Quantification of Microbial Robustness in Yeast

Cecilia Trivelin, Lisbeth Olsson, and Peter Rugbjerg*

ABSTRACT: Stable cell performance in a fluctuating environment is essential for sustainable bioproduction and synthetic cell functionality; however, microbial robustness is rarely quantified. Here, we describe a high-throughput strategy for quantifying robustness of multiple cellular functions and strains in a perturbation space. We evaluated quantification theory on experimental data and concluded that the mean-normalized Fano factor allowed accurate, reliable, and standardized quantification. Our methodology applied to perturbations related to lignocellulosic bioethanol production showed that the industrial bioethanol producing strain *Saccharomyces cerevisiae* Ethanol Red exhibited both higher and more robust growth rates than the laboratory strain CEN.PK and industrial strain PE-2, while a more robust product yield traded off for lower mean levels. The methodology validated that robustness is function-specific and characterized by positive and negative function-specific trade-offs. Systematic quantification of robustness to end-use perturbations will be important to analyze and construct robust strains with more predictable functions.

KEYWORDS: robustness quantification, phenomics, high-throughput, yeast, bioprocess, Fano factor

Microbial robustness ensures predictable, stable synthetic cellular functionality in spite of internal or external perturbations.1–3 Robust cell manufacturing and destination performance will be a key in realizing new synthetic biology modalities and efficient bioproduction.4–6 Robustness is defined for a specific function (or phenotype) and set of perturbations.7 Robustness is therefore different from tolerance, which specifically describes stable growth or survival to various perturbations, e.g., via specific growth rates.6,8 In silico systems biology quantifies robustness by the influence on a cellular function of a frequency-normalized perturbation space relative to a control condition9 with a normally distributed mean and standard deviation. Experimentally, measures of variation represent the stability (dispersion) of quantitative traits across perturbations, but not at different scales,7 for which the dimensionless coefficient of variation (CV) is better suited.10 Yet even if central to realizing predictable, scalable synthetic biology, robustness is seldomly quantified experimentally for strain functions, which may be subject to genetic or environmental perturbations of stochastic or determined behavior.5,11,12 Here, we present and validate a high-throughput methodology to experimentally quantify microbial robustness (script available on GitHub). We show that the methodology can be used to systematically quantify and compare the robustness of different strain functions of interest in a relevant perturbation space and relate them to their performance level. Precise quantification will allow for exploration of trade-offs between robustness and performance of different cellular functions.

RESULTS AND DISCUSSION

Development of a Systematic Method for Quantification of Robustness Based on the Fano Factor. In order to experimentally quantify robustness (R), we first set four important criteria to ensure consistency, reproducibility, and standardization. (1) Testing more perturbations should not change R, only its statistical significance. (2) Positive and negative deviations from the mean or a control performance level10 should contribute negatively to R. (3) Higher R should represent greater robustness. (4) R should be dimensionless and capture cellular function values at different orders of magnitude allowing comparison. To meet these criteria, we evaluated the reported theory.

The first theory quantifies R as the CV based on standard deviation/mean (σ/μ).10 RCV was calculated as

\[ R_{CV} = 1 - CV = 1 - \frac{\sigma}{\bar{x}} \]  

(1)

However, when cellular functions are subjected to different experimental perturbations, CV becomes >1 complicating interpretation. More importantly, as others have, we found that
CV was poorly accurate in describing data dispersion with means between 0 and 1 (Figure S1). CV therefore failed our fourth criterion.

The second theory quantifies R by evaluating a function change in relation to a specific control condition 0, according to an embodiment of Kitano’s formula:

\[
R_{\text{Kitano}} = \int_P \frac{f_{iS}(p)}{f_{iS}(0)} \psi(p) \, dp
\]

(2)

\[R_{\text{Kitano}}\] reports the ratio between the perturbed function \(f_{iS}(p)\) and the control condition \(f_{iS}(0)\) over a space of perturbations \(P\), each multiplied for its frequency \(\psi(p)\). We simplified eq 2 and assumed an equal frequency for each perturbation. However, functions performing better than the control achieved higher robustness (Figure S2); further, defining a control condition performance is not always meaningful. As a result, \(R_{\text{Kitano}}\) failed our criteria 1, 2, and 4.

