Article

Evaluation of Long-Term Fermentation Performance with Engineered *Saccharomyces cerevisiae* Strains

Maarten L. De Mol [1], Victoria Marcoen [1], Isabelle Maryns [1], Nico Snoeck [1], Joeri J. Beauprez [2], Sofie L. De Maeseneire [1,*] and Wim K. Soetaert [1]

[1] Centre for Industrial Biotechnology and Biocatalysis (InBio.be), Department of Biotechnology, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium; maarten.demol@ugent.be (M.L.D.M.)
[2] Inbiose NV, Technologiepark 3, Zwijnaarde, Bus 41, B-9052 Ghent, Belgium
* Correspondence: sofie.demaeseneire@ugent.be

Abstract: The performance of a microbial fermentation on an industrial scale is subjected to the robustness of the strain. Such strains are genetically engineered to optimize the production of desired compounds in minimal time, but they often fail to maintain high productivity levels for many generations, hindering their effective application in industrial conditions. This study focused on assessing the impact of genomic instability in yeasts that were engineered to produce a fluorescent output by incorporating a reporter gene at one or more genomic locations. The fermentation performance of these strains was evaluated over 100 generations in a sequential batch set-up. In order to bridge the gap between strain engineering and industrial implementation, we proposed the use of novel, host-specific parameters to standardize the strain robustness and evaluate potential improvements. It was observed that yeasts carrying multiple copies of the reporter gene exhibited a more pronounced decrease in output, and the genomic integration site significantly influenced the production. By leveraging these new, host-specific parameters, it becomes possible to anticipate strain behavior prior to incurring substantial costs associated with large-scale production. This approach enhances the economic viability of novel microbial fermentation processes and narrows the divide between laboratory findings and industrial applications.

Keywords: biotechnology; fermentation; robustness; yeast; process parameters; strain engineering; genomic integration

1. Introduction

The global fermentation-based chemical industry was valued at USD 71.16 billion in 2022 and is expected to grow [1]. The use of renewable resources as feedstocks and the development of biorefineries to establish a circular economy significantly boosts this growth [2]. While such renewable resources often aim to supply commodity chemicals for the chemical, polymer, coating, agricultural, food/feed, and fuel industries, biobased specialty chemicals for the pharmaceutical industry, for example, can be synthesized by fermentation, often without the use of bulk renewable resources [3–8]. Over the past decade, many microorganisms have been investigated with regard to their industrial robustness, genetic tractability, and production capacity to develop industrially viable fermentation processes [9,10]. The baker's yeast *Saccharomyces cerevisiae* is of particular interest due to its historic track record in beer, wine, and bread production, placing it on the forefront of the fermentation-based industry. In addition to its traditional use, at present, this yeast is employed mainly in bio-ethanol production (e.g., Cargill [11]), but other industrial applications are also surfacing, such as the production of artemisinin (Amyris [12]), malonic acid (Lygos [13]), and opioids (Antheia [14]), and other functional molecules, such as resveratrol, nootkatone, and stevia sweeteners (Evolva [15]).

In order to obtain an economically viable fermentation process, the yeast strain often needs modification. Metabolic engineering and synthetic biology have allowed us to optimize its production [16–18]. For example, the adaptation of baker's yeast to efficiently produce ethanol from lignocellulosic biomass required several genetic modifications to allow the efficient fermentation of the sugars L-xylose and L-arabinose [19,20]. Adapting the yeast's metabolic profile, for example, by introducing heterologous genes, can occur through the transformation of a plasmid or through recombination into the chromosomal DNA, the latter being more robust, as plasmids are self-replicating mobile elements that need selection pressure to be maintained and that can occur in multiple (variable) copies in the yeast's nucleus. The maintenance of this extrachromosomal DNA often gives rise to a greater metabolic burden [21], resulting in a production decrease and fluctuations [22]. The prerequisite of a selection marker, be it auxotrophic or antibiotic, is also a disadvantage for industrial processes, as it is not compatible with several waste stream materials [23]. A viable yeast strain for industrial fermentation requires modifications of the genome itself. The question, hence, arises as to where in the genome these alterations should be established. Although information is available regarding the output level obtained at certain genomic locations [24], no information can be found regarding the robustness in time of these output levels. Generally, a lack of robustness originating from host genomic instability is perceived only at a point where economic feasibility is in jeopardy. Indeed, there is a gap between the optimization of a microbial producer engineered in the lab and its final application in industry [25].

We believe this gap can be decreased by implementing host parameters alongside well-known process parameters, such as titer, yield, and productivity. These host parameters, introduced in this paper, are based on the behavior of the modified microorganism, characterize the microbial host's (in)stability, and they allow the visualization of potential improvement in the production capacity. To do so, we employed the reporter gene yeast-enhanced cyan fluorescent protein (yECFP), which synthesizes a fluorescent protein gene. The number of yECFP proteins is positively correlated with the fluorescent signal, thus allowing us to track the production and visualize (the lack of) strain robustness. The yECFP reporter gene was introduced in a single- to multi-copy set-up on several genomic loci for over 100 generations, uncovering beneficial and detrimental effects. The 100 generations were set as a target to mimic the time required for a yeast-based industrial production process [12]. In practice, the number of generations is dependent on the overall process developed, with factors such as the seed train, production volume, or type of fermentation. Our preliminary study reveals the effect of different genomic locations, both on an output and a stability level. To the best of our knowledge, such host parameters and stability effects have not been examined previously, although they are of vital importance to further exploit yeast, and by extension, to other microorganisms as a modern production host for chemically complex molecules.

2. Materials and Methods

2.1. Materials

Tryptone, yeast extract, yeast nitrogen base without amino acids, and Agar NOBLE were procured from LED-Techno (Heusden-Zolder, Belgium), Cargill (Gent, Belgium), and Becton Dickinson (Erembodegem, Belgium), respectively. Complete supplement mixture was purchased from MP Biomedicals (Groot-Bijgaarden, Belgium). Glucose and fructose were procured from VWR (Leuven, Belgium), while galactose was obtained from Carbosynth (Compton, UK). All other chemicals were purchased from Sigma-Aldrich (Diegem, Belgium), unless stated otherwise.

