (2019). In vivo measurements reveal a single 5′-intron is sufficient to increase protein expression level in Caenorhabditis elegans. *Science Reports*, *9*, 9192.

25. Martin, Y., Page, M., Blanchet, C., & de Jong, H. (2019). WellInverter: A web application for the analysis of fluorescent reporter gene data. *Bmc Bioinformatics [Electronic Resource]*, *20*, 309.

26. Taguchi, Y. h. (2016). Principal component analysis based unsupervised feature extraction applied to budding yeast temporally periodic gene expression. *BioData Mining*, *9*, 22.

27. Rajab, J. M., MatJafri, M. Z., & Lim, H. S. (2013). Combining multiple regression and principal component analysis for accurate predictions for column ozone in Peninsular Malaysia. *Atmospheric Environment*, *71*, 36–43.

28. Theis, F. J., Latif, N., Wong, P., & Frishman, D. (2011). Complex principal component and correlation structure of 16 yeast genomic variables. *Molecular Biology and Evolution*, *28*, 2501–2512.

29. Corrigan, A., Russell, N., Welge, M., Auvil, L., Bushell, C., White, B. A., & Murphy, R. A. (2018). The use of random forests modelling to detect yeast-mannan sensitive bacterial changes in the broiler cecum. *Scientific Reports*, *8*, 13270.

30. Franzoni, V., & Kozak, Y., in: Gervasi, O., Murgante, B., Misra, S., Garau, C., Blečić, I., Taniar, D., Apduhan, B. O., Rocha, A. M. A. C., Tarantino, E., & Torre, C. M. (Eds.), *Computational Science and Its Applications – ICCSA 2021*, Springer International Publishing, Cham 2021, pp. 436–447.

31. Phan, Q., DuBois, A., Osborne, J., & Tomasino, E. (2022). Effects of yeast product addition and fermentation temperature on lipid composition, taste and mouthfeel characteristics of pinot noir wine. *Horticulturae*, *8*, 52.

32. Tang, F., & Ishwaran, H. (2017). Random forest missing data algorithms. *Statistical Analysis and Data Mining*, *10*, 363–377.

33. Powell, C. D., Quain, D. E., & Smart, K. A. (2000). The impact of media composition and petite mutation on the longevity of a polyploid brewing yeast strain. *Letters in Applied Microbiology*, *31*, 46–51.

34. Mortimer, R. K., & Johnston, J. R. (1959). Life span of individual yeast cells. *Nature*, *183*, 1751–1752.

35. Johnston, J. R. (1966). Reproductive capacity and mode of death of yeast cells. *Antonie Van Leeuwenhoek*, *32*, 94–98.

36. Eigenfeld, M., Kerpes, R., & Becker, T. (2021). Understanding the impact of industrial stress conditions on replicative aging in Saccharomyces cerevisiae. *Frontiers in Fungal Biology*, *2*.

37. Powell, C. D., Quain, D. E., & Smart, K. A. (2003). The impact of brewing yeast cell age on fermentation performance, attenuation and flocculation. *FEMS Yeast Research*, *3*, 149–157.

38. Bühligen, F., Lindner, P., Fetzer, I., Stahl, F., Scheper, T., Harms, H., & Müller, S., Analysis of aging in lager brewing yeast during serial repitching, 2014.

39. Kuřec, M., Baszczyňski, M., Lehnert, R., Mota, A., Teixeira, J. A., & Brányik, T. (2009). Flow cytometry for age assessment of a yeast population and its application in beer fermentations. *Journal of the Institute of Brewing*, *115*, 253–258.

40. Hotelling, H. (1957). The relations of the newer multivariate statistical methods to factor analysis. *British Journal of Statistical Psychology*, *10*, 69–79.

41. Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R., IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium 2004, pp. 1049–1052 vol.*1042*.

42. Criminisi, A., Konukoglu, E., & Shotton, J. (2011). Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trends® in Computer Graphics and Vision*, *7*.

43. Breiman, L., Friedman, J., Olshen, R., & Stone, C. J., Classification and Regression Trees, Taylor & Francis Ltd, UK 1984.

44. Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*, 18–22.

45. Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

46. R Core Team, R Foundation for Statistical Computing, Vienna 2013.

47. Filzmoser, P., & Geschwandtner, M., 2018.

48. Filzmoser, P., Garrett, R. G., & Reimann, C. (2005). Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences*, *31*, 579–587.

49. Wang, Y., Lo, W.-C., & Chou, C.-S. (2017). A modeling study of budding yeast colony formation and its relationship to budding pattern and aging. *Plos Computational Biology*, *13*, e1005843.

50. Adams, J. (1977). The interrelationship of cell growth and division in haploid and diploid cells of Saccharomyces cerevisiae. *Experimental Cell Research*, *106*, 267–275.

51. Yang, J., Dungrawala, H., Hua, H., Manukyan, A., Abraham, L., Lane, W., Mead, H., & Wright, J., Schneider, B. (2011). Cell size and growth rate are major determinants of replicative lifespan. *Cell Cycle*, *10*, 144–155.

