Droplet-Based Microfluidic Systems for High-Throughput Single DNA Molecule Isothermal Amplification and Analysis

Linas Mazutis,[†] Ali Fallah Araghi,[†] Oliver J. Miller,[†] Jean-Christophe Baret,[†] Lucas Frenz,[†] Agnes Janoshazi,[†,◇] Valérie Taly,[†] Benjamin J. Miller,[‡] J. Brian Hutchison,[‡] Darren Link,[‡] Andrew D. Griffiths,[†,*] and Michael Ryckelynck[†,*]

Institut de Science et d'Ingénierie Supramoléculaire (ISIS), Université de Strasbourg, CNRS UMR 7006, 8 allée Gaspard Monge, 67083 Strasbourg Cedex, France, RainDance Technologies, Inc., 44 Hartwell Avenue, Lexington, Massachusetts 02421, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Université de Strasbourg, CNRS UMR 7104, 1 rue Laurent Fries, 67404 Illkirch Cedex, France

We have developed a method for high-throughput isothermal amplification of single DNA molecules in a droplet-based microfluidic system. DNA amplification in droplets was analyzed using an intercalating fluorochrome, allowing fast and accurate "digital" quantification of the template DNA based on the Poisson distribution of DNA molecules in droplets. The clonal amplified DNA in each 2 pL droplet was further analyzed by measuring the enzymatic activity of the encoded proteins after fusion with a 15 pL droplet containing an in vitro translation system.

Digital PCR is based on the Poisson distribution of DNA molecules in microtiter plate wells (eq 1; in which $P(X = k)$ is the probability to have k DNA molecules per well, and λ is the mean number of DNA molecules per well).

$$P(X = k) = \frac{e^{-\lambda}}{k!}\lambda^k \qquad (1)$$

At low λ (<0.3), the vast majority of wells contain no more than a single DNA molecule, and fitting the number of PCR competent wells to eq 1 allows the DNA concentration to be calculated. Digital PCR is used, for example, to detect low concentrations of mutations associated with colorectal cancer for diagnosis.[1] However, the large number of reactions required results in high reagent costs. Reaction volumes have been reduced by ~1000-fold (down to tens of nanolitres) using microfluidic systems by performing PCRs separated spatially in continuous flow[2,3] or within compartments defined by elastomeric valves.[4] However, single

DNA molecules can also be compartmentalized in microdroplets in water-in-oil emulsions,[5] which act as microreactors with volumes down to 1 fL.[6] Digital PCR in emulsions (emulsion PCR) is already used to quantify rare mutations using BEAMing[7] and to prepare the template for two commercialized "next-generation" DNA sequencing systems.[8]

Direct, quantitative screening using emulsion PCR is, however, compromised by the polydispersity of bulk emulsions. Furthermore, it is difficult to add reagents to droplets after they are formed,[6] which limits the range of assays that can be performed on the amplified DNA. However, both of these problems can potentially be overcome using droplet-based microfluidic systems that allow the production of highly monodisperse droplets[9] and pairwise droplet electrocoalescence.[10–14]

Digital PCR has previously been used to quantify DNA in droplets in microfluidic systems.[15–18] This manuscript, however, describes the digital quantification of DNA using a droplet-based microfluidic system and isothermal "hyperbranched rolling circle

* To whom correspondence should be addressed. (A.D.G.) Phone: +33 (0)390 245 171. Fax: +33 (0)390 245 115. E-mail: griffiths@isis.u-strasbg.fr. (M.R.) Phone: +33 (0)390 245 217. Fax: +33 (0)390 245 115. m.ryckelynck@isis.u-strasbg.fr.
    † Institut de Science et d'Ingénierie Supramoléculaire.
    ‡ RainDance Technologies.
    ◇ Institut de Génétique et de Biologie Moléculaire et Cellulaire.