Oligonucleotides were purchased from Integrated DNA Technologies (Leuven, Belgium). Sequencing services were conducted by Macrogen (Amsterdam, The Netherlands) or LGC Genomics (Berlin, Germany). Deoxynucleotides, agarose, and ethidium bromide were purchased from ThermoFisher Scientific (Aalst, Belgium). QIAprep Spin Miniprep kit

(QIAgen, Antwerp, Belgium) was used for plasmid isolation. Analytik Jena kits were used for PCR purification (Jena, Germany).

Amplification of DNA fragments for cloning, sequencing, or transformation purposes was performed by the high-fidelity DNA polymerase PrimeSTAR HS (Takara, Westburg, Leusden, The Netherlands). In the case of verification of *E. coli* colonies following electroporation, standard Taq polymerase from NEB (Bioké, Leiden, The Netherlands) was used. Yeast colonies underwent a freeze/microwave/freeze cycle prior to PCR. Polymerases used for checking *S. cerevisiae* colonies were standard Taq (NEB, Ipswich, MA, USA), OneTaq® 2X Mastermix (NEB), and EmeraldAmp (Takara, Saint-Germain-en-Laye, France).

2.2. Strains

E. coli One Shot Top10 Electrocomp™ cells (ThermoFisher Scientific) were used for cloning purposes only, as were electrocompetent MDS™ 42LowMut Δ*recA* Scarab cells (Scarab Genomics LLC, Madison, WI, USA). *Saccharomyces cerevisiae* S288C (*MATa HIS3Δ1 LEU2Δ0 LYS2Δ0 URA3Δ0*) BY4742 was obtained from Euroscarf (University of Frankfurt, Frankfurt am Main, Germany) and used as the wild-type strain.

2.3. Media and Culture Conditions

The culture medium lysogeny broth (LB) consisting of 1% tryptone, 0.5% yeast extract, and 0.5% sodium chloride (VWR) was used for growth of *E. coli*. Lysogeny broth agar (LBA) was similarly composed, with the addition of 12 g·L^{-1} agar (Biokar Diagnostics, Pantin Cedex, France). Media were supplemented with the antibiotics ampicillin (100 μg·mL^{-1}) or kanamycin (50 μg·mL^{-1}) for plasmid maintenance.
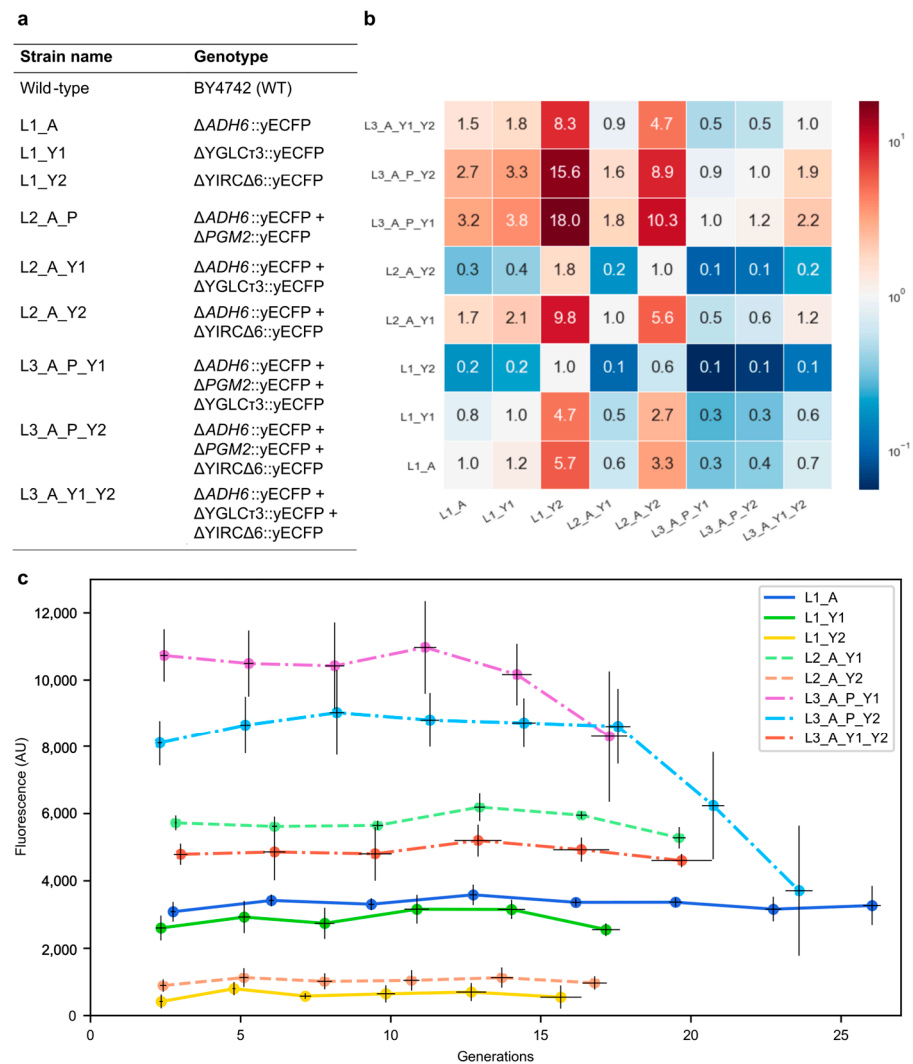
Growth medium for *S. cerevisiae* was based on the synthetic-defined yeast medium with complete supplement mixture (SD-CSM), containing 6.7 g·L^{-1} yeast nitrogen base without amino acids, 20 g·L^{-1} Agar NOBLE when solid cultures are used, 22 g·L^{-1} fructose, 0.22 g·L^{-1} glucose, 0.22 g·L^{-1} galactose, and 0.79 g·L^{-1} CSM. The addition of CSM varied depending on which auxotrophies were needed in the growth medium. All strains were incubated at 30 °C and shaken at 200 rpm.

2.4. Strain Construction

Homologous regions (HR) to guide the yECFP transcription unit to the desired locus were amplified from yeast genomic DNA with the PrimeSTAR HS polymerase, according to the manufacturer's protocol. This polymerase was also used for Circular Polymerase Extension Cloning (CPEC) to clone the homologous regions into an *E. coli* storage vector backbone, with an annealing temperature of 55 °C and an elongation time of 1 min/kb [26]. Colonies obtained after electroporation were verified by PCR with OneTaq® 2X mastermix (manufacturer's protocol) prior to cryovial storage and plasmid isolation. The yECFP gene was ordered from IDT as gBlock. The *TEF1* promoter and *CYC1* terminator are native to *Saccharomyces cerevisiae* and were amplified from genomic DNA similar to the HR. The transcription unit was assembled by CPEC and replaced the coding sequence of the target location within the HR plasmids. Downstream of the yECFP transcription unit, a selection marker was incorporated to select correct transformants after genomic integration.