Therefore, we evaluated an approach quantifying R as the dispersion of data around the function means using the Fano factor (Figure 1). The Fano factor is commonly used to study transcriptional bursting and noise in gene expression by identifying deviation from Poissonian behavior and has been proposed for robustness before, but not actually deployed. For each function \(i\), a strain \(S\), and a perturbation space \(P\), \(R_{iS,P}\) was calculated as \(\sigma^2/\bar{x}\):

\[
R_{iS,P} = -\frac{\text{Fano factor}}{\text{mean}} = -\frac{\sigma^2}{\bar{x}} \cdot \frac{1}{m}
\]

(3)

To allow comparison of \(R\) between functions, we normalized the different Fano factors to the mean of the functions they describe \((m)\) across all strains (Figure 1). We set the upper limit for \(R\) to 0 (highest robustness) and the problem of \(0 < \bar{x} < 1\) was solved. As the Fano factor remained finite for mean values approaching zero, the weight of the mean on \(R\) was higher than for the CV. This quantification strategy was frequency independent, dimensionless, and free from arbitrary control conditions, thereby meeting all criteria. The use of \(m\), however, still means that the \(R\) values calculated using the mean-normalized Fano factor by definition always be relative to the investigated data case.

**Quantifying Robustness of Five Cellular Functions: Case Study of Lignocellulosic Bioethanol Production.** To validate the methodology for quantifying microbial robustness, we used lignocellulosic bioethanol production (Materials and Methods). We included the Saccharomyces cerevisiae CEN.PK113–7D laboratory strain, and the industrial strains Ethanol Red and PE-2, whose robust ethanol production and growth is advantageous in starch and sugar cane fermentation.18–20 We measured five relevant cellular functions (maximum specific growth rate, lag phase, cell dry weight, biomass, and ethanol yields) across 29 different perturbations (single-component lignocellulose growth conditions) in a 96-well plate high-throughput setup (Materials and Methods).

The three strains exhibited different production and growth functionality when exposed to the 29 perturbations (Figure 2A). Across the lignocellulosic perturbation space, Ethanol Red performed better than CEN.PK and PE-2 in all functions, except for ethanol yield. Aldehydes had a negative effect on all five measured functions, while pentoses resulted in unchanged or improved functions. Lactic, levulinic, formic, and acetic acid reduced cell dry weight and biomass yield (Figure 2B).

We next quantified the robustness and found the maximum specific growth rate, ethanol yield, and cell dry weight as significantly higher in Ethanol Red (cell dry weight (\(p\)-value < 0.005), maximum specific growth rate (\(p\)-value < 7 \times 10^{-8}), ethanol yield (\(p\)-value < 0.001)) \((t\) test) (Figure 2B), supported by data less dispersed around the mean. These robust cellular functions were accompanied by a more fragile biomass yield, most robust in CEN.PK (\(p\)-value < 0.02), and lag phase most robust in PE-2 (\(p\)-value < 0.002) (Figure 2B). PE-2 achieved high mean ethanol yield, but its robustness was the lowest in part due to positive effects from pentoses. The observed very robust growth and production functions of Ethanol Red could explain its application in first-generation bioethanol plants that share our perturbations mainly except aldehydes, but came with a cost of lower average performance. High performance and high robustness are sometimes considered mutually exclusive.21 In theory, these two properties are believed to trade off with one another in several biological systems.22,23 Our quantification methodology identified possible robustness and performance trade-offs in lignocellulose-based bioethanol production (Figure 2C). For example, Ethanol Red traded its ethanol yield performance for high robustness, and vice versa for PE-2. Oppositely, we found that PE-2 traded robustness for performance for its maximum specific growth rate. On the basis of these findings, we observed that correlations between performance and robustness are both function and strain dependent. Larger data sets are needed to establish such correlations. In our investigation, the mean-normalized Fano factor was the best option to
calculate robustness, but we noticed that its suitability could be limited by the normalization with $m$ (the mean of the function performance among the strains). The normalization is required by criteria four; however, it further adjusts $R$ to the strains tested in the study.