(1) Vogelstein, B.; Kinzler, K. W. Proc. Natl. Acad. Sci. U.S.A. 1999, 96, 9236–9241.
(2) Li, H.; Xue, G.; Yeung, E. S. Anal. Chem. 2001, 73, 1537–1543.
(3) Dettloff, R.; Yang, E.; Rulison, A.; Chow, A.; Farinas, J. Anal. Chem. 2008, 80, 4208–4213.
(4) Ottesen, E. A.; Hong, J. W.; Quake, S. R.; Leadbetter, J. R. Science 2006, 314, 1464–1467.
(5) Tawfik, D. S.; Griffiths, A. D. Nat. Biotechnol. 1998, 16, 652–656.
(6) Griffiths, A. D.; Tawfik, D. S. Trends Biotechnol 2006, 24, 395–402.
(7) Dressman, D.; Yan, H.; Traverso, G.; Kinzler, K. W.; Vogelstein, B. Proc. Natl. Acad. Sci. U.S.A. 2003, 100, 8817–8822.
(8) Mardis, E. R. Annu. Rev. Genomics Hum. Genet. 2008, 9, 387–402.
(9) Christopher, G. F.; Anna, S. L. J. Phys. D: Appl. Phys. 2007, 40, R319–R336.
(10) Chabert, M.; Dorfman, K. D.; Viovy, J. L. Electrophoresis 2005, 26, 3706–3715.
(11) Ahn, K.; Agresti, J.; Chong, H.; Marquez, M.; Weitz, D. A. Appl. Phys. Lett. 2006, 88, 264105.
(12) Link, D.; Grasland-Mongrain, E.; Duri, A.; Sarrazin, F.; Cheng, Z.; Cristobal, G.; Marquez, M.; Weitz, D. Angew. Chem., Int. Ed. 2006, 45, 2556–2560.
(13) Priest, C.; Herminghaus, S.; Seemann, R. Appl. Phys. Lett. 2006, 89, 134101.
(14) Frenz, L.; El Harrak, A.; Pauly, M.; Begin-Colin, S.; Griffiths, A.; Baret, J. C. Angew. Chem., Int. Ed. 2008, 47, 6817–6820.
(15) Beer, N. R.; Hindson, B. J.; Wheeler, E. K.; Hall, S. B.; Rose, K. A.; Kennedy, I. M.; Colston, B. W. Anal. Chem. 2007, 79, 8471–8475.
(16) Beer, N. R.; Wheeler, E. K.; Lee-Houghton, L.; Watkins, N.; Nasarabadi, S.; Hebert, N.; Leung, P.; Arnold, D. W.; Bailey, C. G.; Colston, B. W. Anal. Chem. 2008, 80, 1854–1858.
(17) Kiss, M. M.; Ortoleva-Donnelly, L.; Beer, N. R.; Warner, J.; Bailey, C. G.; Colston, B. W.; Rothberg, J. M.; Link, D. R.; Leamon, J. H. Anal. Chem. 2008, 80, 8975–8981.
(18) Schaerli, Y.; Wootton, R. C.; Robinson, T.; Stein, V.; Dunsby, C.; Neil, M. A.; French, P. M.; Demello, A. J.; Abell, C.; Hollfelder, F. Anal. Chem. 2009, 81, 302–306.

analytical chemistry
feature

Microfluidics: On the Slope of Enlightenment

Rajendrani Mukhopadhyay

Now that the hype has blown over, will microfluidics live up to its promise of providing marketable applications?

"One could argue that miniaturized chemical analysis systems are just a fashionable craze. However, it is difficult to foresee the impact a new technological concept will have, when it is in its early stages of development." —Andreas Manz and colleagues (*Chimia* 1991, 45, 103–105).

Pessimism and optimism walk hand in hand in the field of microfluidics. Some experts feel that the research endeavor is in danger of falling into a rut. Andreas Manz, currently at the University of Freiburg (Germany), feels a twinge of disappointment with the current state of affairs. He had bought into the hope carried by the first wave of patents, scientific publications, and start-up companies that occurred before 2000. But now he thinks, "It's very well possible that microfluidics will end up in a never-ending cycle where it eventually reaches a steady state, just running through all kinds of [academic] applications or all sorts of technical innovations with no major outlet."

Others take a different attitude. Microfluidics, in and of itself, won't sell. What will sell are the solutions it can provide to certain analytical problems. "When you go to the commercial landscape, technology doesn't make money," states the University of Washington's Daniel Chiu, who helped found the company Cellectricon. "Applications make money."

The wildly varying views of the field are dizzying. But Holger Becker at the microfluidic Chip Shop (Germany) explains that this rift is perfectly normal. He says that microfluidics is tracking the Gartner Hype Cycle (*Med. Device Technol.* 2008, 19, 21–24).

THE HYPE CYCLE

Gartner, an information technology research and advisory company, introduced their Hype Cycle Model in 1995. The model has five stages: a technology trigger, a peak of inflated expectations, a trough of disillusionment, a slope of enlightenment, and finally, a plateau of productivity. Over the years, Gartner has applied the model to many arenas, including the automotive industry, Web 2.0, and higher-education institutions.

