Quantification of microbial robustness

Cecilia Trivellin¹, Lisbeth Olsson¹, Peter Rugbjerg^{1,2,*}

¹Department of Biology and Biological Engineering, Division of Industrial Biotechnology, Chalmers University of Technology, Kemivägen 10, 412 96, Gothenburg, Sweden

²Enduro Genetics ApS, Copenhagen, Denmark

*Correspondence: rugbjerg@chalmers.se

Graphical Abstract

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 quantifications 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 *Saccharomyces cerevisiae* Ethanol Red exhibited both higher and more robust growth rates than CEN.PK and 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

Introduction

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^{4,5}. Robustness is defined for a specific function (or phenotype) and set of perturbations⁶. Robustness is therefore different from tolerance, which specifically describes stable growth or survival to various perturbations e.g. via specific growth rates⁷. *In silico* systems biology quantifies robustness by the influence on a cellular function of a frequency-normalized perturbation space relative to a control condition⁸ 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 scales⁶, for which the dimensionless coefficient of variation (CV) is better suited⁹. 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^{5,10,11}. Here, we present and validate a high-throughput methodology to experimentally and systematically quantify microbial robustness with script available from 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 tradeoffs 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

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 level⁸ 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^{8,9}.

The first theory quantifies R as the CV based on standard deviation/mean ($\sigma/\bar{x}$)⁹. R_{CV} was calculated as:

$$R_{CV} = 1 - CV = 1 - \frac{\sigma}{\bar{x}} \text{ (Eq 1)}$$

However, when cellular functions are subjected to different experimental perturbations, CV becomes > 1 complicating interpretation. More importantly, as others have, we found CV was poorly accurate in describing data dispersion with means between 0 and 1 (Figure S1)¹². 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 Kitano's formula⁸:

$$R_{Kitano} = \int_p \frac{f_{i,S}(p)}{f_{i,S}(0)} \psi(p) dp \text{ (Eq 2)}$$

R_{Kitano} reports the ratio between the perturbed function $f_{i,S}(p)$ and the control condition $f_{i,S}(0)$ over a space of perturbations P , each multiplied for its frequency $\psi(p)$. We simplified Eq2 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_{Kitano} failed our criteria 1,2 and 4.

MEASUREMENT OF CELLULAR FUNCTIONS

ROBUSTNESS QUANTIFICATION

$R = -\text{mean normalized Fano factor}$

Figure 1: Relevant functions are measured upon exposure to various perturbations (colored dots) and robustness is calculated as the negative mean-normalized Fano factor. A control condition (e.g. 20 g/L glucose) is needed for calculation of R_{Kitano} .

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^{6,12–14}, but not actually deployed. For each function i , a strain S , and a perturbation space P , $R_{S,i,P}$ was calculated as $\sigma^2/\bar{x}$ ¹⁵:

$$R_{S,i,P} = - \frac{\text{Fano factor}}{\text{mean}} = - \frac{\sigma^2}{\bar{x}} \cdot \frac{1}{m} \text{ (Eq 3)}$$

To allow comparison between functions, we normalized the Fano factor to the mean m of 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 means 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.

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 (see Supporting Information, Material and methods). We included the *Saccharomyces cerevisiae* CEN.PK113-7D laboratory strain¹⁶, and the industrial strains Ethanol Red and PE-2, whose robust ethanol production and growth is advantageous in starch and sugarcane fermentation^{17–19}. 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 (Methods and materials).

Figure 2: Quantification of microbial robustness with the Fano factor. (A) Function evaluation in a large perturbation space containing components found in lignocellulosic hydrolysates. CDW: cell dry weight; $\mu_{\max}$: maximum specific growth rate. All points are individual biological replicates ($n = 3$). Lag phase missing points: cultures did not grow within 48 hr. (B) Robustness quantification for each function. Error bars: standard error of the mean ($n = 3$). (C) Robustness and performance trade-offs for each function and strain.

The three strains exhibited different production and growth functionality when exposed to the 29 perturbations (Figure 2A). Across the lignocellulose 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 2A).

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 < 7e-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 from aldehydes¹⁸, but came with a cost of lower average performance. High performance and high robustness are sometimes considered mutually exclusive²⁰. In theory, these two properties are believed to trade off with one another in several biological systems^{21,22}. 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. Based on these findings, we observed that correlations between performance and robustness are both function and strain depended. Larger datasets are needed to establish such correlations.

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 sub-optimal state for both. In strain engineering, robust yields of products are sometimes preferred over higher but unstable yields⁵. 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 (e.g. genetic robustness) will also be highly relevant to prevent declines of performance e.g. due to accumulation of mutations and heterogeneity^{5,23} e.g. by quantifying robustness over many cellular divisions.

High-throughput-friendly methodology as described here will be useful for quickly quantifying robustness of multiple strain variants and functions. The methodology can uncover which perturbations affect the robustness of specific functions, how robustness relates positively and negatively to performance for each function (i.e. trade-offs). Through our methodology we validated that robustness is function-specific⁶, rather than universal strain value, as previously theorized.

Future work will also show whether more robust starting strains are better 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.

Supporting Information

Figure S1: Robustness estimation with Eq1.

Figure S2: Robustness estimation with Eq2.

Material and methods.

Author information

Conflicts of interest

The authors declare no competing financial interest.

Notes

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. (2004) Biological robustness. *Nat. Rev. Genet.* 5, 826–837.
- (2) Wehrs, M., Tanjore, D., Eng, T., Lievense, J., Pray, T. R., and Mukhopadhyay, A. (2019) Engineering Robust Production Microbes for Large-Scale Cultivation. *Trends*

Microbiol. 27, 524–537.

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

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

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

(6) Félix, M.-A., and Barkoulas, M. (2015) Pervasive robustness in biological systems. *Nat. Rev. Genet.* 2015 168 16, 483–496.

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

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

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

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

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

(12) Zheng, L., Chen, M., and Nie, Q. (2012) External noise control in inherently

stochastic biological systems. *J. Math. Phys.* 53, 115616.

(13) Munsky, B., Neuert, G., and Van Oudenaarden, A. (2012) Using Gene Expression Noise to Understand Gene Regulation. *Science* 336, 183.

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

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

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

(17) 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., and Thevelein, J. M. (2013) Development of a D-xylose fermenting and inhibitor tolerant industrial *Saccharomyces cerevisiae* strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. *Biotechnol. Biofuels* 6, 1–24.

(18) Wallace-Salinas, V., and Gorwa-Grauslund, M. F. (2013) Adaptive evolution of an industrial strain of *Saccharomyces cerevisiae* for combined tolerance to inhibitors and temperature. *Biotechnol. Biofuels* 2013 61 6, 1–9.

(19) da Conceição, L. E. F. R., Saraiva, M. A. F., Diniz, R. H. S., Oliveira, J., Barbosa, G. D., Alvarez, F., da Mata Correa, L. F., Mezaadri, H., Coutrim, M. X., Afonso, R. J. de C. F., Lucas, C., Castro, I. M., and Brandão, R. L. (2015) Biotechnological potential of yeast isolates from cachaça: the Brazilian spirit. *J. Ind.*

Microbiol. Biotechnol. 42, 237–246.

(20) Jin, Y., and Sendhoff, B. (2003) Trade-Off between Performance and Robustness: An Evolutionary Multiobjective Approach. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 2632, 237–251.

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

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

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