A systematic framework for assessing robustness of several performance indicators will improve our understanding of how cellular functions respond to relevant perturbations, e.g., by favoring robustness, performance, or a suboptimal state for both. In strain engineering, robust yields of products are sometimes preferred over higher but unstable yields. In synthetic biology, engineered strains should carry new functions that perform robustly under anticipated perturbations. Relevant functions include biosensor signals, gene expression reporters and heterologous proteins production.

Quantifying robustness to scale-up or long-term cultivation (such as genetic robustness) will also be highly relevant to prevent declines of performance due to accumulation of mutations and heterogeneity, e.g., by quantifying robustness over many cellular divisions.

The high-throughput methodology described here will be useful for quickly quantifying robustness of multiple strain variants, environments and functions. For example, when engineering strains for heterologous expression, one could compare the robustness and performance trade-offs of product yield under various environmental and genetic conditions (e.g., different gene homologues, promoters, growth conditions). By doing so at laboratory scale, it may be possible to better understand and screen for robust cellular functions, rather than high-performing but unstable functions. When a strain is
selected with a robust cell function, shifting conditions (e.g.,
different concentrations of inhibitors in the growth media) would not affect functionality. Thus, different industrially
“robust strains” may be optimized toward robustness of production functionality across the narrow environmental and stochastic perturbation space of the bioprocess, making such strains robust to typical fluctuations in the process and thus not necessarily robust in foreign environments. Quantification methodology can uncover which perturbations affect the robustness of specific functions and how robustness relates positively and negatively to performance for each function (i.e., trade-offs). Through our methodology, we validated that robustness is function-specific, rather than a universal strain value, as previously theorized.

Future work will also show whether strains with robust functions are better starting points for subsequent enhancements via metabolic or evolutionary engineering. Robustness quantification may also include perturbation frequency, as well as more strains and perturbations. At present, our methodology may be applied to phenomics databases to compare numerous traits and strains.

MATERIAL AND METHODS

Strains. The strains used in the study were Saccharomyces cerevisiae CEN.PK113–7D25 (Scientific Research and Development GmbH, Oberursel, Germany), PE-226 (wild-type strain isolated during sugar cane-to-ethanol production in Brazil), and Ethanol Red (kindly provided by Société Industrielle Lesaffre, Division Leaf).

Media Preparation. Delft minimal medium27 was used for strain cultivation. The medium was prepared with 5 g/L (NH₄)₂SO₄, 3 g/L KH₂PO₄, 1 g/L MgSO₄·7H₂O, 1 mL (in 1 L solution) trace mineral solution (Table S1), and 1 mL (in 1 L solution) vitamin solution (Table S2). The medium was adjusted to pH 5 with KOH and buffered with 250 mM potassium hydrogen phthalate. Multiple compounds were added to the minimal medium to mimic the composition of lignocellulosic hydrolysates obtained mostly from spruce, corn starch, and wheat straw (Table S3).

Fermentation Experiments. The strains were preserved in glycerol (final concentration 16%) at −80 °C. Precultures were prepared by inoculating 10 µL glycerol stocks in 5 mL of the above-described Delft medium (20 g/L glucose), followed by incubation at 30 °C with 200 rpm shaking.