The microfluidics "technology trigger" happened in the early 1990s when the early publications by Manz and his colleagues kicked off the field. Everyone agrees that the field's "peak of
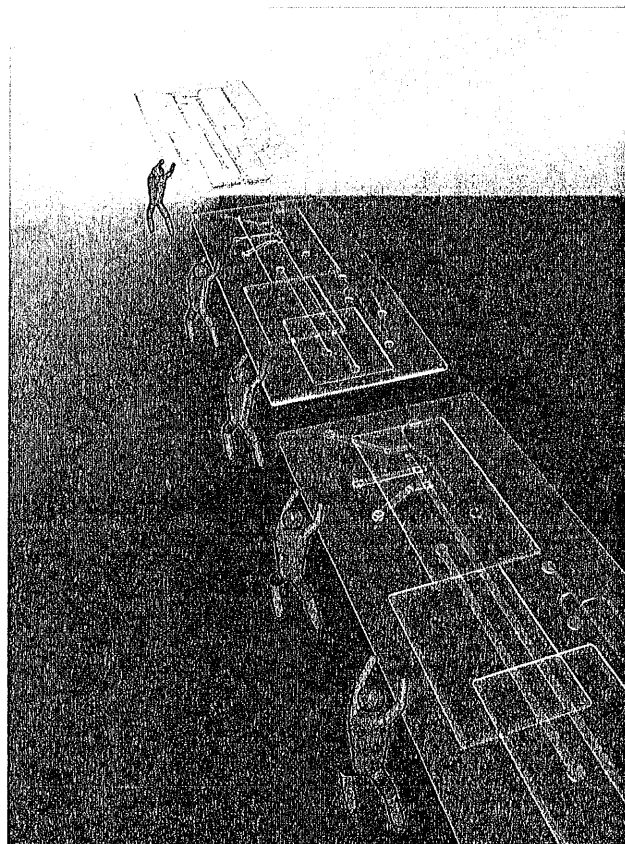
inflated expectations" crested later that decade when the hype about what microfluidics could achieve was at its zenith. According to Gartner, during such a peak "the early publicity produces a number of success stories, often accompanied by scores of failures." In the case of microfluidics, the poster child of a successful application was the Agilent Technologies–Caliper Life Sciences Bioanalyzer.

Microfluidics experts think that the field's "trough of disillusionment" probably happened in the early 2000s, when the technology failed to meet inflated expectations. "Because of the bad rap microfluidics got at the end of the 1990s, with the promises and the lack of delivery, it was a very challenging funding environment for microfluidics entrepreneurs in the wake of that,"

Bacterial cell curvature through mechanical control of cell growth

Matthew T Cabeen[1], Godefroid Charbon[1], Waldemar Vollmer[2], Petra Born[2], Nora Ausmees[3], Douglas B Weibel[4] and Christine Jacobs-Wagner[1,5,6,]*

[1]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT, USA, [2]Institute for Cell and Molecular Biosciences, Medical School, Newcastle University, Newcastle upon Tyne, UK, [3]Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden, [4]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA, [5]Section of Microbial Pathogenesis, Yale School of Medicine, New Haven, CT, USA and [6]Howard Hughes Medical Institute, Yale University, New Haven, CT, USA

The cytoskeleton is a key regulator of cell morphogenesis. Crescentin, a bacterial intermediate filament-like protein, is required for the curved shape of *Caulobacter crescentus* and localizes to the inner cell curvature. Here, we show that crescentin forms a single filamentous structure that collapses into a helix when detached from the cell membrane, suggesting that it is normally maintained in a stretched configuration. Crescentin causes an elongation rate gradient around the circumference of the sidewall, creating a longitudinal cell length differential and hence curvature. Such curvature can be produced by physical force alone when cells are grown in circular microchambers. Production of crescentin in *Escherichia coli* is sufficient to generate cell curvature. Our data argue for a model in which physical strain borne by the crescentin structure anisotropically alters the kinetics of cell wall insertion to produce curved growth. Our study suggests that bacteria may use the cytoskeleton for mechanical control of growth to alter morphology.

*Corresponding author. Department of Molecular, Cellular, and Developmental Biology, Yale University, KBT 1032, PO Box 208103, New Haven, CT 06520, USA. Tel.: +203 432 5170; Fax: +203 432 9960; E-mail: christine.jacobs-wagner@yale.edu

Introduction

The cytoskeleton has a central function in the morphogenesis of eukaryotes and prokaryotes alike. It can achieve this function in two main ways. One is by directly deforming the cell membrane, as in lamella (Pollard and Borisy, 2003), and the other is by affecting cell growth (Smith and Oppenheimer, 2005; Cabeen and Jacobs-Wagner, 2007; Fischer *et al*, 2008). In bacteria, as in plants and fungi, cell growth depends on cell wall growth. The bacterial peptidoglycan cell wall is a covalently closed meshwork of rigid glycan strands cross-linked by relatively flexible peptide bridges (Vollmer *et al*, 2008). It thereby forms a strong but elastic fabric that resists internal turgor pressure and prevents cell bursting. Under normal growth conditions, cells are turgid, as the force exerted by turgor pressure maintains the peptidoglycan in a stretched configuration (Koch, 1984; Koch *et al*, 1987; Baldwin *et al*, 1988). To expand, bonds in the peptidoglycan must be hydrolyzed to generate insertion sites for new wall material. The stretching from turgor pressure helps overcome molecular cohesion, facilitating bond hydrolysis and thereby providing a source of energy for cell wall growth (Koch *et al*, 1982; Harold, 2002). To preserve the integrity of the peptidoglycan fabric, cell wall hydrolases and synthases coordinate their activities by forming complexes (Romeis and Höltje, 1994; von Rechenberg *et al*, 1996; Höltje, 1998; Schiffer and Höltje, 1999; Vollmer *et al*, 1999; Bertsche *et al*, 2006).