From these assembled plasmids, linear DNA was amplified, encompassing the homologous regions, transcription unit, and selection marker, with the high fidelity PrimeSTAR HS polymerase and introduced into yeast cells according to the high efficiency Gietz transformation [27]. Sequences of the linear integration cassettes are provided in Supplementary Figure S1 and Supplementary Table S1. Colonies were verified by EmeraldAmp (manufacturer's protocol) to confirm the replacement of the coding sequence with the yECFP transcription unit and selection marker. The presence of LoxP sites neighboring the selection marker allowed its excision by the Cre recombinase. This recombinase was introduced in the yeast strains on a plasmid (pSH47, Euroscarf), also according to the Gietz transoformation. Galactose induction to express the Cre recombinase was performed overnight with a galactose concentration of 20 g·L^{-1}. In the morning, a 10^{-5} dilution series

was plated on medium selective for pSH47 (SD-CSM lacking uracil) but not selective for the marker present in the genome. Colonies were subsequently checked for presence/absence of growth on medium selective for the genomic marker by spotting them on (non-)selective SD-CSM plates. Correct colonies were verified by PCR, and pSH47 was removed by means of counterselection with 5-fluoroorotic acid (1 g·L^{-1}). An overview of all strains is given in Figure 1.

a

Strain name	Genotype
Wild-type	BY4742 (WT)
L1_A	Δ*ADH6*::yECFP
L1_Y1	Δ YGLCτ3::yECFP
L1_Y2	Δ YIRCΔ6::yECFP
L2_A_P	Δ*ADH6*::yECFP + Δ*PGM2*::yECFP
L2_A_Y1	Δ*ADH6*::yECFP + Δ YGLCτ3::yECFP
L2_A_Y2	Δ*ADH6*::yECFP + Δ YIRCΔ6::yECFP
L3_A_P_Y1	Δ*ADH6*::yECFP + Δ*PGM2*::yECFP + Δ YGLCτ3::yECFP
L3_A_P_Y2	Δ*ADH6*::yECFP + Δ*PGM2*::yECFP + Δ YIRCΔ6::yECFP
L3_A_Y1_Y2	Δ*ADH6*::yECFP + Δ YGLCτ3::yECFP + Δ YIRCΔ6::yECFP

Figure 1. Fluorescence output of the first sequential batch experiment. (**a**) Overview of *Saccharomyces cerevisiae* strains used in this study. 'L1_A', for example, indicates a '1 copy of yECFP at locus *ADH6*' strain. A: *ADH6* locus, P: *PGM2* locus, Y1: YGLCτ3 locus, Y2: YIRCΔ6 locus. Δ stands for knock-out of the subsequent gene or retrotransposon, while : indicates the knock-in of yECFP at this location. (**b**) Heatmap of relative output levels among strains. The strains on the vertical axis are compared with the strains on the horizontal axis. Strain L1_Y1, for example, has an output 4.7 times the output of strain L1_Y2. Exact ratios are given in Supplementary Table S2. (**c**) Output profiles of different yeast strains over time with 8 replicates for each strain. Fluorescence, given in arbitrary units (AU), was corrected for both wild-type fluorescence and medium background. The generations were calculated based on the relationship between optical density and cell number, while the error bars represent the standard deviation.

2.5. In Vivo Fluorescence Measurements

Yeast strains were grown in a black 96-well plate (Greiner, Vilvoorde, Belgium), where the medium was refreshed approximately every 12 h (sequential batch set-up). The medium

used for all strains was SD-CSM medium devoid of auxotrophies. Prior to the start of the experiment, strains were grown in two batches to allow for medium adaptation. For every batch, 10 µL (5%) grown culture was used to inoculate the next batch. Plates were continuously shaken at 30 °C in an orbital mode. At the start and end of each batch, the optical density (OD) was measured at 600 nm, as well as the yECFP fluorescence (435 nm/485 nm) with the Tecan Infinite machine 200 Pro. Prior to measurement, the 96-well plate was shaken for 60 s, with an amplitude of 2 mm and a frequency of 280.8 rpm. The gain for yECFP measurement was manually set at 100.

2.6. Calculations

The number of generations a culture reached was estimated through the correlation of the optical density and cell number:

$$n = \frac{\log(c) - \log(C)}{\ln(2)} \tag{1}$$

where n represents the number of generations, c the number of cells at the end of the batch, and C the number of cells at the beginning of the batch. A correlation between the colony-forming units and the optical density was used to calculate the number of cells in the culture (Supplementary Figure S2).

The G_{ST} value was obtained by performing a two-tailed standard *t*-test of the yECFP/OD$_{600}$ compared against all previous data points. In cases where the null hypothesis H_0: $\mu_1 = \mu_2$ was rejected ($p \leq 0.05$), the previous data point was chosen as G_{ST} value. Prior to the two-tailed standard *t*-test, the normal distribution was checked by means of the Shapiro–Wilk test and an F-test for homo- or heteroscedasticity (H_0: $\sigma_1^2 = \sigma_2^2$).

3. Results

To investigate the production stability of genetically engineered baker's yeast over time (generations), the yeast-enhanced cyan fluorescent protein (yECFP) reporter gene was introduced into the genome of the wild-type yeast BY4742 at one or more locations (Figure 1). A first experiment was conducted by examining several strains carrying one to three copies of yECFP on different genomic locations, potentially giving rise to genetic instability. The coding sequences or retrotransposons present at these locations were knocked out, and a transcription unit for yECFP was knocked in simultaneously. First, the alcohol dehydrogenase gene *ADH6*, located on chromosome 13, was chosen, as industrial processes often want a maximal product yield and, thus, no loss of carbon through side products such as ethanol [28]. Aside from this gene, two retrotransposon locations, YIRCΔ6 (chromosome 9) and YGLCτ3 (chromosome 7), were examined. Retrotransposons are mobile elements that can propagate in the genome, thus potentially giving rise to strain instability [29]. Both retrotransposon locations were previously investigated pertaining to their initial output level by Flagfeldt and colleagues [24]. As a preliminary study, experiments were carried out in a sequential batch set-up on a 96-well plate with medium replacement every 12 h to allow for sufficient replicates and to minimize the effects of process conditions.

