226. Mi H, Lazareva-Ulitsky B, Loo R et al (2005) The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res 33:D284–D288
227. Croft D, O'Kelly G, Wu G et al (2011) Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res 39:D691–D697
228. Punta M, Coggill PC, Eberhardt RY et al (2012) The Pfam protein families database. Nucleic Acids Res 40:290–301
229. Hunter S, Jones P, Mitchell A et al (2012) InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res 40:D306–D312
230. cDonagh EM, Whirl-Carrillo M, Garten Y et al (2011) From pharmacogenomic knowledge acquisition to clinical applications: the PharmGKB as a clinical pharmacogenomic biomarker resource. Biomark Med 5:795–806
231. Zhang Y, James M, Middleton FA, Davis RL (2005) Transcriptional analysis of multiple brain regions in Parkinson's disease supports the involvement of specific protein processing, energy metabolism, and signaling pathways, and suggests novel disease mechanisms. Am J Med Genet 137B:5–16. doi:10.1002/ajmg.b.30195
232. Lesnick TG, Papapetropoulos S, Mash DC et al (2007) A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. PLoS Genet 3:12
233. Moran LB, Duke DC, Deprez M et al (2006) Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. Neurogenetics 7:1–11
234. Zheng B, Liao Z, Locascio JJ et al (2010) PGC-1α, a potential therapeutic target for early intervention in Parkinson's disease. Sci Transl Med 2:52ra–73ra
235. Scherzer CR, Eklund AC, Morse LJ et al (2007) Molecular markers of early Parkinson's disease based on gene expression in blood. PNAS 104:955–960
236. Sforza D, Hartenstein P, Lacan G, et al. (2008) Gene expression changes in multiple brain regions of a mouse MPTP model of Parkinson's disease. Fairfax, VA 22030 USA
237. Foti R, Zucchelli S, Biagioli M et al (2010) Parkinson disease-associated DJ-1 is required for the expression of the glial cell line-derived neurotrophic factor receptor RET in human neuroblastoma cells. J Biol Chem 285:18565–18574
238. Rajagopalan D, Agarwal P (2005) Inferring pathways from gene lists using a literature-derived network of biological relationships. Bioinformatics 21:788–793. doi:10.1093/bioinformatics/bti069
239. Ulitsky I, Shamir R (2009) Identifying functional modules using expression profiles and confidence-scored protein interactions. Bioinformatics 25:1158–1164. doi:10.1093/bioinformatics/btp118
240. Cerami E, Demir E, Schultz N et al (2010) Automated network analysis identifies core pathways in glioblastoma. PLoS One 5: e8918. doi:10.1371/journal.pone.0008918
241. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. PNAS 98:5116–5121. doi:10.1073/pnas.091062498
242. Aerts S, Lambrechts D, Maity S et al (2006) Gene prioritization through genomic data fusion. Nat Biotechnol 24:537–544. doi:10.1038/nbt1203
243. Ma X, Lee H, Wang L, Sun F (2007) CGI: a new approach for prioritizing genes by combining gene expression and protein–protein interaction data. Bioinformatics 23:215–221. doi:10.1093/bioinformatics/btl569
244. Zaykin DV, Zhivotovsky LA, Czika W et al (2007) Combining *p*-values in large-scale genomics experiments. Pharm Stat 6:217–226. doi:10.1002/pst.304
245. Marot G, Foulley J-L, Mayer C-D, Jaffrézic F (2009) Moderated effect size and *P*-value combinations for microarray meta-analyses. Bioinformatics 25:2692–2699
246. Aittokallio T, Schwikowski B (2006) Graph-based methods for analysing networks in cell biology. Brief Bioinform 7:243–255. doi:10.1093/bib/bbl022
247. Emmert-Streib F, Dehmer M (2011) Networks for systems biology: conceptual connection of data and function. IET Syst Biol 5:185. doi:10.1049/iet-syb.2010.0025
248. Junker BH, Schreiber F (2008) Analysis of biological networks. Science 1–28. doi:10.1002/9780470253489
249. Guimerà R, Sales-Pardo M, Amaral LAN (2007) Classes of complex networks defined by role-to-role connectivity profiles. Nat Phys 3:63–69. doi:10.1038/nphys489
250. Wang R-S, Albert R (2011) Elementary signaling modes predict the essentiality of signal transduction network components. BMC Syst Biol 5:44
251. Klipp E, Liebermeister W (2006) Mathematical modeling of intracellular signaling pathways. BMC Neurosci 7(Suppl 1):S10. doi:10.1186/1471-2202-7-S1-S10
252. Hucka M, Finney a, Sauro HM (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19:524–531. doi:10.1093/bioinformatics/btg015
253. Hoops S, Sahle S, Gauges R et al (2006) COPASI—a COmplex PAthway SImulator. Bioinform (Oxford, England) 22:3067–3074. doi:10.1093/bioinformatics/btl485
254. Van Eunen K, Kiewiet J a L, Westerhoff HV, Bakker BM (2012) Testing biochemistry revisited: how in vivo metabolism can be understood from in vitro enzyme kinetics. PLoS Comput Biol 8: e1002483. doi:10.1371/journal.pcbi.1002483
255. Kowald A, Hamann A, Zintel S et al (2012) A systems biological analysis links ROS metabolism to mitochondrial protein quality control. MechAgeing Dev 133:331–337. doi:10.1016/j.mad.2012.03.008
256. Berndt N, Bulik S, Holzhütter H-G (2012) Kinetic modeling of the mitochondrial energy metabolism of neuronal cells: the impact of reduced α-ketoglutarate dehydrogenase activities on ATP production and generation of reactive oxygen species. International journal of cell biology 2012:757594. doi: 10.1155/2012/757594
257. Craddock TJ a, Tuszynski J a, Chopra D et al (2012) The zinc dyshomeostasis hypothesis of Alzheimer's disease. PLoS One 7: e33552. doi:10.1371/journal.pone.0033552
258. Ambert N, Greget R, Haeberlé O et al (2010) Computational studies of NMDA receptors: differential effects of neuronal activity on efficacy of competitive and non-competitive antagonists. Open Access Bioinforma 2:113–125. doi:10.2147/OAB.S7246
259. Berardini TZ, Li D, Muller R, et al. (2012) Assessment of community-submitted ontology annotations from a novel database-journal partnership. Database : the journal of biological databases and curation 2012:bas030. doi: 10.1093/database/bas030
260. Meyer P, Alexopoulos LG, Bonk T et al (2011) Verification of systems biology research in the age of collaborative competition. Nat Biotechnol 29:811–815. doi:10.1038/nbt.1968
261. Meyer P, Hoeng J, Rice JJ et al (2012) Industrial methodology for process verification in research (IMPROVER): toward systems biology verification. Bioinform (Oxford, England) 28:1193–1201. doi:10.1093/bioinformatics/bts116
262. Nielsen MA (2011) Reinventing discovery: the new era of networked science. Princeton University Press, Princeton, NJ
263. Kelder T, Van Iersel MP, Hanspers K et al (2012) WikiPathways: building research communities on biological pathways. Nucleic Acids Res 40:D1301–D1307. doi:10.1093/nar/gkr1074
264. Matsuoka Y, Ghosh S, Kikuchi N, Kitano H (2010) Payao: a community platform for SBML pathway model curation. Bioinformatics 26:1381–1383. doi:10.1093/bioinformatics/btq143
265. Le Novère N, Hucka M, Mi H et al (2009) The systems biology graphical notation. Nat Biotechnol 27:735–741