The location and timing of new wall insertion is critical for generating nonspherical shapes, as turgor pressure is nondirectional (Harold, 2007; den Blaauwen *et al*, 2008). In most rod-shaped bacteria, elongation occurs by insertion of peptidoglycan material along the sidewall of the cell, with little to no insertion at the poles (de Pedro *et al*, 1997; Daniel and Errington, 2003). Cytoskeletal proteins of the actin family (MreB and its homologs), which form helical structures along the long axis of the cell beneath the cytoplasmic membrane (Jones *et al*, 2001; Shih *et al*, 2003; Figge *et al*, 2004), are thought to spatially restrict insertion of new peptidoglycan to the sidewall (Daniel and Errington, 2003; Figge *et al*, 2004; Dye *et al*, 2005; Kruse *et al*, 2005; Carballido-Lopez *et al*, 2006; Divakaruni *et al*, 2007).

Vibrioid (curved rod) or helical morphologies are also common among bacteria, but their morphogenetic mechanisms are largely unresolved. It is known that the intermediate filament-like protein crescentin is required for the crescent shape of *Caulobacter crescentus* as in its absence, the cells are straight rod shaped (Ausmees *et al*, 2003). Crescentin localizes along the inner cell curvature, beneath the cytoplasmic membrane, but how crescentin confers cell curvature remains unclear.

Physical modelling has proposed that filaments attached to the cell wall could produce a tension leading to curved or helical morphology (Wolgemuth *et al*, 2005). Our findings suggest a model in which the attachment of a crescentin structure at the cell envelope results in a compressive force that sets up a gradient of growth rate across the short cell axis to contribute to cell curvature.

Results

Crescentin assembles into a single coiled structure that is maintained in an extended configuration along the membrane

Although crescentin polymerizes into filaments *in vitro* (Ausmees *et al*, 2003), it was unclear whether it also formed

GENETICS

DinB Upregulation Is the Sole Role of the SOS Response in Stress-Induced Mutagenesis in *Escherichia coli*

Rodrigo S. Galhardo[*], Robert Do[*], Masami Yamada[†], Errol C. Friedberg[‡], P. J. Hastings[*], Takehiko Nohmi[†] and Susan M. Rosenberg[*,§,1]

[*] Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030-3411, [‡] Department of Pathology, University of Texas Southwestern Medical Center, Dallas, Texas 5390-9072, [†] Division of Genetics and Mutagenesis, National Institute of Health Sciences, Tokyo 158-8501, Japan and [§] Departments of Biochemistry and Molecular Biology and Molecular Virology and Microbiology, The Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas 77030-3411

[1] *Corresponding author:* Department of Molecular and Human Genetics, Baylor College of Medicine, 1 Baylor Plaza, Room S809, Mail Stop BCM225, Houston, TX 77030-3411.
E-mail: smr@bcm.edu

Stress-induced mutagenesis is a collection of mechanisms observed in bacterial, yeast, and human cells in which adverse conditions provoke mutagenesis, often under the control of stress responses. Control of mutagenesis by stress responses may accelerate evolution specifically when cells are maladapted to their environments, *i.e.*, are stressed. It is therefore important to understand how stress responses increase mutagenesis. In the *Escherichia coli* Lac assay, stress-induced point mutagenesis requires induction of at least two stress responses: the RpoS-controlled general/starvation stress response and the SOS DNA-damage response, both of which upregulate DinB error-prone DNA polymerase, among other genes required for Lac mutagenesis. We show that upregulation of DinB is the only aspect of the SOS response needed for stress-induced mutagenesis. We constructed two *dinB*(o^c) (operator-constitutive) mutants. Both produce SOS-induced levels of DinB constitutively. We find that both *dinB*(o^c) alleles fully suppress the phenotype of constitutively SOS-"off" *lexA*(Ind⁻) mutant cells, restoring normal levels of stress-induced mutagenesis. Thus, *dinB* is the only SOS gene required at induced levels for stress-induced point mutagenesis. Furthermore, although spontaneous SOS induction has been observed to occur in only a small fraction of cells, upregulation of *dinB* by the *dinB*(o^c) alleles in all cells does not promote a further increase in mutagenesis, implying that SOS induction of DinB, although necessary, is insufficient to differentiate cells into a hypermutable condition.