The optical density at 600 nm (OD₆₀₀) of overnight cultures was monitored in a GENESYS 10 spectrophotometer (Thermo Scientific). After 24 h, the cultures (in exponential phase) were inoculated in square polystyrene 96-half-deepwell microtiter plates (CR1496dg; Enzyscreen) at a starting OD₆₀₀ of 0.02. Strain growth was monitored in a Growth Pro plates (CR1496dg; Enzyscreen) at a starting OD₆₀₀ of 0.02.

The strains were cultivated in Delft medium (20 g/L glucose) to stationary phase and centrifuged at 5000 rpm for 5 min. The cell pellet was resuspended in 1 mL water and OD₆₀₀ was measured. A series of five 1:2 dilutions were made and filtered through a predrilled and weighed 0.45 µm poly(ether sulfone) membrane (Sartorius). OD₆₀₀ of the dilutions was measured. The filters containing the samples were dried for 10 min in a microwave oven (350 W) and weighed again. Calibration curves were constructed with CDW and OD₆₀₀. The slope values for each strain were used subsequently to calculate the CDW of the growth profiler cultures.

Determination of Sugars and Ethanol. Culture medium obtained by filtration from the first cultivation was used to determine the sugars and ethanol content. Ethanol was measured with the K-ETOH Ethanol Assay Kit, glucose with the K-GLUHK-220A d-Glucose HK Assay Kit, mannose with the K-MANGL d-Mannose/d-Fructose/d-Glucose Assay kit, xylose with the K-XYLOSE d-Xylose Assay Kit, and galactose and arabinose with the K-ARGA L-Arabinose/D-Galactose Assay Kit (all Megazyme). The assays are based on enzymatic reactions, which produce NADH, whose absorbance at 340 nm can be read in a spectrophotometer. The amount of NADH is stoichiometric with the amount of the compound of interest, making it possible to calculate the concentration (g/L) of the respective compounds.

Determination of Performance Values. The growth data in GV units were imported in R software for visualization and determination of growth parameters. The maximum specific growth rate (µmax) was determined using all_splines function. Duration of the lag phase was determined by calculating the x-coordinate of the intersection between the line with µmax slope passing through the inflection point and the line passing through y₀ = 0 parallel to the x-axis. In the wells where no growth was detected, R² was <0.99, so the µmax was set to 0 while the lag phase was set to NA.

Ethanol yield and biomass yield were calculated based on total consumed sugars as follows:

\[
y_e = \frac{\text{ethanol produced (g)}}{\text{initial sugars (g)} - \text{final sugars (g)}} \tag{4}
\]

\[
y_b = \frac{\text{CDW (g)}}{\text{initial sugars (g)} - \text{final sugars (g)}} \tag{5}
\]

Scripts with line-by-line explanation available on Github (https://github.com/cectri/Quantiﬁcation-of-microbial-robustness).
Robustness Calculation. Eq 3 was applied to the case study database mentioned above. The following variables were considered: $\mu_{\text{max}}$ (1/h), lag phase (h), CDW (g/L), biomass yield (g biomass/g consumed substrate), and ethanol yield (g produced/g consumed). Robustness and performance values were calculated and plotted in R. Statistical difference among the different cellular functions and strains was determined with an unpaired, two-sided t test and p-values were adjusted with Holm–Bonferroni method.

Scripts with line-by-line explanation available on Github (https://github.com/cectri/Quantification-of-microbial-robustness).

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.1c00615.