52. Stephenson, F. H., Chapter 3 – Cell growth, in: Stephenson, F. H. (Ed.), Calculations for Molecular Biology and Biotechnology (Second Edition), Academic Press, Boston 2010, pp. 45–81.

53. Forthofer, R. N., Lee, E. S., & Hernandez, M., 5 – Probability Distributions, in: Forthofer, R. N., Lee, E. S., Hernandez, M. (Eds.), *Biostatistics (Second Edition)*, Academic Press, San Diego 2007, pp. 103–133.

54. Ahrens, J. H., & Dieter, U. (1982). Computer generation of poisson deviates from modified normal distributions. *ACM Transactions on Mathematical Software*, *8*, 163–179.

55. Steinkraus, K. A., Kaeberlein, M., & Kennedy, B. K. (2008). Replicative aging in yeast: The means to the end. *Annual Review of Cell and Developmental Biology*, *24*, 29–54.

56. Eigenfeld, M., Kerpes, R., & Becker, T. (2021). Recombinant protein linker production as a basis for non-invasive determination of single-cell yeast age in heterogeneous yeast populations. *RSC Advances*, *11*, 31923–31932.

57. Powell, C. D., Quain, D. E., & Smart, K. A. (2003). Chitin scar breaks in aged *Saccharomyces cerevisiae*. *Microbiology (Reading, England)*, *149*, 3129–3137.

58. Azbarova, A. V., Galkina, K. V., Sorokin, M. I., Severin, F. F., & Knorre, D. A. (2017). The contribution of Saccharomyces cerevisiae replicative age to the variations in the levels of Trx2p, Pdr5p, Can1p and Idh isoforms. *Scientific Reports*, *7*, 13220.

59. Narziß, L., *Abriss der Bierbrauerei*, Wiley-VCH 2017.

60. Johnston, G. C., Pringle, J. R., & Hartwell, L. H. (1977). Coordination of growth with cell division in the yeast Saccharomyces cerevisiae. *Experimental Cell Research*, *105*, 79–98.

61. Krainer, F. W., Gmeiner, C., Neutsch, L., Windwarder, M., Pletzenauer, R., Herwig, C., Altmann, F., Glieder, A., & Spadiut, O. (2013). Knockout of an endogenous mannosyltransferase increases the homogeneity of glycoproteins produced in Pichia pastoris. *Scientific Reports*, *3*, 3279.

62. Rossetti, B. J., Wilbert, S. A., Mark Welch, J. L., Borisy, G. G., & Nagy, J. G. (2019). Semi-blind sparse affine spectral unmixing of autofluorescence-contaminated micrographs. *Bioinformatics*, *36*, 910–917.

63. Landini, G., Martinelli, G., & Piccinini, F. (2020). Colour deconvolution: Stain unmixing in histological imaging. *Bioinformatics*, *37*, 1485–1487.

64. Jiménez-Sánchez, D., Ariz, M., Morgado, J. M., Cortés-Domínguez, I., & Ortiz-de-Solórzano, C. (2019). NMF-RI: Blind spectral unmixing of highly mixed multispectral flow and image cytometry data. *Bioinformatics*, *36*, 1590–1598.

65. Megjhani, M., Correa de Sampaio, P., Leigh Carstens, J., Kalluri, R., & Roysam, B. (2017). Morphologically constrained spectral unmixing by dictionary learning for multiplex fluorescence microscopy. *Bioinformatics*, *33*, 2182–2190.

66. Artigue, H., & Smith, G. (2019). The principal problem with principal components regression. *Cogent Mathematics & Statistics*, *6*, 1622190.

67. Hagiwara, T., Ushimaru, T., Tainaka, K.-i., Kurachi, H., & Yoshimura, J. (2011). Apoptosis at inflection point in liquid culture of budding yeasts. *Plos One*, *6*, e19224.

68. Molon, M., Woznicka, O., & Zebrowski, J. (2018). Cell wall biosynthesis impairment affects the budding lifespan of the *Saccharomyces cerevisiae* yeast. *Biogerontology*, *19*, 67–79.

69. Chaudhari, R. D., Stenson, J. D., Overton, T. W., & Thomas, C. R. (2012). Effect of bud scars on the mechanical properties of *Saccharomyces cerevisiae* cell walls. *Chemical Engineering Science*, *84*, 188–196.

70. Kong, Y., Zhao, Y., Yu, Y., Su, W., Liu, Z., Fei, Y., Ma, J., & Mi, L. (2022). Single cell sorting of young yeast based on label-free fluorescence lifetime imaging microscopy. *Journal of Biophotonics*, 1–10.

71. Lima-Pérez, J., Rodríguez-Gómez, D., Loera, O., Viniegra-González, G., & López-Pérez, M. (2018). Differences in growth physiology and aggregation of Pichia pastoris cells between solid-state and submerged fermentations under aerobic conditions. *Journal of Chemical Technology & Biotechnology*, *93*, 527–532.

SUPPORTING INFORMATION

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