The Dienes Phenomenon: Competition and Territoriality in Swarming *Proteus mirabilis*[▽]

A. E. Budding,[1][†][*] C. J. Ingham,[2][†] W. Bitter,[1] C. M. Vandenbroucke-Grauls,[1] and P. M. Schneeberger[2]

VU Medical Centre, Amsterdam, The Netherlands,[1] and Jeroen Bosch Hospital, 's Hertogenbosch, The Netherlands[2]

When two different strains of swarming *Proteus mirabilis* encounter one another on an agar plate, swarming ceases and a visible line of demarcation forms. This boundary region is known as the Dienes line and is associated with the formation of rounded cells. While the Dienes line appears to be the product of distinction between self and nonself, many aspects of its formation and function are unclear. In this work, we studied Dienes line formation using clinical isolates labeled with fluorescent proteins. We show that round cells in the Dienes line originate exclusively from one of the swarms involved and that these round cells have decreased viability. In this sense one of the swarms involved is dominant over the other. Close cell proximity is required for Dienes line formation, and when strains initiate swarming in close proximity, the dominant Dienes type has a significant competitive advantage. When one strain is killed by UV irradiation, a Dienes line does not form. Killing of the dominant strain limits the induction of round cells. We suggest that both strains are actively involved in boundary formation and that round cell formation is the result of a short-range killing mechanism that mediates a competitive advantage, an advantage highly specific to the swarming state. Dienes line formation has implications for the physiology of swarming and social recognition in bacteria.

The gram-negative bacterium *Proteus mirabilis* is well known for its ability to differentiate into hyperflagellated, motile, and elongated swarmer cells that rapidly spread over a surface. When cultured on a nutrient agar plate, a strain of *P. mirabilis* typically is able to colonize the whole plate within 24 h. This phenomenon is both of interest in terms of the differentiation and survival strategy of the organism and of practical importance, as contamination of agar plates by swarming *P. mirabilis* is a common problem in diagnostic microbiology. When two different strains of *P. mirabilis* swarm on the same agar plate, a visible demarcation line with lower cell density forms at the intersection, and this line is known as a Dienes line (5, 6). A Dienes line is seen when both strains are swarming; it is not a property of the smaller vegetative cells (5). When two identical isolates meet, the swarming edges merge without formation of a Dienes line. This phenomenon has been used in epidemiological typing of clinical isolates (4, 20, 23, 28) and raises interesting questions concerning its mechanism and biological importance. Dienes typing in the clinical environment suggests that the number of Dienes types is large; 81 types were found in one study alone (25). Incompatibility between swarming strains may not be unique to *P. mirabilis*; a comparable process also appears to occur in *Pseudomonas aeruginosa* (19). Like many other bacteria, some *P. mirabilis* strains produce bacteriocins, termed proticines (3). It has been shown that Dienes line formation is not directly caused by a proticine, nor has any other secreted substance or cellular lysate been found to contribute to this phenomenon (27). There does, however, seem to be a circumstantial link between proticine production and Dienes line formation. In the 1970s Senior typed strains of *P. mirabilis* based on their proticine production and sensitivity (23, 24). He found that proticine production is not related to proticine sensitivity (apart from strains being resistant to their own proticine) but that there is a good correlation between a combined proticine production-sensitivity (P/S) type and Dienes line formation. Strains with the same P/S type do not form a Dienes line, whereas strains with different P/S types do. The more closely related the P/S types of two strains, the less clearly defined the Dienes line is, suggesting that relatedness of strains plays a central role. In contrast, a strain of *Proteus vulgaris* has been shown to be Dienes compatible with a strain of *P. mirabilis* with the same P/S type (24). Furthermore, Dienes incompatibility between otherwise identical strains can be triggered by phage lysogeny (2). In contrast to P/S type, the polysaccharide (O) or flagellar (H) serotype has been shown to have no relationship to Dienes line formation (25, 27).

Research into the mechanism governing Dienes line formation in *Proteus* has revealed some important features. The Dienes line region contains many large and often rounded cells (5). The nature of these cells remains controversial. Dienes suggested that round cells originated from both strains but that viable round cells always originated from one of the two intersecting strains. He concluded from this that nonviable round cells should therefore be cells of the other strain (6). In contrast, Wolstenhome suggested that the round cells originated mostly from one of the two swarms involved and that only 50% of these cells were viable (32). It has also been noted that round cells occasionally seem to develop into stable L forms lacking a full cell wall and that they can grow to form tiny L-type colonies (5). Additionally, extracellular DNase has been found at the site of the Dienes line (1). The presence of this enzyme has been interpreted to be a result of cell lysis, as *P. mirabilis* is known to contain large amounts of DNase (22, 26). Recently, a cluster of six genes, termed *ids* (identification of

---

* Corresponding author. Mailing address: VU Medical Centre, Department of Medical Microbiology and Infection Prevention, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands. Phone: 0031-20-4440457. Fax: 0031-20-4440473. E-mail: d.budding@vumc.nl.

† A.E.B. and C.J.I. contributed equally to this work.