Robustness of ethanol yield, eq 1 (Figure S1); Robustness of biomass yield, eq 2 (Figure S2); Mineral solution composition (Table S1); Vitamin solution composition (Table S2); List of chemicals (and their relative concentrations) added to the medium to mimic the composition of lignocellulosic hydrolysates (Table S3) (PDF)

AUTHOR INFORMATION

Corresponding Author

Peter Rugbjerg — Department of Biology and Biological Engineering, Division of Industrial Biotechnology, Chalmers University of Technology, Gothenburg 412 96, Sweden; Enduro Genetics ApS, Copenhagen 2200, Denmark; orcid.org/0000-0003-2561-5063; Email: rugbjerg@chalmers.se

Authors

Cecilia Trivellin — Department of Biology and Biological Engineering, Division of Industrial Biotechnology, Chalmers University of Technology, Gothenburg 412 96, Sweden

Lisbeth Olsson — Department of Biology and Biological Engineering, Division of Industrial Biotechnology, Chalmers University of Technology, Gothenburg 412 96, Sweden

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.1c00615

Author Contributions

CT performed the experiments and the data analysis and drafted the manuscript. LO and PR supervised and contributed to the discussion of the collected data. All authors conceived the study and contributed to the correction of the manuscript prior to submission.

Notes

The authors declare no competing financial interest. The developed scripts for robustness quantification are available at Github: https://github.com/cectri/Quantification-of-microbial-robustness.

ACKNOWLEDGMENTS

Financial support by Novo Nordisk Foundation Grant Distinguished Investigator 2019—Research within biotechnology-based synthesis & production (#0055044) is gratefully acknowledged. Société Industrielle Lesaffre, Division Leaf, is kindly acknowledged for providing the Ethanol Red strain.

REFERENCES

(1) Kitano, H. Biological robustness. Nat. Rev. Genet. 2004, S, 826–837.

(2) Wehrs, M.; Tanjore, D.; Eng, T.; Lieveense, J.; Pray, T. R.; Mukhopadhyay, A. Engineering Robust Production Microbes for Large-Scale Cultivation. Trends Microbiol. 2019, 27, 524–537.

(3) Gong, Z.; Nielsen, J.; Zhou, Y. J. Engineering Robustness of Mycelial Cell Factories. Biotechnol. J. 2017, 12, 1700014.

(4) Li, Y.; Cheng, Z.; Zhao, C.; Gao, C.; Song, W.; Liu, L.; Chen, X. Reprogramming Escherichia coli Metabolism for Bioplastics Synthesis from Waste Cooking Oil. ACS Synth. Biol. 2021, 10, 1966–1979.

(5) Rugbjerg, P.; Olsson, L. The future of self-selecting and stable fermentations. J. Ind. Microbiol. Biotechnol. 2020, 47, 993–1004.

(6) Olsson, L.; Rugbjerg, P.; Torelllo Pianale, L.; Trivellin, C. Robustness: linking strain design to viable bioprocesses. Trends Biotechnol. 2022, DOI: 10.1016/j.tibtech.2022.01.004.

(7) Félix, M.-A.; Barkoulas, M. Pervasive robustness in biological systems. Nat. Rev. Genet. 2015, 16, 483–496.

(8) Song, X.; Li, Y.; Wu, Y.; Cai, M.; Liu, Q.; Gao, K.; Zhang, X.; Bai, Y.; Xu, H.; Qiao, M. Metabolic engineering strategies for improvement of ethanol production in cellulolytic Saccharomyces cerevisiae. FEMS Yeast Res. 2018, 18, 90.

(9) Kitano, H. Towards a theory of biological robustness. Mol. Syst. Biol. 2007, 3, 137.

(10) Yang, L.; Srinivasan, S.; Mahadevan, R.; Cluett, W. R. Characterizing metabolic pathway diversification in the context of perturbation size. Metab. Eng. 2015, 28, 114–122.

(11) Rugbjerg, P.; Sommer, M. O. A. Overcoming genetic heterogeneity in industrial fermentations. Nat. Biotechnol. 2019, 37, 869–876.

(12) Zakrzewska, A.; Van Eikenhorst, G.; Burggraaff, J. E. C.; Vis, D. J.; Hoefsloot, H.; Delneri, D.; Oliver, S. G.; Brul, S.; Smits, G. J. Genome-wide analysis of yeast stress survival and tolerance regulation to analyze the central trade-off between growth rate and cellular robustness. Mol. Biol. Cell 2011, 22, 4435–4446.