ANALYTICAL SCIENCES MARCH 2013, VOL. 29

2013 © The Japan Society for Analytical Chemistry

Notes

Parallel Real-Time PCR on a Chip for Genetic Tug-of-War (gTOW) Method

Toyohiro NAITO,[*1,*2†] Ai YATSUHASHI,[*1] Noritada KAJI,[*1,*2] Taeko ANDO,[*3] Kazuo SATO,[*4] Hisao MORIYA,[*5] Hiroaki KITANO,[*6] Takao YASUI,[*1,*2] Manabu TOKESHI,[*2,*7] and Yoshinobu BABA[*1,*2,*8]

*1 *Department of Applied Chemistry, Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8603, Japan*
*2 *FIRST Research Center for Innovative Nanobiodevices, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8603, Japan*
*3 *College of Science and Engineering Department of Micro System Technology, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu, Shiga 525-8577, Japan*
*4 *Department of Mechanical Engineering, Faculty of Engineering, Aichi Institute of Technology, 1247 Yachigusa, Yakusa-cho, Toyota 470-0392, Japan*
*5 *Research Core for Interdisciplinary Sciences (RCIS), Okayama University, 3-1-1 Tsushima-Naka, Kita, Okayama 700-8530, Japan*
*6 *The Systems Biology Institute, Falcon Building 5F, 5-6-9 Shirokanedai, Minato, Tokyo 108-0071, Japan*
*7 *Division of Biotechnology and Macromolecular Chemistry, Faculty of Engineering, Hokkaido University, Kita 13 Nishi 8, Kita, Sapporo 060-8628, Japan*
*8 *Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 2217-14 Hayashi-cho, Takamatsu 761-0395, Japan*

A microchip-based real-time polymerase chain reaction (PCR) device has been developed for the genetic tug-of-war (gTOW) method that provides quantitative data for research on biorobustness and systems biology. The device was constructed of a silicon glass chip, a temperature controlling Peltier element, and a microscope. A parallel real-time amplification process of target genes on the plasmids and the housekeeping genes in a model eukaryote *Saccharomyces cerevisiae* were detected simultaneously, and the copy number of the target genes were estimated. The device provides unique quantitative data that can be used to augment understanding of the system-level properties of living cells.

(Received December 24, 2012; Accepted January 13, 2013; Published March 10, 2013)

Introduction

Intracellular biochemical parameters, such as the gene expression level of gene products, are considered to have been optimized during the evolutionary process to allow biological activities to be carried out in an organism's habitat. These parameters should have some permissible range so that the systems have robustness against perturbations. However, little is known about the permissible ranges of parameters in real cells because there has been no experimental technique to comprehensively measure the limits of intracellular parameters. Gene knockout experiments are used to reduce the expression level of target genes.[1,2] These experiments provide phenotypical information that reveals the functions of target genes. However, such experiments do not provide quantitative information associated with the limit of expression of the target genes in order to maintain function of the biological systems. Promoter-swapping experiments, in which the promoter of the target gene is changed into a strong promoter, are used to increase the expression level of the target genes, and this method also has provided much useful information for predicting the functions of target genes, as well as genetic interactions between target genes.[3-5] However, it is also difficult to determine the upper limit of the expression of the target genes because this method ignores the native expression level and regulation of the target genes.[6-9] Genetic screening by a genetic tug-of-war (gTOW) method allows the upper limit of each target gene in living cells to be measured by increasing the copy number of that gene.[10,11] Each target gene and its native regulatory DNA elements (promoter and terminator) are used as a unit so that the increased copy number of that gene can be determined quantitatively, and the gene expression level is expected to increase according to the copy number. Knowledge obtained from the gTOW method is important for understanding of biology at the system level.[12-14]

The gTOW method analysis is useful exploring the foundation of biological systems such as biological robustness.[15-17] Biological robustness is proposed as the main concept for the application of systems biology in cancer research. Biological robustness is a fundamental feature of evolvable complex systems that must be robust against environmental and genetic perturbation to be evolvable. At the same time, robust systems face fragility and performance setback as an inherent trade-off.

† To whom correspondence should be addressed.
E-mail: naito.toyohiro@f.mbox.nagoya-u.ac.jp

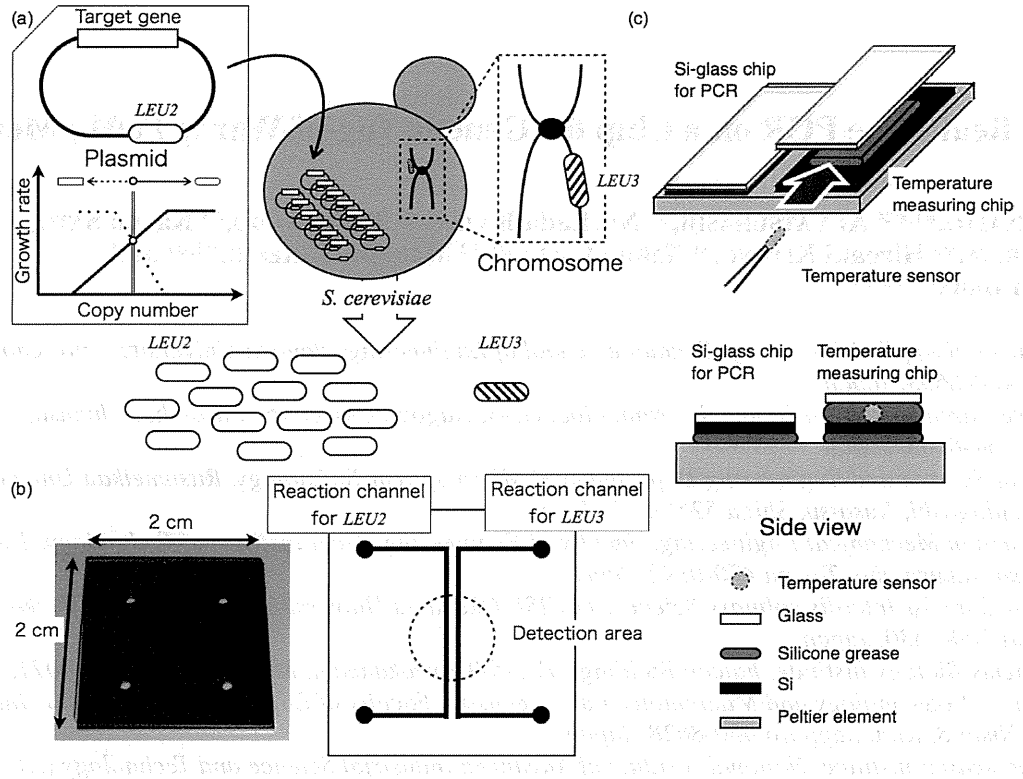

Fig. 1 Schematic illustrations of the PCR device for gTOW method. (a) Samples used in the gTOW method. Two kinds of genes were used, one is *LEU2* gene coded on plasmid introduced to *S. cerevisiae* and the other gene is *LEU3* coded on a chromosomes of *S. cerevisiae*. After a tug-of-war between the increasing activity of the copy number from *LEU2* and the decreasing activity of the copy number from *LEU3*, we could define the copy number of the plasmid. (b) Design of the PCR chip. Two channels were aligned in a single microscopic field to allow direct comparison of the relative quantities of *LEU2* and *LEU3*. (c) Schematic illustration of the temperature measurement system. To measure the temperature at the microfluidic channel, a temperature sensor was sandwiched between the silicon wafer and glass cover next to the chip for PCR.