Gene Expression Profiling and the Use of Genome-Scale In Silico Models of *Escherichia coli* for Analysis: Providing Context for Content[∇]

Nathan E. Lewis,[1] Byung-Kwan Cho,[1] Eric M. Knight,[2] and Bernhard O. Palsson[1]*

*Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive, Mail Code 0412, La Jolla,
California 92093-0412,[1] and Center for Systems Biology, University of Iceland, Vatnsmyravegar 16,
101 Reykjavik, Iceland[2]*

One of the most widely used high-throughput technologies is the oligonucleotide microarray. From the initial development of microarrays, high expectations were held for their use to aid in answering biological questions, due to their ability to measure mRNA abundances on a genome scale. However, accumulating experience is revealing that even when questions of sample preparation, data processing, and dealing with the inherently noisy data (81) are set aside, the large amount of data generated has proven difficult to analyze and interpret (12). It is also often challenging to narrow down specific novel findings based solely on expression profiling data.

Here, we present a downloadable compendium of gene expression profiles for *Escherichia coli* and discuss the experience from one lab in which expression profiling data have been employed in a myriad of studies of *E. coli*. We will try to address two classes of expression profiling data usage: (i) how expression profiling can be analyzed using more traditional statistical methods to provide biological understanding and (ii) how genome-scale models form a context within which expression profiling data content increases in value.

THE DATA

We present a database of 213 expression profiles produced in our laboratory and used in many of the case studies reviewed here. These profiles represent measurements for about 70 combinations of variations in experimental conditions and genetic variations in *E. coli* K-12 MG1655. In this set of experiments, the following parameters were varied: (i) the carbon source, (ii) the terminal electron acceptor, (iii) the temperature, (iv) the number of days the bacteria were grown at mid-log phase, and (v) the genotype (wild-type and gene deletion strains were used). These data and the corresponding minimum information about a microarray experiment (9) may be freely downloaded at http://systemsbiology.ucsd.edu/In_Silico _Organisms/E_coli/E_coli_expression2.

LEVELS OF ANALYSIS

Much insight into the function of *E. coli* has been gained over the past decade through the statistical analysis of cDNA and oligonucleotide arrays. Methods such as the ge-

nome-scale analysis of differential gene expression patterns (21), clustering and classification (17), and gene set enrichment (76) have been successfully deployed. Overall, our experience shows that informative results can also be obtained if data are analyzed in the context of a genome-scale network reconstruction and with the use of an in silico model. Furthermore, greater success in data analysis has been achieved on a fine-grain level for more specific hypothesis generation and assessment than on a coarse-grain genome-scale level. The lessons learned from the analysis of this data set can be classified by the granularity of the analysis and the level of use of the in silico model.

(i) **Genome-scale analysis without using a modeling framework.** *E. coli* grown to mid-log phase on various carbon sources and/or with various gene knockouts has been shown to usually evolve a growth phenotype as predicted computationally. However, only a modest level of understanding of the optimal growth phenotype was obtained through the analysis of gene expression data alone.

(ii) **Regulon- and network-level analysis in the modeling framework.** Several studies in which computational models have been tested to predict global properties of *E. coli* metabolism and transcriptional regulation, to suggest pathway usage in strains that do not grow at the optimal rate, and to predict gene expression changes under growth condition shifts have been conducted. Here, expression profiling data have played a beneficial role in both validating computational predictions and providing necessary information to predict novel regulatory interactions.

(iii) **Gene-level analysis using in silico models.** The most informative use of gene expression profiling data was found when simulations predicted growth phenotypes that conflicted with experimental observations (growth/no growth), suggesting that undiscovered pathways still existed. Microarrays were used in this setting to identify individual genes that may contribute to computationally predicted pathways.

Each of these levels of analysis will be discussed and examples will be provided below.

GENOME-LEVEL ANALYSIS WITHOUT USING IN SILICO MODELS

Characterizing disparate paths to common end points in adaptive evolution. Wild-type *E. coli* K-12 has been evolved to improve the growth rate by using serial passaging in the mid- to late log phase of growth (45, 47). Interestingly, when parallel cultures are evolved in this way, their evolutionary paths are not the same, though they usually evolve to have similar growth

The Stationary-Phase Sigma Factor σ^S Is Responsible for the Resistance of *Escherichia coli* Stationary-Phase Cells to *mazEF*-Mediated Cell Death[▽]

Ilana Kolodkin-Gal and Hanna Engelberg-Kulka*

Department of Molecular Biology, the Hebrew University-Hadassah Medical School, Jerusalem 91120, Israel

Escherichia coli mazEF is a toxin-antitoxin gene module that mediates cell death during exponential-phase cellular growth through either reactive oxygen species (ROS)-dependent or ROS-independent pathways. Here, we found that the stationary-phase sigma factor σ^S was responsible for the resistance to *mazEF*-mediated cell death during stationary growth phase. Deletion of *rpoS*, the gene encoding σ^S from the bacterial chromosome, permitted *mazEF*-mediated cell death during stationary growth phase.