As can be observed from Figure 1, differences in yECFP expression between single-copy strains were present, with the *ADH6* and YGLCτ3 locations resulting in a higher expression level than the YIRCΔ6 location. The effect of a second copy was examined by introducing the same TU for yECFP at the *ADH6* locus in strains L1_Y1 and L1_Y2. Elevated fluorescence levels were observed, although the fluorescence of strain L2_A_Y2 was only slightly enhanced. Moreover, the double-copy strain L2_A_Y2 had a lower fluorescence level than the single-copy strain L1_A. The selection of the YIRCΔ6 locus for efficient production in an industrial set-up is, thus, not recommended. In the case of the Y1 locus, the addition of an extra copy had almost exactly doubled the output (2.06 times L1_Y1), as expected. Increasing the copy number to three revealed several interesting fluorescence profiles. We selected the phosphoglucomutase locus *PGM2* (located on chromosome 13),

which was active in the central carbon metabolism [30], as an additional location to examine intrachromosomal stability; this gene was located on the same chromosome as *ADH6*. The strains L3_A_P_Y1 and L3_A_P_Y2 could, thus, be compared to L3_A_Y1_Y2, which contained three TUs on three different chromosomes. At first glance, a lower output of the latter strain compared with the intrachromosomal strains was observed, confirming the fact that the Y2 locus is not recommended. Adding a third copy of the TU at the *PGM2* locus elevated the output 1.8-fold compared with the L2_A_Y1 combination. However, compared with the L2_A_Y2 combination, the improvement was 9-fold. These results indicated that combinations were not predictable, as synergistic and antagonistic effects could be at play.

Aside from the initial fluorescence of yECFP on these loci, the stability of fluorescence was examined over 15 to 25 generations. For the single-copy strains, a stable output profile was observed within this timeframe. In addition, both two-copy strains showed a robust output over 15 generations or longer. Concerning the three copy strains, a more stable output pattern of the interchromosomal strain (L3_A_Y1_Y2) compared with both intrachromosomal strains (L3_A_P_Y1 and L3_A_P_Y2) was observed.
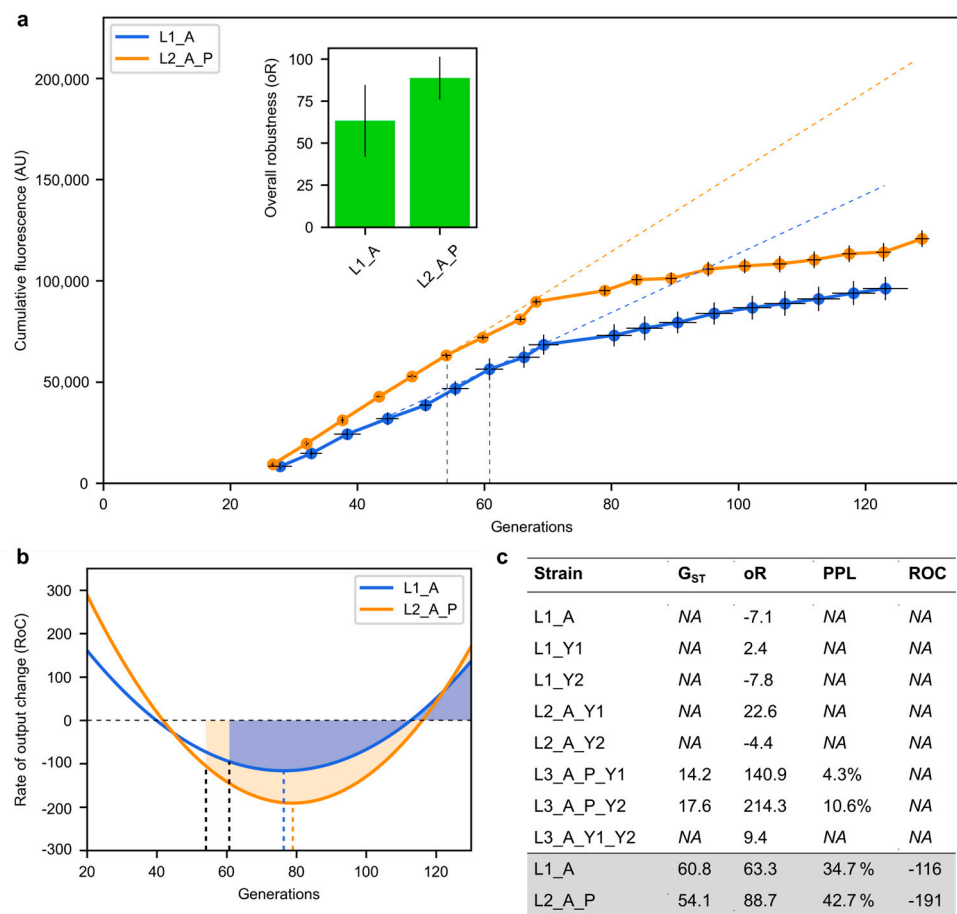
Based on the output profile, several novel parameters could be introduced with the aim to standardize the microbial host assessment pertaining to the stability of their output prior to scale-up to industrial fermentations. At a certain generation, it could be determined that this output deviated from all previous data points. We named the point prior to this the 'last generation of stability', the G_{ST} value, representing the generation from which the output digresses from the initial output. Taking into account the normal distribution of the replicates and homo- or heteroscedasticity, we used a two-tailed standard *t*-test to calculate this parameter for strains L3_A_P_Y1 and L3_A_P_Y2. Their G_{ST} values were 14.2 and 17.6, respectively. In practice this means that the output on these loci was stable until these generations, and host instability occurred from then on. Concerning the other engineered yeast strains, no instability was observed in the first 15 to 25 generations, and, thus, no G_{ST} value was calculated.