Identification of the basic architecture for a robust system and the associated trade-off is essential for understanding their faults and countermeasures against fatal diseases. Although the gTOW method is the only way to provide quantitative data that can be used to determine the system-level properties of living cells and to quantitate biorobustness against gene overexpression, it needs high-throughput, comprehensive, and accurate measurement devices to obtain precisely quantified information.

Microchip devices are one of the most suitable technologies to be introduced to design and build next-generation experimental devices. It may drastically improve the speed and accuracy of measurement through integration and automation of experimental procedures. Especially, real-time polymerase chain reaction (PCR) is an important procedure to understand biological systems through the gTOW method. Although several real-time PCR devices have been reported,[18,19] they were not optimized for the gTOW method. In this paper, we describe development of a microchip-based device that realizes parallel real-time PCR on a chip for the gTOW method.

Experimental

The gTOW method can estimate the upper limit of the gene expression level for each target gene by increasing the copy number of that gene. Two kinds of genes were amplified in the gTOW method to determine the upper limit copy number of target genes; one was leucine biosynthesis gene *LEU2* coded on the plasmid DNA with a target gene and the other was transcriptional regulator of leucine *LEU3* coded on the chromosome of *Saccharomyces cerevisiae* (Fig. 1(a)). By comparing the relative quantity of *LEU2* and *LEU3* genes, we could estimate the copy number of the plasmid per haploid genome. A design of our device is shown in Fig. 1(b). The temperature was controlled by the Peltier element and reaction channels were fabricated on a silicon wafer. The thermal control system was comprised of a power supply (HLE, Nihontecmo Co., Ltd., Fukuoka, Japan), a controller (SP5R7-576, Nihontecmo Co., Ltd.), the Peltier element (TEC-12708, Nihontecmo Co., Ltd.) and LabVIEW-based temperature control and data-logging software. The reaction mixture contained intercalating dye, SYBER Green1, which allowed DNA detection with a fluorescence microscope. In this device design, the amplification process of *LEU2* and *LEU3* genes could be detected simultaneously in a single observation view.

Total DNA used as a template DNA sample was extracted from wild type yeast cells (BY4741) with pTOW that is a 2-micron based plasmid with *leu2d* for the gTOW method.[10] Yeast cells, collected from 200 μL of saturated culture, were suspended in lysis solution (10 mM Na-phosphate (pH 7.5), 1.2 M sorbitol, and 2.5 mg/mL Zymolyase 100T) (Seikagaku, Tokyo, Japan) and incubated with a block incubator (BI-525, ASTEC, Fukuoka, Japan) for 10 min at 37°C to digest the cell wall. Then the cell suspension was heat-shocked at 94°C for

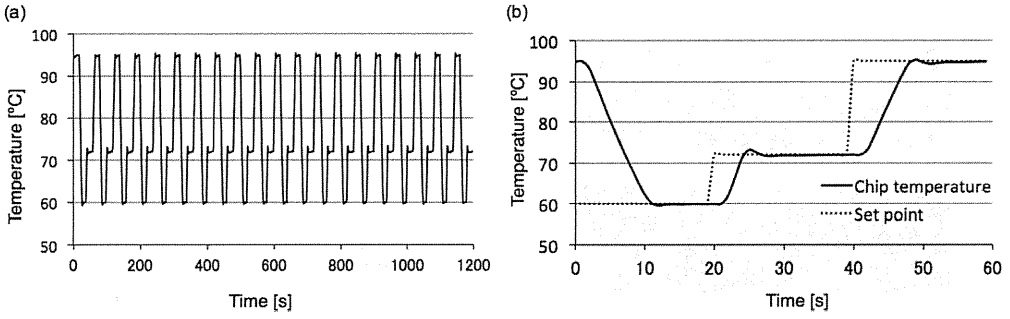

Fig. 2 (a) Temperature profile for 20 cycles: one cycle consisted of heating at 95°C for 20 s, 60°C for 20 s, and 72°C for 20 s. (b) Comparison between chip temperature and set point for 1 cycle.

15 min, –80°C for 5 min, and then 94°C for 15 min. After the cell suspension was chilled and centrifuged, supernatant (containing total DNA) was used for the following two real-time PCRs: LightCycler FastStart DNA MasterPLUS SYBER Green I (Roche Diagnostics K.K., Tokyo, Japan) with LightCycler 2.0 instrument (Roche Diagnostics K.K.) was used for the kinetic PCR. Supernatant (2 μL) was mixed with each reaction mix (18 μL) containing 0.5 μM of *LEU2* primer set (F: GCTAATGTTTTGGCCTCTTC, R: ATTTAGGTGGGTTGGG-TTCT) and *LEU3* primer set (F: CAGCAACTAAGGACAAGG, R: GGTCGTTAATGAGCTTCC) using 0.1% bovine serum albumin (Sigma Aldrich Japan K.K., Tokyo, Japan). The genes were amplified using the following thermal cycling profile; initial denaturation at 95°C for 10 min, followed by 30 cycles with one cycle consisting of 95°C for 10 s, 60°C for 10 s and 72°C for 7 s.

Reaction channels were fabricated on silicon wafers by using standard photolithography and deep reactive ion etching (MUC-21, Sumitomo Precision Products Co., Ltd., Hyogo, Japan). The reaction channel had a rectangular cross section of 100 μm width and 20 μm depth. After dicing the wafers with an automatic dicing saw (DAD522, DISCO, Tokyo, Japan), inlet and outlet holes were made using ultrasonic drilling (SOM-121, Shinoda K.K., Tokyo, Japan). A glass cover (SD-2, Hoya Candeo, Saitama, Japan) was anodically bonded to the silicon substrate at 400°C, 1 kV with the handmade anodic bonding machine consists of a hot plate (MSA Factory, Tokyo, Japan) and a power supply (31601-5N, Apple Electronics Co., Ltd., Shanghai, China). A silicon-glass chip for PCR and silicon-glass plate sandwiching a temperature sensor for temperature measurements were put into silicone grease (TCOH-1002, Taica, Tokyo, Japan) on the Peltier element (Fig. 1(c)). As for real-time PCR on a chip, the same reaction mixture was used with different thermal cycling; initial denaturation at 95°C for 2 min, followed by 30 cycles (95°C for 27 s, 60°C for 28 s and 72°C for 15 s). Fluorescent intensity was measured by a fluorescent microscope (Eclipse T*i*, Nikon, Tokyo, Japan) and open source image analysis software (ImageJ).

Results and Discussion

S. cerevisiae lacking the leucine-generating function was used for the gTOW method. The leucine-generating function of the cells was obtained by the plasmids. Cells with a higher copy number of the plasmids can grow faster under the leucine condition because the activity of the leucine-generating function derived from the plasmid is very weak. On the other hand, the copy number of the plasmid is limited by the target gene on the plasmid. The target gene becomes another genetic selection bias toward decreasing the plasmid copy number for inhibition of growth over the upper limit copy number of the target gene. The copy numbers of the plasmids reach an equilibrium between the increasing activity of the copy number from *LEU2* and the decreasing activity of the copy number from *LEU3*, just like a tug-of-war between the two kinds of genes.