Toxin-antitoxin systems have been found on the chromosomes of many bacteria (8, 10, 23, 27). One of the best studied among these chromosomal toxin-antitoxin systems is *Escherichia coli mazEF*, which was the first to be described as regulatable and responsible for bacterially programmed cell death (2). *E. coli mazEF* is located downstream from the *relA* gene (18, 20), specifying for ppGpp synthase (28). *mazF* specifies the stable toxin MazF, while *mazE* specifies the labile antitoxin MazE, degraded in vivo by the ATP-dependent ClpPA serine protease (2). MazF is a sequence-specific endoribonuclease that preferentially cleaves single-stranded mRNAs at ACA sequences (36, 37) and thereby inhibits translation (3, 37). MazE counteracts the action of MazF. Because MazE is a labile protein, prevention of MazF-mediated action requires the continuous production of MazE. Therefore, stressful conditions that prevent the expression of the chromosomally borne *mazEF* module permit the formation of free MazF and thereby cell death. These stressful conditions include (i) the use of antibiotics that are general inhibitors of transcription and/or translation such as rifampin, chloramphenicol, and spectinomycin (31); (ii) extreme amino acid starvation, leading to the production of ppGpp that inhibits *mazEF* transcription (2, 7); and (iii) DNA damage caused by thymine starvation (32) as well as by DNA-damaging agents like mitomycin C or nalidixic acid (11). The use of these antibiotics and other stressful conditions are well known to cause bacterial cell death (1, 5); we found that such cell death takes place through the action of the *mazEF* module (31, 32). All the groups of stressful conditions were found to trigger *mazEF*-mediated cell death by preventing the continuous synthesis of MazE and thereby reducing its level (2, 31, 32). We were surprised to find that *mazEF*-mediated cell death occurs at the exponential stage of growth but does not occur during stationary phase (11).

We have recently reported that the activation of *E. coli mazEF* by using stressful conditions causes the generation of reactive oxygen species (ROS) (15). ROS have been previously implicated in programmed cell death in eukaryotes (21, 29), including in yeast (12, 17), in the life span of several organisms (25, 33), in the senescence of bacteria (6), and in the mode of action of some antibiotics (13–15). It was previously reported that the stationary-phase sigma factor σ^S, encoded by *rpoS* (16, 19), positively regulates the formation of catalase and is responsible for the elevated levels of this enzyme during stationary growth phase (24, 34, 35). Since catalase detoxifies ROS, we asked whether resistance of stationary-phase cells to *mazEF*-mediated cell death was caused by the elevated levels of catalase produced at that time. So, we tested whether deleting *rpoS* from *E. coli* cells would lead to their death through the *mazEF* system during stationary growth phase. Indeed, as we have predicted, in Δ*rpoS* cells we observed *mazEF*-mediated cell death even during stationary phase of growth.

We used strain MC410*relA*[+] and its Δ*mazEF::kan* derivative (9) and strain MC4100*relA*[+]Δ*rpoS*, which we constructed by P1 transduction from *E. coli* strain K-38Δ*rpoS::tet* (kindly provided by Shosh Altuvia), and its Δ*mazEF* derivative, which we constructed by PCR deletion (4). We used plasmid pQE*katE* (15), bearing the catalase-specifying *katE* gene, which is continuously expressed in the strains described here.

We grew the bacteria in liquid M9 minimal medium with 1% glucose and a mixture of amino acids (10 μg/ml each) (22) and then plated them on rich LB agar plates, as we have described previously (11).

Nalidixic acid, mitomycin C, trimethoprim, rifampin, serine hydroxamate, chloramphenicol, spectinomycin, Trizma base, sodium dodecyl sulfate, DNase, and RNase were obtained from Sigma (St. Louis, MO). Lysozyme was obtained from the United States Biochemical Corporation (Cleveland, OH). Ampicillin was obtained from Biochemie GmbH (Kundl, Austria). Carbonylated proteins were detected using the chemical and immunological reagents from the OxyBlot oxidized protein detection kit (Chemicon, Temecula, CA). Nitrocellulose membranes were obtained from Pall Corporation (New York). Luminol and *p*-coumaric acid were obtained from Sigma (St. Louis, MO), hydrogen peroxidase solution was obtained from Merck (NJ), and AnaeroGen bags were obtained from Gamidor Diagnostics (Petach Tikva, Israel).

* Corresponding author. Mailing address: Department of Molecular Biology, the Hebrew University-Hadassah Medical School, P.O. Box 12272, Jerusalem 91120, Israel. Phone: 972-2-6758250. Fax: 972-2-678-4010. E-mail: hanita@cc.huji.ac.il.

NOTES

The Conserved Sporulation Protein YneE Inhibits DNA Replication in *Bacillus subtilis*[∇][†]

Lilah Rahn-Lee,[1] Boris Gorbatyuk,[1] Ole Skovgaard,[2] and Richard Losick[1]*

Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts,[1] and Department of Science, Systems and Models, Roskilde University, Roskilde, Denmark[2]

Cells of *Bacillus subtilis* triggered to sporulate under conditions of rapid growth undergo a marked decrease in chromosome copy number, which was partially relieved by a mutation in the sporulation-induced gene *yneE*. Cells engineered to express *yneE* during growth were impaired in viability and produced anucleate cells. We conclude that YneE is an inhibitor of DNA replication.