Moreover, the output of a microbial production host during a fermentation process should also be evaluated on a more global scale, in line with the process parameter 'productivity'. Indeed, the 'overall microbial robustness parameter' (oR) can be introduced; this parameter is calculated as the difference in output between the beginning and the end of the fermentation divided by the number of generations in between. Ideally the output level of a microbial production host should not decrease over time. The desired oR values, thus, center around zero (stable profile) or are negative (output increasing over time). An oR value around zero was calculated for all strains except L3_A_P_Y1 and L3_A_P_Y2. These three-copy yeast strains, showing instability at some point during the fermentation, had positive oR values, indicating a decrease in output over time. The general trend was, thus, an elevated output instability for the higher-copy strains where the transcription units were located on the same chromosome. The (lack of) robustness observed with the G_{ST} values was, thus, confirmed in this oR value.

Although being the most unstable production strains, the three-copy strains overall had the highest output, which could be concluded by looking at their cumulative output (Supplementary Figure S3). Considering this cumulative production, a straight line over the generations would be the desirable outcome. This was the case for all strains except for L3_A_P_Y1 and L3_A_P_Y2, indicating that these strains lost potential production capacity over generations. This can best be visualized by plotting a straight line based on the slope of all data points until the G_{ST} value in an output-versus-generations plot is reached, hereafter referred to as the maximal production capacity of a certain strain. In such a cumulative output graph, the effects of heterogeneity are visualized, as well as the extent of 'potential product loss' (PPL). The latter was calculated as the difference between the maximal production capacity and the actual production output at a given time point. In this case, the strains L3_A_P_Y1 and L3_A_P_Y2 had PPLs of 4.3% and 10.6%, respectively, compared with the maximal production capacity at the final data point. The

PPL can be used as a measure for potential strain improvement regarding strain stability. In addition to visualizing the important PPL parameter, the cumulative plot is less susceptible to variations in production. The standard deviations presented in Figure 1 were calculated on eight biological replicates on the same time point, while the standard deviations from the cumulative graph were calculated by the sums of each replicate until that time point.

The presence of three copies of a gene spread over the genome of the baker's yeast clearly had its effect on the heterogeneity. However, one can expect a similar effect of strains carrying two copies, or even a single copy, given the burden they can cause. To study this, a second sequential batch experiment was performed for over 100 generations with the strains L1_A and L2_A_P. The latter strain contained two copies of the yECFP transcription unit on the same chromosome. The G_{ST} values for these strains were calculated as mentioned above and were 60.8 for L1_A and 54.1 for L2_A_P. These G_{ST} values were higher than those of the three-copy strains; this was in agreement with the previous experiment, where it was shown that increased copy numbers led to increased instability. The cumulative output of both strains is given in Figure 2. Similar to the first experiment, the actual cumulative titers were lower than the estimated maximal production capacity, indicating room for improvement during the strain optimization pertaining to robustness. The PPL of strain L1_A was 34.7% at generation 123, while that of L2_A_P was 42.7% at generation 129.

Strain	G_{ST}	oR	PPL	ROC
L1_A	NA	-7.1	NA	NA
L1_Y1	NA	2.4	NA	NA
L1_Y2	NA	-7.8	NA	NA
L2_A_Y1	NA	22.6	NA	NA
L2_A_Y2	NA	-4.4	NA	NA
L3_A_P_Y1	14.2	140.9	4.3%	NA
L3_A_P_Y2	17.6	214.3	10.6%	NA
L3_A_Y1_Y2	NA	9.4	NA	NA
L1_A	60.8	63.3	34.7%	-116
L2_A_P	54.1	88.7	42.7%	-191

Figure 2. Fluorescence output of the second sequential batch experiment. (**a**) Cumulative yECFP output of strains L1_A and L2_A_P (n = 4) during the second sequential batch experiment. The dotted lines represent the maximal production capacity, based on the slope and intercept of the initial and G_{ST} time point. Inset: Overall robustness values (oR) for both strains. Fluorescence, given in arbitrary units (AU), was corrected for both wild-type fluorescence and medium background. The generations were calculated based on the relationship between optical density and cell number, while the error bars represent the standard deviation. (**b**) Rate of output change (ROC) graphs for both

strains during the second sequential batch experiment. The ROC was calculated as the first derivative of the third-degree polynomial fit of the output over generations data. The dashed black lines indicate the G_{ST} time point, from which the potential product loss (PPL; shaded area) is visualized. The dashed color lines indicate the vertices of these parabolic derivatives. (**c**) Values of the novel parameters for each strain tested in this manuscript. Grey shading indicates strains during the second sequential batch experiment. G_{ST}: last generation of stability, oR: overall robustness, PPL: potential product loss, ROC: rate of output change, NA: not applicable.

Aside from determining the G_{ST} value, the overall robustness parameter oR, and the potential product loss (PPL), the 'rate of output change' (ROC) was also important. In order to visualize this change, the output over generations data was fitted to a third-degree polynomial from which the first derivative was taken, resulting in a parabolic plot. As we have observed from the second sequential batch experiment, the initial fluorescence was stable until the G_{ST}, declined afterwards, and 'stabilized' at a lower output, which conformed to our parabolic first derivative. This parameter is represented in Figure 2 for strains L1_A and L2_A_P during the second sequential batch experiment. These plots give an in-depth description of the course of the output during the fermentation, as the curvature of the parabola is representative for the robustness of a strain's output. A robust strain has a rate of output change near the zero-line, indicating that the single-copy strain L1_A was more robust than the double-copy strain L2_A_P, as was concluded earlier. This curvature of the parabola could be represented by the rate of output change at the vertex, resulting in -116 for L1_A and -191 for L2_A_P. A higher ROC value for a strain, thus, indicates a more robust strain. Aside from determining the rate of output change of a strain, the potential product loss (PPL) was also visualized by the area between the curve and the zero-line from the G_{ST} time point on. An overview of the values of the novel parameters for each strain during both experiments is given in Figure 2.

4. Discussion

To the best of our knowledge, no analogous studies have been performed, and no parameters pertaining to strain robustness and population heterogeneity for industrial fermentation processes have been described. In the case of *E. coli*, Goormans et al. investigated 49 intergenic regions throughout its genome pertaining to their expression capacity [31], while Flagfeldt and colleagues investigated the expression level of 20 genomic locations of baker's yeast by using LacZ as a reporter gene and observed differences among integration sites [24]. Our study incorporated two of those genomic locations, the retrotransposon sites YIRCΔ6 and YGLCτ3, and employed the same *TEF1* promoter and *CYC1* terminator for the expression of the reporter gene but differed in the reporter gene (yECFP instead of LacZ), the growth medium (defined SD-CSM compared with complex YPD), and the wild-type yeast used (S288c versus CEN.PK113-5D). Flagfeldt et al. noted a comparable expression level of LacZ at the YIRCΔ6 and YGLCτ3 integration sites, while our study clearly indicated a higher yECFP fluorescence for YGLCτ3. This discrepancy could be due to experimental differences, such as complex versus defined media. However, the variations of the single-copy strains at the start of the fermentation (Figure 1b) were in the range with the maximal 8.7-fold difference published by Flagfeldt and colleagues.