First, we evaluated the temperature control device. The Peltier element was computer-controlled with a proportional integral derivative (PID) control algorithm, and we optimized the PID setting to achieve a high temperature ramp rate, high accuracy, and small temperature overshoot. The temperature program was set to 20 cycles (heating at 95°C for 20 s, 60°C for 20 s, 72°C 20 s). As shown in Fig. 2(a), the temperature control device worked reliably throughout 20 thermal cycles and it also functioned for over 30 cycles in subsequent experiments. Total process time could be extended to 3600 s by the temperature control program, which was an adequate time scale for the PCR process. The actual holding time in the reaction channel was 95°C for 13 s, 60°C for 12 s, and 72°C for 17 s, all with the accuracy of ±0.3°C; the rate for changing the temperature was 4°C s^{-1} and it took 2 s to change the heating-cooling mode (Fig. 2(b)). The temperature control system achieved the same level of the temperature ramp rate and accuracy as the commercial product we used, and the temperature overshoot was negligible.

We tried to amplify two kinds of genes simultaneously in a single observation view, and we observed increasing fluorescent intensity of the reaction channel for *LEU2* compared to that of the reaction channel for *LEU3* (Fig. 3(a)). To estimate the relative fluorescent intensity result from amplification of *LEU2* gene to amplification of *LEU3* gene, we measured fluorescent intensity profiles of each image. The relative fluorescent intensity of a *LEU2* channel *versus* a *LEU3* channel was 2.2 times larger after the 30th cycle, whereas there was no change between fluorescent intensities of the channel for *LEU2* and *LEU3* after 10 cycles (Fig. 3(b)). Amplification curves of *LEU2* and *LEU3* genes were obtained using chips (Fig. 3(c)). After fitting with a logistic curve, the relative quantities of *LEU2* and *LEU3* were compared at 40% of their maximum value. The copy number of the plasmid determined by the amplification curves obtained by fabricated devices was 365 copies; this value was quite different from the 107 copies obtained in a bulk experiment (Fig. 3(d)). The difference is attributed to the amplification curves of the on-chip PCR experiment being of low precision due to observation with unfixed focus. Three possible causes for unfixed focus are: non-uniform thickness of

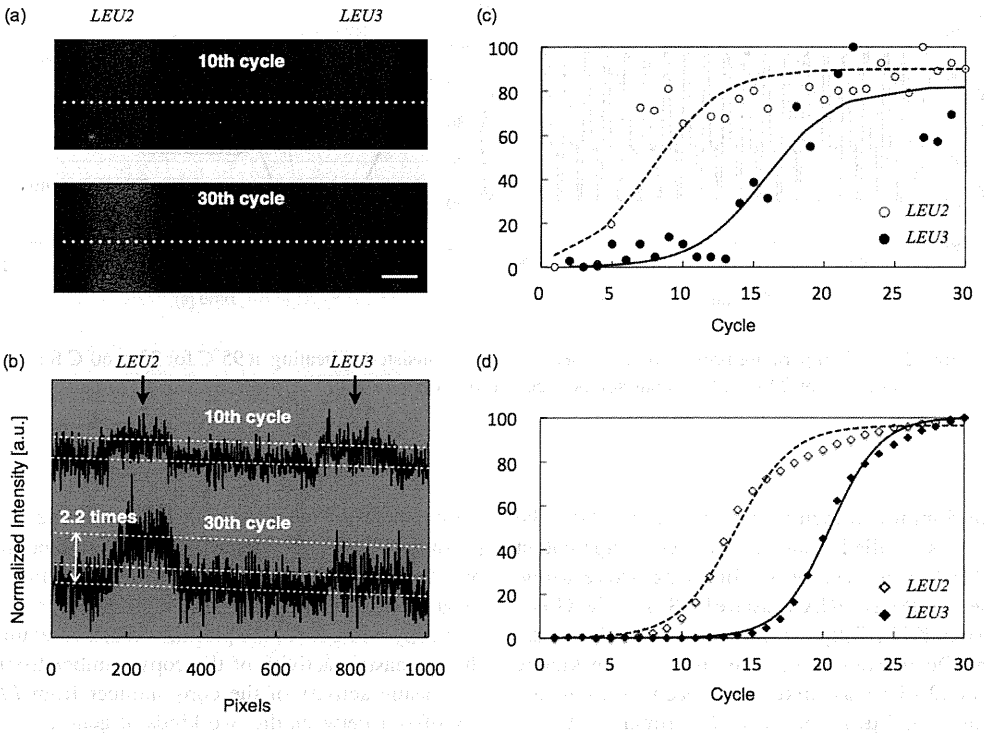

Fig. 3 (a) Fluorescent images of microfluidic channel for directly comparing PCR at the 10th cycle (upper) and 30th cycle (lower). (b) Normalized fluorescent intensity of images from (c). (c) Amplification curves for *LEU2* and *LEU3* genes by the chip. Open and solid circles indicate normalized measured values of *LEU2* and *LEU3*, respectively. The logistic curve was fitted to the measurement values (dotted line, *LEU2*; solid line, *LEU3*). (d) Amplification curves for *LEU2* and *LEU3* genes obtained by a commercial PCR instrument. Open and solid squares indicate the normalized measured values of *LEU2* and *LEU3*, respectively. The logistic curve was fitted to the measurement values (dotted line, *LEU2*; solid line, i*LEU3*).

the silicone grease, an unstable heating system, and thermal expansion of the silicon-glass chip. The non-uniform grease thickness might cause inhomogeneous heating of reaction channels and that would cause slight changes in reaction time and reaction temperature. Although the silicone grease we used was a conductive thermal silicone sheet pad to prepare homogeneous silicone layers, the thickness of grease was changed by manually putting chips on the Peltier element. To obtain higher precision, it is necessary to use special holders with a detector to fix in place the PCR chip with the temperature control system including the Peltier element that is put on the silicone grease under uniform pressure. The problem of thermal expansion of silicon-glass chips was also caused by detection with a manual procedure. To solve the problem we will have to develop a fluorescent capture system that synchronizes with the temperature control software to detect the fluorescent signal of reaction channels periodically. In addition, the PCR device needs significant improvements, that is, microfluidic techniques to integrate experimental procedures and functions into the PCR device for automation and exclusion of caused by manual handling experimental errors. We have developed two key components, a solution mixing device[20,21] and a reaction channel sealing system,[22] for accelerating automation of the procedures to obtain more quantitative data. We will integrate these key components with the PCR device developed here in the near future. There is a possibility that the logistic curve is not suitable for curve fitting in this case. We attempted to make a fitting curve using a logistic curve because DNAs were double a cycle during PCR, and we used the same data analysis method for on-chip and in-bulk PCRs. However, the fitting curve in Fig. 3(d) does not seem to be the best fitting curve. Concentrations of PCR solutions change with increasing number of thermal cycles, which changes the amplification efficiency. We considered that using the best suited curve fitting might reduce the analytical error between on-chip and in-bulk PCRs.

There are several issues that need to be addressed regarding our parallel PCR device. One is that the device used only 2 μL PCR solution, only 10 percent of the sample volume required for a bulk experiment for amplifications of DNA. The device can be extended to an even more parallelized PCR device that can also prepare calibration curves during the amplifications of the genes to determine the upper limit copy number of target genes after fewer thermal cycles with the same or less sample volume for bulk experiments.

Conclusions

We developed a real-time PCR chip for gTOW method and amplified two kinds of genes using the chip. The device could be improved, however, such as by making a special chip holder, changing the channel design and modifying channel surfaces to prevent non-specific adsorption of an enzyme so that the device can be easily used by systems biologists. The device has good potential for promoting studies in systems biology.

Acknowledgements

This research was partially supported by the Japan Society for the Promotion of Science (JSPS) through its "Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)" and Grant-in-Aid for JSPS Fellows.