(13) Zheng, L.; Chen, M.; Nie, Q. External noise control in inherently stochastic biological systems. J. Math. Phys. 2012, 53, 115616.

(14) Munksy, B.; Neupert, G.; Van Oudenaarden, A. Using Gene Expression Noise to Understand Gene Regulation. Science 2012, 336, 183.

(15) Raj, A.; van Oudenaarden, A. Single-Molecule Approaches to Stochastic Gene Expression. Annu. Rev. Biophys. 2009, 38, 255.

(16) Grav, L. M.; Sergeeva, D.; Lee, J. S.; Marin De Mas, I.; Lewis, N. E.; Andersen, M. R.; Nielsen, L. K.; Lee, G. M.; Kildegaard, H. F. Minimizing Clonal Variation during Mammalian Cell Line Engineer-ing for Improved Systems Biology Data Generation. ACS Synth. Biol. 2018, 7, 2148–2159.

(17) Strucko, T.; Magdenoska, O.; Mortensen, U. H. Benchmarking two commonly used Saccharomyces cerevisiae strains for heterologous vanillin-β-glucoside production. Metab. Eng. Commun. 2015, 2, 99–108.

(18) Demeke, M. M.; Dietz, H.; Li, Y.; Foulquié-Moreno, M. R.; Mutturi, S.; Deprez, S.; Den Abt, T.; Bonini, B. M.; Liden, G.; Dumortier, F.; Verplaetse, A.; Boles, E.; Thevelein, J. M. Development of a D-xylose fermenting and inhibitor tolerant industrial Saccha-romyces cerevisiae strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. Biotechnol. Biofuels 2013, 6, 1–24.

(19) Wallace-Salinas, V.; Gorwa-Grauslund, M. F. Adaptive evolution of an industrial strain of Saccharomyces cerevisiae for combined tolerance to inhibitors and temperature. Biotechnol. Biofuels 2013, 6, 1–9.

(20) da Conceição, L. E. F. R.; Saraiva, M. A. F.; Diniz, R. H. S.; Oliveira, J.; Barbosa, G. D.; Alvarez, F. J.; de Mata Correa, L. F.; Mezadri, H.; Coutrim, M. X.; de Afonso, R. J. C. F.; Lucas, C.; Castro, I. M.; Brandão, R. L. Biotechnological potential of yeast isolates from cachaça: the Brazilian spirit. J. Ind. Microbiol. Biotechnol. 2015, 42, 237–246.
(21) Jin, Y.; Sendhoff, B. Trade-Off between Performance and Robustness: An Evolutionary Multiobjective Approach. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 2003, 2632, 237−251.

(22) Kitano, H. Violations of robustness trade-offs. Mol. Syst. Biol. 2010, 6, 384.

(23) Whitacre, J. M. Biological robustness: Paradigms, mechanisms, systems principles. Front. Genet. 2012, 3, 67.

(24) Binder, D.; Drepper, T.; Jaeger, K. E.; Delvigne, F.; Wiechert, W.; Kohlheyer, D.; Grünberger, A. Homogenizing bacterial cell factories: Analysis and engineering of phenotypic heterogeneity. Metab. Eng. 2017, 42, 145−156.

(25) Entian, K. D.; Kötter, P. 25 Yeast Genetic Strain and Plasmid Collections. Methods Microbiol. 2007, 36, 629−666.

(26) Basso, L. C.; De Amorim, H. V.; De Oliveira, A. J.; Lopes, M. L. Yeast selection for fuel ethanol production in Brazil. FEMS Yeast Res. 2008, 8, 1155−1163.

(27) Bruinenberg, P. M.; Van Dijken, J. P.; Scheffers, W. A. An enzymic analysis of NADPH production and consumption in Candida utilis. J. Gen. Microbiol. 1983, 129, 965−971.