In addition to the initial output of (single-copy) strains, the microbial population should homogenously express the desired protein or compound to ensure reproducibility and output stability, which is crucial for industrial processes. As was observed in this study, an increasing number of gene copies or generations had a negative effect on the yECFP output of the yeast strains. While various biological factors, such as mutations and epigenetics [32], or physicochemical factors, such as environmental spatiotemporal changes [33], could contribute to varying extents to this decrease in output, the baker's yeast's native homologous recombination machinery was highly efficient [34] and could play a major role. It was, thus, essential to take into account the genetic tractability of microorganisms and their robustness pertaining to industrial fermentation processes. Our results confirmed previous expectations that a higher-copy-number strain is more

likely to show more variation in output, as it is more prone to genetic instability [35]. The introduction of heterologous genes and the knock-out of native genes introduces a burden upon the cells, which is being counter-selected by growth [36]. Therefore, cells diminishing their target output have an advantage over cells spending resources on this target output, resulting in biological heterogeneity. During DNA replication, for example, double-stranded DNA breaks can, in a multi-copy strain, be wrongly repaired by one of the additional copies, altering the genome of the cell [35]. This type of heterogeneity was observed not only for our fluorescent output but also for most other industrial processes, as the microbial hosts are being pushed towards maximal production of a desired compound instead of towards maximal growth [37].

As this was a preliminary study with a simple, fluorescent reporter gene to showcase the need for strain robustness and the characterization thereof, future studies should elaborate on these findings and validate them with more representative genomic modifications. The yECFP is a short protein, requiring plain post-translational modifications and structural folding [38]. On the one hand, more complex proteins, such as monoclonal antibody fragments, could be used to investigate the effect of production over time and to validate the strain-specific parameters postulated in this manuscript. On the other hand, the biosynthesis of metabolites might tax the microbial cell in a different way than do proteins [39]. Therefore, biochemical pathways of β-carotene or violacein, for example, two pharmaceutically interesting pigments requiring a multi-gene pathway for their biosynthesis [40,41], could be introduced at one or more genomic locations. Due to the increased complexity of biochemical pathways compared with single transcription units, the lack of robustness in output is expected to be more pronounced for such studies.

A final remark on the employed set-up revolves around the fermentation conditions. This study was performed on a very small scale with 96-well plates, while industrial fermentations require large capacities, meaning that the conditions experienced by the yeast strains differ significantly from bioreactor conditions. Typically, microbial populations are more strictly controlled in a bioreactor environment, as the pH, dissolved oxygen, agitation, and respiration are monitored and could be used to steer the fermentation process. In addition, the shear stress experienced by the microorganisms in a bioreactor is higher due to the mechanical stirring, and spatiotemporal heterogeneity often occurs due to compartmentalization in the vessel, leading to fluctuating concentrations of nutrients, dissolved gases, and metabolites [42]. Nevertheless, this preliminary study clearly demonstrated that yeast strains might drop their productivity over time, jeopardizing the industrial fermentation processes.

5. Conclusions

It is clear that the robustness of the output of a microbial host over the generations is an important issue to consider in the development of a biotechnological production process. The difference in initial output between strains containing a sole copy of the reporter gene yECFP was in line with the current literature; however, no results were previously published about the temporal heterogeneity of the yeast strains engineered to produce a heterologous protein. At a multi-copy level, the outcome became unpredictable, as both synergistic and antagonistic effects were observed. While validation with the production of more complex proteins or pathways for the biosynthesis of metabolites, as well as in bioreactor conditions, is still warranted, the results presented here aimed to open the debate on strain robustness and the gap between strain engineering and their application in industrial processes. To the best of our knowledge, this study is the first of its kind, and no analogous studies have been published.

To standardize the observed lack of robustness, new and host-specific parameters were introduced. The 'last generation of stability' (G_{ST}) is determined as the generation after which the output digresses from the previous output level and, thus, becomes unstable. The 'rate of output change' (ROC) describes the manner in which a strain produces a desired product, while the 'overall robustness' (oR) is calculated as the difference between the initial

and final output divided by the number of generations between both. Finally, the 'potential product loss' (PPL) parameter indicates the difference between the maximal production capacity and the actual production output. Together with the process parameters titer, yield, and productivity, these strain-specific parameters aim to speed up strain development in a robust manner.

To overcome a drop in production over time, the time between strain construction and industrial application is often kept as short as possible. The results presented here postulate that the location of genomic integration of heterologous genes significantly influences the robustness of newly developed strains and should, thus, be carefully considered during the design of microbial cell factories. Not only does such a design warrant closer inspection of integration sites, it should also examine the combinatorial effect in cases where multiple copies of the heterologous gene(s) are desired. A streamlined strategy between the fermentation process set-up and strain development should be targeted to establish robust microbial cell factories, leading to improved bioprocesses for the synthesis of proteins as well as bulk and specialty chemicals by means of microbial fermentation.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/fermentation9080721/s1, Supplementary Figure S1: Annotated Genbank file of the integration cassette for the genomic modification of baker's yeast at the *ADH6* locus; Supplementary Figure S2: Correlation between the optical density and the colony forming units of the wild-type yeast; Supplementary Figure S3: Cumulative yECFP output profiles of the first sequential batch experiment; Supplementary Table S1: Homologous regions for genomic integration used in this study; Supplementary Table S2: Calculated relative expression strengths among strains at generation 5.