References

1. G. Giaever, A. M. Chu, L. Ni, C. Connelly, L. Riles, S. Véronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. André, A. P. Arkin, A. Astromoff, M. E. Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K.-D. Entian, P. Flaherty, F. Foury, D. J. Garfinkel, M. Gerstein, D. Gotte, U. Güldener, J. H. Hegemann, S. Hempel, Z. Herman, D. F. Jaramillo, D. E. Kelly, S. L. Kelly, P. Kötter, D. LaBonte, D. C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S. Loon Ooi, J. L. Revuelta, C. J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D. D. Shoemaker, S. Sookhai-Mahadeo, R. K. Storms, J. N. Strathern, G. Valle, M. Voet, G. Volckaert, C.-Y. Wang, T. R. Ward, J. Wilhelmy, E. A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J. D. Boeke, M. Snyder, P. Philippsen, R. W. Davis, and M. Johnston, *Nature*, **2002**, *418*, 387.
2. S. Mnaimneh, A. P. Davierwala, J. Haynes, J. Moffat, W. T. Peng, W. Zhang, X. Yang, J. Pootoolal, G. Chua, A. Lopez, M. Trochesset, D. Morse, N. J. Krogan, S. L. Hiley, Z. Li, Q. Morris, J. Grigull, N. Mitsakakis, C. J. Roberts, J. F. Greenblatt, C. Boone, C. A. Kaiser, B. J. Andrews, and T. R. Hughes, *Cell*, **2004**, *118*, 31.
3. X. Pan, P. Ye, D. S. Yuan, X. Wang, J. S. Bader, and J. D. Boeke, *Cell*, **2006**, *124*, 1069.
4. A. H. Tong and C. Boone, *Methods Mol. Biol.*, **2006**, *313*, 171.
5. A. H. Tong, G. Lesage, G. D. Bader, H. Ding, H. Xu, X. Xin, J. Young, G. F. Berriz, R. L. Brost, M. Chang, Y. Chen, X. Cheng, G. Chua, H. Friesen, D. S. Goldberg, J. Haynes, C. Humphries, G. He, S. Hussein, L. Ke, N. Krogan, Z. Li, J. N. Levinson, H. Lu, P. Ménard, C. Munyana, A. B. Parsons, O. Ryan, R. Tonikian, T. Roberts, A. M. Sdicu, J. Shapiro, B. Sheikh, B. Suter, S. L. Wong, L. V. Zhang, H. Zhu, C. G. Burd, S. Munro, C. Sander, J. Rine, J. Greenblatt, M. Peter, A. Bretscher, G. Bell, F. P. Roth, G. W. Brown, B. Andrews, H. Bussey, and C. Boone, *Science*, **2004**, *303*, 808.
6. H. Liu, J. Krizek, and A. Bretscher, *Genetics*, **1992**, *132*, 665.
7. R. Akada, J. Yamamoto, and I. Yamashita, *Mol. Gen. Genet.*, **1997**, *254*, 267.
8. L. F. Stevenson, B. K. Kennedy, and E. Harlow, *Proc. Natl. Acad. Sci. U. S. A.*, **2001**, *98*, 3946.
9. R. Sopko, D. Huang, N. Preston, G. Chua, B. Papp, K. Kafadar, M. Snyder, S. G. Oliver, M. Cyert, T. R. Hughes, C. Boone, and B. Andrews, *Mol. Cell*, **2006**, *21*, 319.
10. H. Moriya, Y. Shimizu, Y. Yoshida, and H. Kitano, *PLoS Genet.*, **2006**, *2*, e111.
11. K. Kaizu, H. Moriya, and H. Kitano, *PLoS Genet.*, **2010**, *6*, e1000919.
12. J. K. Nicholson, *Mol. Sys. Biol.*, **2006**, *2*, 52.
13. Q. Cui, Y. Ma, M. Jaramillo, H. Bari, A. Awan, S. Yang, S. Zhan, L. Liu, M. O'Connor-MacCourt, E. O. Purisima, and E. Wang, *Mol. Sys. Biol.*, **2007**, *3*, 152.
14. J. J. Hornberg, F. J. Bruggeman, H. V. Westerhoff, and J. Lankelma, *BioSystems*, **2006**, *83*, 81.
15. H. Kitano, *Nature*, **2002**, *420*, 6912.
16. H. Kitano, *Science*, **2002**, *295*, 5560.
17. M. Ahdesmäki, H. Lähdesmäki, R. Pearson, H. Huttunen, and O. Yli-Harja, *BMC Bioinformatics*, **2005**, *6*, 117.
18. C. Zhang, J. Xu, W. Ma, and W. Zheng, *Biotechnol. Adv.*, **2006**, *24*, 243.
19. C. Zhang and D. Xing, *Nucleic Acids Res.*, **2007**, *35*, 4223.
20. T. Yasui, Y. Omoto, K. Osato, N. Kaji, N. Suzuki, T. Naito, M. Watanabe, Y. Okamoto, M. Tokeshi, E. Shamoto, and Y. Baba, *Lab Chip*, **2011**, *11*, 3356.
21. T. Yasui, Y. Omoto, K. Osato, N. Kaji, N. Suzuki, T. Naito, Y. Okamoto, M. Tokeshi, E. Shamoto, and Y. Baba, *Anal. Sci.*, **2012**, *28*, 57.
22. T. Naito, R. Arayanarakool, S. L. Gac, T. Yasui, N. Kaji, M. Tokeshi, A. van den Berg, and Y. Baba, *Lab Chip*, **2013**, *13*, 452.

Identification of dosage-sensitive genes in *Saccharomyces cerevisiae* using the genetic tug-of-war method

Koji Makanae,[1] Reiko Kintaka,[2] Takashi Makino,[3] Hiroaki Kitano,[4,5,6] and Hisao Moriya[1,7]

[1]*Research Core for Interdisciplinary Sciences, Okayama University, Okayama 700-8530, Japan;* [2]*Graduate School of Natural Science and Technology, Okayama University, Okayama 700-8530, Japan;* [3]*Graduate School of Life Sciences, Tohoku University, Sendai 980-8578, Japan;* [4]*Sony Computer Science Laboratories, Tokyo 141-0022, Japan;* [5]*The Systems Biology Institute, Tokyo 108-0071, Japan;* [6]*Okinawa Institute of Science and Technology Graduate University, Okinawa 904-0495, Japan*

Gene overexpression beyond a permissible limit causes defects in cellular functions. However, the permissible limits of most genes are unclear. Previously, we developed a genetic method designated genetic tug-of-war (gTOW) to measure the copy number limit of overexpression of a target gene. In the current study, we applied gTOW to the analysis of all protein-coding genes in the budding yeast *Saccharomyces cerevisiae*. We showed that the yeast cellular system was robust against an increase in the copy number by up to 100 copies in >80% of the genes. After frameshift and segmentation analyses, we isolated 115 dosage-sensitive genes (DSGs) with copy number limits of 10 or less. DSGs contained a significant number of genes involved in cytoskeletal organization and intracellular transport. DSGs tended to be highly expressed and to encode protein complex members. We demonstrated that the protein burden caused the dosage sensitivity of highly expressed genes using a gTOW experiment in which the open reading frame was replaced with GFP. Dosage sensitivities of some DSGs were rescued by the simultaneous increase in the copy numbers of partner genes, indicating that stoichiometric imbalances among complexes cause dosage sensitivity. The results obtained in this study will provide basic knowledge about the physiology of chromosomal abnormalities and the evolution of chromosomal composition.

[Supplemental material is available for this article.]

Intracellular biochemical parameters, such as gene expression levels and protein activities, are highly optimized to maximize the performance of biological systems (Zaslaver et al. 2004; Dekel and Alon 2005; Wagner 2005). These parameters, however, have certain permissive ranges to protect the function of the system against perturbations such as environmental changes, mutations, and noise in biochemical reactions. This robustness against fluctuations in parameters is considered a common design principle of biological systems (Alon et al. 1999; Little et al. 1999; von Dassow et al. 2000). When gene expression fluctuates beyond the robustness of cellular systems, various defects occur in the systems. However, the differences in the expression limits of different genes and the factors influencing these differences are unclear.

We previously developed the genetic tug-of-war (gTOW) method to measure the limit of gene overexpression (Moriya et al. 2006, 2011, 2012). Using gTOW, we can assess the limit of gene overexpression as the copy number limit (CNL) of the target gene as follows. A target gene with its native regulatory sequences is cloned into a plasmid for gTOW. The plasmid carries a 2-micron origin, *URA3*, and *LEU2* with a truncated promoter (*leu2d*). Yeast cells are transformed by the plasmid, and the transformants are first selected in medium lacking uracil (−Ura). The cells are then transferred into medium lacking both uracil and leucine (−Leu−Ura). In this medium, *leu2d* becomes a selection bias to increase the plasmid copy number in the cells because the cells with higher *leu2d* (plasmid) copy numbers grow faster. As the copy number increases, the copy number of the target gene also increases, and the gene becomes proportionally overexpressed according to the increased copy number. If the gene has an overexpression limit at which cellular function is halted when the limit is crossed (i.e., inducing cellular death), then the plasmid copy number must be less than the limit, and the target gene becomes a selection bias to decrease the plasmid copy number. Biases arising from *leu2d* (that increases the plasmid copy number) and the target gene (that decreases the plasmid copy number) determine the plasmid copy number in the cells (thus, we designated this method "genetic tug-of-war"). Because the bias to increase the plasmid copy number by *leu2d* is always the same, the copy number should be associated with the CNL of overexpression of the target gene. The plasmid copy number determined under the −Leu−Ura condition is considered the CNL of overexpression of the target gene if the copy number is significantly lower than that of the empty vector control (which is usually ~100 copies per haploid genome). As the plasmid copy number and the cellular max growth rate under the −Leu−Ura condition are correlated with each other, max growth rate can also be an indicator of the CNL of the target gene. Ideally, in gTOW, the protein level expressed from the target gene increases according to the copy number increase. However, if the transcription factors for the target gene are diluted or if there is feedback in expression regulation, then the copy number increase might not be linearly reflected in the protein level. In this study, we thus designated it on the basis of the overexpression limit measured by gTOW as the "CNL of overexpression" to distinguish the limit of protein overexpression. We previously determined the CNLs of cell cycle regulatory genes in the budding yeast and fission yeast and found that their CNLs were diverse, ranging from less than two to more than 100 (Moriya et al. 2006, 2011).

Several genome-wide analyses revealed the genes that cause cellular dysfunction upon overexpression (Gelperin et al. 2005; Sopko et al. 2006). These analyses were performed using promoter

[7]**Corresponding author**
E-mail hisaom@cc.okayama-u.ac.jp
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.146662.112. Freely available online through the *Genome Research* Open Access option.

23:300–311 © 2013, Published by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/13; www.genome.org

swapping in which each target open reading frame (ORF)/protein is highly expressed by the strong inducible *GAL1* promoter. The results obtained by promoter swapping and gTOW are known to be different (Moriya et al. 2006; Krantz et al. 2009) because the former technique causes absolute overexpression and the latter causes relative overexpression from the native level. The promoter swapping approach is useful for determining what happens when a target protein abundantly exists within the cell. Conversely, it is difficult to argue how much the target is overexpressed when cellular dysfunction is observed. As gTOW increases the copy number of the target gene with its native promoter, this argument is possible. We thus consider that gTOW is a useful method for evaluating the robustness of cellular systems by assessing how much gene expression is fluctuated from the native level when the system halts (Moriya et al. 2012). The advantage of gTOW is that one cannot only isolate genes causing cellular dysfunctions upon overexpression but also quantitate the limits of gene overexpression that are associated with cellular robustness. In addition, we consider that gTOW is useful for evaluating cellular dysfunction triggered by the fluctuation of the gene copy number.

In this study, we performed a genome-wide CNL measurement of genes of the budding yeast *Saccharomyces cerevisiae* using gTOW to reveal the profile of CNLs of all genes in this organism and determine why the yeast cellular systems are sensitive to minor increases in the copy numbers of those genes. First, we isolated 786 genes with significantly low CNLs. Further, we isolated genes with extremely low CNLs (10 or fewer copies per haploid genome), which we designated "yeast dosage-sensitive genes" (DSGs). Our results indicated that the yeast cellular system was robust against copy number variations (overexpression) in most genes but fragile against variations in a specific set of genes. Yeast DSGs tended to encode protein complex components, as well as proteins involved in cytoskeletal organization and intracellular transport. Our experimental evidence suggested that protein burden and stoichiometric imbalance are the primary causes of dosage sensitivity. These findings may have an interesting evolutionary implication in that DSGs function to constrain and secure the integrity of eukaryotic genomes during evolution.

Results

gTOW6000: Analysis of all protein-coding genes in *S. cerevisiae* using gTOW

To analyze all protein-coding genes in the *S. cerevisiae* genome using gTOW, we performed a series of experiments as summarized in Figure 1 (for details, see the Methods). We amplified all protein-coding genes (5806) with their native regulatory regions in the yeast strain BY4741 chromosome using polymerase chain reaction (PCR) and then cloned the genes into pTOWug2-836 (Supplemental Fig. S1; Moriya et al. 2012). Because not all promoter regions were identified, we cloned genes with their upstream and downstream sequences up to their neighboring genes (as an example, see Supplemental Fig. S2A,B). Cells harboring the gTOW plasmids with each target gene were cultivated in −Ura and −Leu−Ura media. We then measured max growth rate under the −Leu−Ura condition using online monitoring of cellular growth, and the plasmid copy numbers under the −Ura and −Leu−Ura conditions using quantitative PCR. We analyzed at least two independent plasmid clones for each gene. The reproducibility between each duplicate is shown in Supplemental Figure S3. To this point, we have succeeded in analyzing >95% of the genes in the