Author Contributions: Conceptualization, M.L.D.M., N.S., S.L.D.M., J.J.B. and W.K.S.; methodology, M.L.D.M., N.S. and S.L.D.M.; investigation, M.L.D.M., V.M. and I.M.; writing—original draft preparation, M.L.D.M. and N.S.; writing—review and editing, S.L.D.M.; visualization, M.L.D.M. and N.S.; supervision, S.L.D.M. and W.K.S.; funding acquisition, M.L.D.M., S.L.D.M. and W.K.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data used throughout this study are provided in the Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. *Fermentation Chemicals Market Size, Share & Trends Analysis Report by Application (Plastics & Fibers, Industrial), by Product (Organic Acids, Alcohols), by Region (Asia Pacific, North America), and Segment Forecasts, 2023—2030*; Grand View Research: San Francisco, CA, USA, 2023.
2. Ewing, T.A.; Nouse, N.; van Lint, M.; van Haveren, J.; Hugenholtz, J.; van Es, D.S. Fermentation for the Production of Biobased Chemicals in a Circular Economy: A Perspective for the Period 2022–2050. *Green Chem.* **2022**, *24*, 6373. [CrossRef]
3. Wang, J.; Lin, M.; Xu, M.; Yang, S.-T. Anaerobic Fermentation for Production of Carboxylic Acids as Bulk Chemicals from Renewable Biomass. In *Anaerobes in Biotechnology. Advances in Biochemical Engineering/Biotechnology*; Hatti-Kaul, R., Mamo, G., Mattiasson, B., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; Volume 156, pp. 323–361.
4. Rosenboom, J.-G.; Langer, R.; Traverso, G. Bioplastics for a circular economy. *Nat. Rev. Mater.* **2022**, *7*, 117–137. [CrossRef] [PubMed]
5. Kumar, S.; Krishnan, S.; Samal, S.K.; Mohanty, S.; Nayak, S.K. Itaconic acid used as a versatile building block for the synthesis of renewable resource-based resins and polyesters for future prospective: A review. *Polym. Int.* **2017**, *66*, 1349–1363. [CrossRef]

6. Liew, F.M.; Martin, M.E.; Tappel, R.C.; Heijstra, B.D.; Mihalcea, C.; Köpke, M. Gas Fermentation—A Flexible Platform for Commercial Scale Production of Low-Carbon-Fuels and Chemicals from Waste and Renewable Feedstocks. *Front. Microbiol.* **2016**, *7*, 694. [CrossRef] [PubMed]

7. Martínez, J.L.; Liu, L.; Petranovic, D.; Nielsen, J. Pharmaceutical protein production by yeast: Towards production of human blood proteins by microbial fermentation. *Curr. Opin. Biotechnol.* **2012**, *23*, 965–971. [CrossRef]

8. Meghwanshi, G.K.; Kaur, N.; Verma, S.; Dabi, N.K.; Vashishtha, A.; Charan, P.D.; Purohit, P.; Bhandari, H.; Bhojak, N.; Kumar, R. Enzymes for pharmaceutical and therapeutic applications. *Biotechnol. Appl. Biochem.* **2020**, *67*, 586–601. [CrossRef] [PubMed]

9. Matsushita, K.; Azuma, Y.; Kosaka, T.; Yakushi, T.; Hoshida, H.; Akada, R.; Yamada, M. Genomic analyses of thermotolerant microorganisms used for high-temperature fermentations. *Biosci. Biotechnol. Biochem.* **2016**, *80*, 655–668. [CrossRef]

10. Gustavsson, M.; Lee, S.Y. Prospects of microbial cell factories developed through systems metabolic engineering. *Microb. Biotechnol.* **2016**, *9*, 610–617. [CrossRef]

11. Meurens, N.A.A.; De Troostembergh, J.-C.M.P.; Horbach, B.; Kruse, M.J.C. Alcoholic Fermentation Process in the Presence of a High Alcohol Tolerant Yeast and a Maltotriose Positive Yeast. Global Patent WO2013181496A1, 31 May 2013.

12. Paddon, C.J.; Westfall, P.J.; Pitera, D.J.; Benjamin, K.; Fisher, K.; McPhee, D.J.; Leavell, M.D.; Tai, A.; Main, A.; Eng, D.; et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **2013**, *496*, 528–532. [CrossRef]

13. Dietrich, J.A.; Fortman, J.L.; Steen, E.J. Recombinant Host Cells for the Production of Malonate. U.S. Patent 9,816,114, 14 November 2017.

14. Galanie, S.; Thodey, K.; Trenchard, I.J.; Interrante, M.F.; Smolke, C.D. Complete Biosynthesis of Opiods in Yeast. *Science* **2015**, *349*, 1095–1100. [CrossRef]

15. Olsson, K.; Carlsen, S.; Semmler, A.; Simón, E.; Mikkelsen, M.D.; Møller, B.L. Microbial production of next-generation stevia sweeteners. *Microb. Cell Factories* **2016**, *15*, 207. [CrossRef]

16. Jiang, T.; Li, C.; Teng, Y.; Zhang, R.; Yan, Y. Recent advances in improving metabolic robustness of microbial cell factories. *Curr. Opin. Biotechnol.* **2020**, *66*, 69–77. [CrossRef] [PubMed]

17. Goel, A.; Wortel, M.T.; Molenaar, D.; Teusink, B. Metabolic shifts: A fitness perspective for microbial cell factories. *Biotechnol. Lett.* **2012**, *34*, 2147–2160. [CrossRef]

18. Gong, Z.; Nielsen, J.; Zhou, Y.J. Engineering Robustness of Microbial Cell Factories. *Biotechnol. J.* **2017**, *12*, 1700014. [CrossRef] [PubMed]

19. Wisselink, H.W.; Toirkens, M.J.; del Rosario, F.B.M.; Winkler, A.A.; van Dijken, J.P.; Pronk, J.T.; van Maris, A.J.A. Engineering of Saccharomyces cerevisiae for Efficient Anaerobic Alcoholic Fermentation of l-Arabinose. *Appl. Environ. Microbiol.* **2007**, *73*, 4881–4891. [CrossRef] [PubMed]

20. Verhoeven, M.D.; Lee, M.; Kamoen, L.; van den Broek, M.; Janssen, D.B.; Daran, J.-M.G.; van Maris, A.J.A.; Pronk, J.T. Mutations in PMR1 Stimulate Xylose Isomerase Activity and Anaerobic Growth on Xylose of Engineered Saccharomyces Cerevisiae by Influencing Manganese Homeostasis. *Sci. Rep.* **2017**, *7*, 46155. [CrossRef]