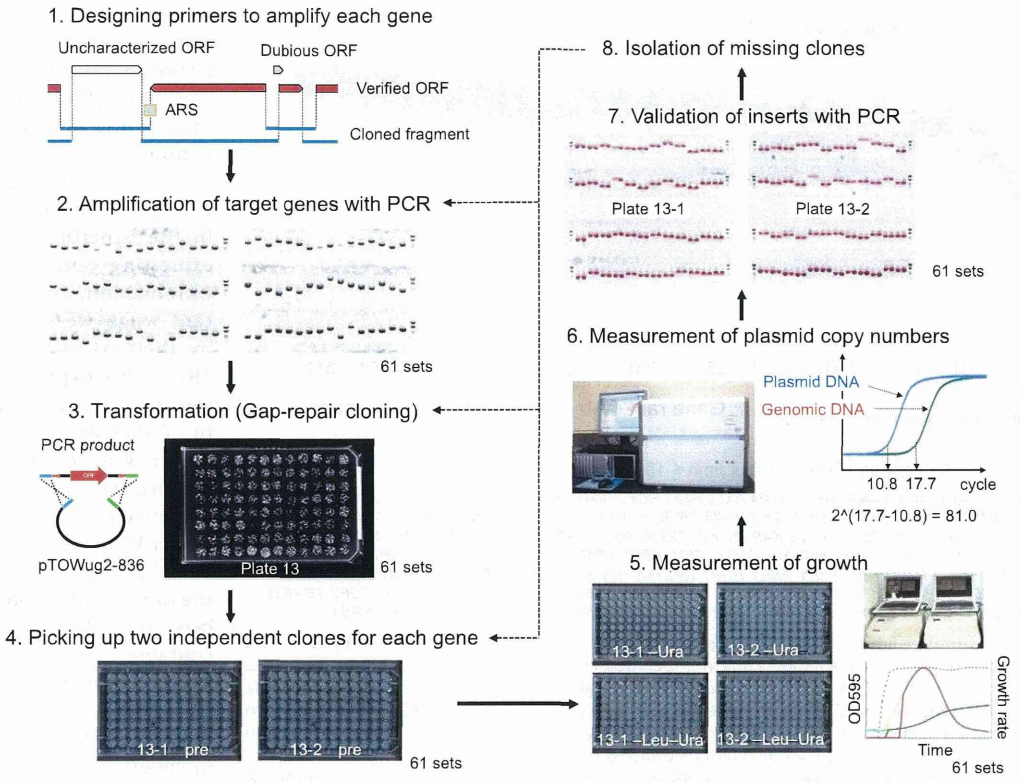

Figure 1. Scheme of genome-wide analysis of protein-coding genes in *S. cerevisiae* with gTOW (gTOW6000). Each step of gTOW6000 is shown. For steps 2–7, the representative data of plate no. 13 are given as an example. The details of each step are described in Methods.

yeast genome (the entire data can be found in Supplemental Table S1). Hereafter, we will refer to this analysis as "gTOW6000."

Figure 2 shows the copy number under the −Leu−Ura condition determined in gTOW6000. gTOW6000 was performed using 96-well microplates. We handled 244 plates, as we analyzed two clones under two culture conditions for each gene. For the purpose of data quality control and to obtain a negative control, several empty vector experiments were performed for each plate (a total of 230 measurements) (Supplemental Table S2). The average of the empty vector experiments is shown as the orange line in Figure 2. To identify genes with significantly lower limits than the empty vector control, we evaluated the copy number data under the −Leu−Ura condition using Student's *t*-test. In total, 919 genes had *P*-values <0.05, and 786 of them had lower copy numbers than the vector average (genes surrounded by a blue-dotted rectangle in Fig. 2). We thus considered the copy numbers of these genes under the −Leu−Ura condition to be their CNLs of overexpression. The average copy number of these genes was less than 85. This finding conversely indicates that the other 5000 genes have similar or higher CNLs than the detectable CNL in gTOW using pTOWug2-836, and suggests that the yeast cellular system is generally robust against a nearly 100-fold increase in the copy number of any one of 80% of its genes. Although some genes displayed much higher limits than the vector average, there was no reproducibility between the two clones (Pearson's correlation coefficient between the duplicates of genes with average copy numbers of >250 was −0.26). We thus concluded that the findings were reflective of experimental errors.

In gTOW, there should be a correlation between the CNLs and max growth rates of low limit genes (Moriya et al. 2006, 2012). In addition, there should be a correlation between the copy numbers under the −Ura and −Leu−Ura conditions (Moriya et al. 2006). These expectations were confirmed in gTOW6000 (Supplemental Table S3). We next calculated the copy number causing 50% growth inhibition in gTOW6000. To reduce the effect of experimental errors, we first calculated the moving averages of max growth rates and CNLs for 100 of the 786 genes with significantly low CNLs (Supplemental Fig. S4A). To approximate the relationship between CNL and max growth rate (Supplemental Fig. S4B), we derived a first dimension equation as follows: CNL = 49.24 × [max growth rate] ($R^2 = 0.98$). From the equation, the copy number that gave 50% growth inhibition (max growth rate = 1.11) was calculated to be 54.7 copies. If the target gene has a very low limit, then the cells expressing the gTOW plasmid cannot grow under the −Leu−Ura condition because they cannot produce sufficient amounts of leucine (Moriya et al. 2006). We next evaluated the lower limit copy number resulting in no growth in gTOW6000. We calculated the moving averages of max growth rates as described previously in this section. For each bin, we then counted the number of genes displaying no growth (max growth rate is set as 0.1; see Methods) in both of the duplicated experiments (i.e., frequency of no-growth) (Supplemental Fig. S5A). To approximate the relationship between frequency of no-growth and CNL (Supplemental Fig. S5B), we derived the following equation: [frequency of no-growth] = $-0.0002 \times CNL^3 + 0.0476 \times CNL^2 - 3.6046 \times CNL + 101.53$ ($R^2 = 0.996$). We used this equation to calculate that a gene with a CNL of 18.4 could not grow in 50% of cases in the gTOW experiment.

By use of genome-wide screening, Sopko et al. (2006) previously isolated 767 *S. cerevisiae* genes that caused cellular growth defects when overexpressed by the *GAL1* promoter. As we isolated a similar number of genes with low CNLs (786 genes), we compared two data sets. As shown in Figure 3A, only 161 of the 786 genes isolated by gTOW6000 overlapped with those in the study by Sopko et al. (2006), although the overlap was significant ($P < 1.5 \times 10^{-8}$, chi-square test). The difference possibly arose from the difference in the experimental systems for overexpressing genes, as is discussed in the Introduction. The difference was significant when we separated isolated genes by their native expression levels (Fig. 3B). Highly expressed genes were significantly isolated as genes with low CNLs in gTOW6000 ($P = 1.322 \times 10^{-15}$ in the Mann-Whitney *U*-test), whereas this finding was not replicated in the study by Sopko et al. (2006) ($P = 0.7378$ in the Mann-Whitney *U*-test). Another difference between the two experiments was the proportions of protein complex members. The 786 genes isolated by gTOW contained significant numbers of protein complex members (Table 1), whereas the 767 genes isolated by Sopko et al. (2006) did not contain many protein complex members (Table 1). This might reflect the fact that protein complex members tend to be highly expressed (Supplemental

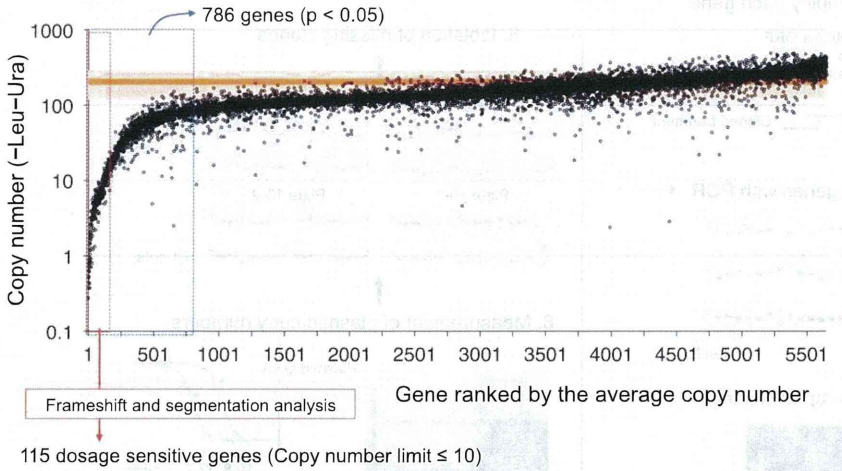

Figure 2. Copy number limits (CNLs) of *S. cerevisiae* genes determined by gTOW analysis. Genes were ordered according to their average copy number determined by gTOW under the −Leu−Ura condition. Each gene has two data points because of the duplication of the experiment. The orange line and the transparent zone around the line indicate the average copy number with the empty vector and the standard deviation, respectively. Genes that showed significantly lower limits than those observed in the vector experiments (786 genes, $P < 0.05$) are surrounded by the blue dotted rectangle. Genes with CNLs of 10 and less (dosage-sensitive genes [DSGs]) are surrounded by the red-dotted rectangle. A confident set of DSGs isolated after frameshift and segmentation analyses (Fig. 4) is shown. The entire data set is given in Supplemental Table S1.

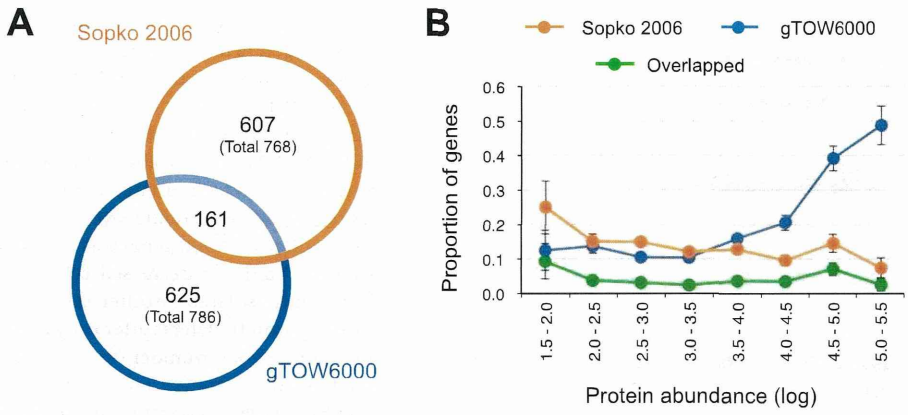

Figure 3. Comparison of gTOW6000 data with data of another overexpression analysis performed using promoter swapping. (*A*) Overlap of genes identified by the overexpression analyses performed by Sopko et al. (2006) and in this study. (*B*) Distribution of genes identified by overexpression analysis ordered by their native protein levels. Each bin contains genes ordered by their native protein levels (Ghaemmaghami et al. 2003). The protein abundance unit is molecules per cell. Error bars, SEM.

Fig. S6). From these results, we considered that gTOW6000 would provide additional clues to understand the cellular effects of gene overexpression, as this method isolated a different subset of genes from previous promoter swapping experiments. Of the 161 overlapped genes (Fig. 3A), the highly expressed genes among the 786 gTOW6000 genes were excluded (Fig. 3B), and the complex members of 767 genes isolated by Sopko et al. (2006) were enriched (Table 1), probably due to the characteristics of the opposite data sets.

Isolation of low limit genes (yeast DSGs)

To further understand the characteristics of low limit genes, we performed additional experiments to isolate a confident set of genes with CNLs of 10 or less. We introduced a frameshift mutation in each of the 182 genes to confirm whether the expression of the protein but not that of the DNA and RNA elements determined the limit (Fig. 4A). Frameshift analysis could also determine whether either of the bidirectionally overlapped genes was the cause of the low CNL (for example, see Supplemental Fig. S7A). Among the 155 genes with CNLs of 20 or less, the frameshift mutants of 140 of these genes displayed more than fivefold higher CNLs than the wild-type genes or their CNLs increased to the vector level (~100 copies) (Fig. 4B; Supplemental Table S4). We thus verified that the original target ORFs of these 140 genes determined the CNLs (denoted as "fs verified" in Supplemental Tables S1, S4).

We further analyzed the 15 genes in frameshift mutants that did not exhibit increased limits (12 of them are indicated by red circles in Fig. 4B). They were categorized as four different types of genes as follows. (1) One of the overlapping ORFs appeared to cause the low limits. The cloned regions contained two overlapping ORFs in cases of *YFL010C/WWM1–YFL010W-A/AUA1* and *YGL167C/PMR1–YGL168W/HUR1*. Because the frameshift mutants of *WWM1* and *PMR1* displayed increased CNLs, we concluded that these genes were responsible for the low CNLs. The result for *YGL167C* is shown in Supplemental Figure S7A as an example. (2) Because both clones containing one of the two neighboring genes (*YNL024C-A/KSH1–YNL025C/SSN8*) exhibited low CNLs but the frameshift mutations did not increase the CNL of either gene (Supplemental Fig. S7B), we concluded that an RNA gene (*NME1*) caused the low limits. (3) For genes for which the frameshift mutations did not increase their CNLs but the cause could not be ascertained from their genome annotations, we segmented the fragments into 5′ UTR and ORF-3′ UTR fragments and measured their limits (Fig. 4A). Both the 5′ and 3′ segmented fragments of *CPS1*, *FHL1*, *GRX3*, *HOM3*, *TPK1*, and *TPK3* (underlined in blue in Fig. 4C) displayed increased copy numbers. These ORFs may have been expressed from ATGs other than the annotated ones. (4) The segmented fragments (ORF-3′ UTR) of *ASE1*, *DIE2*, *IRC8*, and *SFP1* did not exhibit increased CNLs (underlined in red, Fig. 4C). For *DIE2* and *IRC8*, we

Table 1. Characteristics of DSGs

	Protein complex members[a]	Genes with no. of PPIs ≥1[b]	Genes with no. of PPIs ≥5[b]	Intrinsic protein disorder (≥150)[c]	Yeast ohnologs[d]	Essential genes[e]
Yeast DSG[f] (limit ≤10)	69.6% (80/115)	75.7% (87/115)	36.5% (42/115)	23.5% (27/115)	34.8% (40/115)	26.1% (30/115)
P-value	9.05×10^{-7}	7.80×10^{-10}	$1. \times 10^{-10}$	3.43×10^{-7}	2.21×10^{-5}	—
gTOW6000 786 genes	61.5% (483/786)	60.3% (474/786)	25.7% (202/786)	24.8% (195/786)	27.2% (214/786)	27.4% (215/786)
P-value	$<2.2 \times 10^{-16}$	9.37×10^{-15}	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$	3.23×10^{-10}	9.90×10^{-8}
Overlapped 161 genes	62.7% (101/161)	64.6% (104/161)	29.8% (48/161)	32.3% (52/161)	34.8% (56/161)	28.0% (45/161)
P-value	4.06×10^{-5}	1.37×10^{-5}	1.40×10^{-7}	4.91×10^{-12}	3.75×10^{-7}	0.01705
Sopko 767 genes	46.0% (353/767)	57.4% (440/767)	21.1% (162/767)	26.2% (201/767)	21.8% (167/767)	20.9% (160/767)
P-value	—	3.92×10^{-9}	3.08×10^{-7}	$<2.2 \times 10^{-16}$	3.91×10^{-2}	—
All genes	46.5% (2690/5783)	47.4% (2742/5783)	14.9% (863/5783)	13.6% (786/5783)	19.0% (1098/5783)	20.2% (1168/5783)

[a]Protein complex components (mips; ftp://ftpmips.gsf.de/yeast/catalogues/complexcat/complexcat_data_18052006).
[b]Protein–protein interactions (dip; http://dip.doe-mbi.ucla.edu).
[c]Intrinsic protein disorder (Vavouri et al. 2009).
[d]Yeast ohnolog (http://wolfe.gen.tcd.ie/ygob/).
[e]Essential genes (http://www-deletion.stanford.edu/YDPM/YDPM_index.html).
[f]Complete data set for yeast DSGs is given in Supplemental Table S5.