21. Karim, A.S.; Curran, K.A.; Alper, H.S. Characterization of Plasmid Burden and Copy Number in Saccharomyces Cerevisiae for Optimization of Metabolic Engineering Applications. *FEMS Yeast Res.* **2013**, *13*, 107–116. [CrossRef]

22. Ryan, O.W.; Skerker, J.M.; Maurer, M.J.; Li, X.; Tsai, J.C.; Poddar, S.; Lee, M.E.; DeLoache, W.; Dueber, J.E.; Arkin, A.P.; et al. Selection of Chromosomal DNA Libraries Using a Multiplex CRISPR System. *Elife* **2014**, *3*, e03703. [CrossRef]

23. Hahn-Hägerdal, B.; Karhumaa, K.; Fonseca, C.; Spencer-Martins, I.; Gorwa-Grauslund, M.F. Towards industrial pentose-fermenting yeast strains. *Appl. Microbiol. Biotechnol.* **2007**, *74*, 937–953. [CrossRef]

24. Flagfeldt, D.B.; Siewers, V.; Huang, L.; Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast* **2009**, *26*, 545–551. [CrossRef] [PubMed]

25. Jansen, M.L.A.; Bracher, J.M.; Papapetridis, I.; Verhoeven, M.; De Bruijn, H.; De Waal, P.P.; Van Maris, A.J.A.; Klaassen, P.; Pronk, J.T. Saccharomyces cerevisiae strains for second-generation ethanol production: From academic exploration to industrial implementation. *FEMS Yeast Res.* **2017**, *17*, fox044. [CrossRef]

26. Quan, J.; Tian, J. Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways. *PLoS ONE* **2009**, *4*, e6441. [CrossRef] [PubMed]

27. Gietz, R.D.; Schiestl, R.H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2007**, *2*, 31–34. [CrossRef]

28. Pyne, M.E.; Kevvai, K.; Grewal, P.S.; Narcross, L.; Choi, B.; Bourgeois, L.; Dueber, J.E.; Martin, V.J.J. A yeast platform for high-level synthesis of tetrahydroisoquinoline alkaloids. *Nat. Commun.* **2020**, *11*, 3337. [CrossRef] [PubMed]

29. Rowley, P.A. The Frenemies within: Viruses, Retrotransposons and Plasmids That Naturally Infect Saccharomyces Yeasts. *Yeast* **2017**, *34*, 279–292. [CrossRef] [PubMed]

30. Liu, J.-J.; Zhang, G.-C.; Kong, I.I.; Yun, E.J.; Zheng, J.-Q.; Kweon, D.-H.; Jin, Y.-S. A Mutation in PGM2 Causing Inefficient Galactose Metabolism in the Probiotic Yeast Saccharomyces boulardii. *Appl. Environ. Microbiol.* **2018**, *84*, e02858. [CrossRef]

31. Goormans, A.R.; Snoeck, N.; Decadt, H.; Vermeulen, K.; Peters, G.; Coussement, P.; Van Herpe, D.; Beauprez, J.J.; De Maeseneire, S.L.; Soetaert, W.K. Comprehensive Study on Escherichia Coli Genomic Expression: Does Position Really Matter? *Metab. Eng.* **2020**, *62*, 10–19. [CrossRef]

32. Zheng, D.-Q.; Zhang, K.; Wu, X.-C.; Mieczkowski, P.A.; Petes, T.D. Global analysis of genomic instability caused by DNA replication stress in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E8114–E8121. [CrossRef]

33. Nadal-Rey, G.; McClure, D.D.; Kavanagh, J.M.; Cornelissen, S.; Fletcher, D.F.; Gernaey, K.V. Understanding Gradients in Industrial Bioreactors. *Biotechnol. Adv.* **2021**, *46*, 107660. [CrossRef]

34. Bernardi, B.; Wendland, J. Homologous Recombination: A GRAS Yeast Genome Editing Tool. *Fermentation* **2020**, *6*, 57. [CrossRef]

35. Saka, K.; Takahashi, A.; Sasaki, M.; Kobayashi, T. More than 10% of yeast genes are related to genome stability and influence cellular senescence via rDNA maintenance. *Nucleic Acids Res.* **2016**, *44*, 4211–4221. [CrossRef]

36. Wu, G.; Yan, Q.; Jones, J.A.; Tang, Y.J.; Fong, S.S.; Koffas, M.A. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol.* **2016**, *34*, 652–664. [CrossRef]

37. Min, B.E.; Hwang, H.G.; Lim, H.G.; Jung, G.Y. Optimization of industrial microorganisms: Recent advances in synthetic dynamic regulators. *J. Ind. Microbiol. Biotechnol.* **2017**, *44*, 89–98. [CrossRef]

38. Kaishima, M.; Ishii, J.; Matsuno, T.; Fukuda, N.; Kondo, A. Expression of varied GFPs in Saccharomyces cerevisiae: Codon optimization yields stronger than expected expression and fluorescence intensity. *Sci. Rep.* **2016**, *6*, 35932. [CrossRef] [PubMed]

39. Kastberg, L.L.B.; Ard, R.; Jensen, M.K.; Workman, C.T. Burden Imposed by Heterologous Protein Production in Two Major Industrial Yeast Cell Factories: Identifying Sources and Mitigation Strategies. *Front. Fungal Biol.* **2022**, *3*, 827704. [CrossRef]

40. Fathi, Z.; Tramontin, L.R.R.; Ebrahimipour, G.; Borodina, I.; Darvishi, F. Metabolic engineering of Saccharomyces cerevisiae for production of β-carotene from hydrophobic substrates. *FEMS Yeast Res.* **2020**, *21*, foaa068. [CrossRef] [PubMed]

41. Ahmed, A.; Ahmad, A.; Li, R.; Al-Ansi, W.; Fatima, M.; Mushtaq, B.S.; Basharat, S.; Li, Y.; Bai, Z. Recent Advances in Synthetic, Industrial and Biological Applications of Violacein and Its Heterologous Production. *J. Microbiol. Biotechnol.* **2021**, *31*, 1465–1480. [CrossRef] [PubMed]

42. Vrábel, P.; van der Lans, R.G.; Luyben, K.C.; Boon, L.; Nienow, A.W. Mixing in large-scale vessels stirred with multiple radial or radial and axial up-pumping impellers: Modelling and measurements. *Chem. Eng. Sci.* **2000**, *55*, 5881–5896. [CrossRef]