The ability of bacteria to adapt to a variety of growth conditions demands mechanisms for controlling DNA replication and coordinating chromosome duplication with the cell cycle. Bacteria initiate one and only one round of DNA replication at the origin of chromosomal DNA replication, *oriC*, for each cycle of cell division (1). During slow growth, the single copy of *oriC* is duplicated to create two copies of *oriC*. In contrast, under conditions of rapid growth, when the time required to duplicate a chromosome is longer than the generation time, bacteria inherit chromosomes undergoing active replication with multiple replication forks. In this case, replication initiation occurs synchronously across all origins, leading to multifork replication with 2^n ($n = 1,2,3$) copies of *oriC* present in each cell (15). Multifork replication ensures that each newly divided cell will receive at least one chromosome. The transition from one growth condition to another, such as when nutrients become limiting and cells enter stationary phase, requires mechanisms to adjust the frequency of replication initiation (10). One noteworthy but poorly understood example of replication control is exhibited by *Bacillus subtilis* during the entry into sporulation. Sporulating cells complete a final round of replication before undergoing a process of asymmetric division that results in a forespore and a mother cell (11, 14). Each of these progeny cells ordinarily inherits a single chromosome, and mature spores, which are derived from the forespore, are known to contain only one chromosome (9, 14).

Cells normally enter the pathway to sporulate in response to nutrient limitation. We have previously shown, however, that artificial induction of synthesis of the kinase KinA causes efficient sporulation even during growth in rich medium (7). KinA transfers phosphate groups into a multicomponent phosphorelay that phosphorylates the response regulator Spo0A,

the master regulator for entry into sporulation (2). The discovery that rapidly growing cells can be triggered to sporulate with high efficiency implies that *B. subtilis* must have an active mechanism to shut down DNA replication, given that rapidly growing cells contain multiple replication forks. An appealing hypothesis for how such a mechanism might work stems from the observation that the *B. subtilis* origin of replication contains multiple binding sites for phosphorylated Spo0A (Spo0A~P) and that many of these sites overlap with binding sites for the replication initiation protein DnaA (5). This suggests that Spo0A~P may impede replication by directly blocking the action of DnaA. It has not been established, however, whether Spo0A~P binds to the origin region in vivo and, if so, whether this binding blocks the initiation of replication. We wondered, therefore, whether Spo0A~P might additionally, or alternatively, block replication indirectly by turning on the expression of a gene or genes whose product is an inhibitor of replication.

In previous work, we identified genes under the direct control of Spo0A, a small number of which encode proteins of unknown function and mutants of which do not exhibit conspicuous defects in spore formation (7, 13). We used this as a starting point to investigate whether any previously uncharacterized genes under Spo0A~P control inhibit DNA replication. Here we report the discovery that a Spo0A~P-induced gene, *yneE*, for which we introduce the name *sirA* (for *s*porulation *i*nhibitor of *r*eplication) encodes a protein that inhibits DNA replication.

To determine if the induction of sporulation by actively growing cells blocks replication, we turned on the synthesis of KinA in cells in the midexponential phase of growth by placing *kinA* under the control of an IPTG (isopropyl-β-D-thiogalactopyranoside)-inducible promoter and treating the cells with IPTG. We measured chromosome content by staining for DNA and analyzing the cells with a flow cytometer. Either samples were fixed immediately for a "snapshot" of the total amount of DNA present at the time that the sample was collected or the cells were allowed to continue growing in the presence of chloramphenicol. Because chloramphenicol blocks protein synthesis, it allows ongoing rounds of replication to

---

* Corresponding author. Mailing address: Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Ave., Cambridge, MA 02138. Phone: (617) 495-4905. Fax: (617) 496-4642. E-mail: Losick@mcb.harvard.edu.
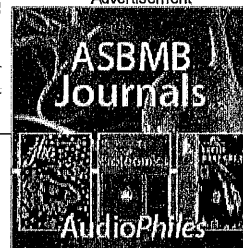
jbc ONLINE

HOME HELP FEEDBACK SUBSCRIPTIONS ARCHIVE SEARCH TABLE OF CONTENTS
Institution: Harvard Libraries | Sign In via User Name/Password

Reflections

Genetic Suppressors and Recovery of Repressed Biochemical Memory

Jon Beckwith

From the Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115

This Article

▸ **Full Text (PDF)**
▸ **All Versions of this Article:**
  284/19/12585 *most recent*
  X800017200v1
▸ **Submit a Letter to Editor**
▸ **Alert me when this article is cited**
▸ **Alert me when eLetters are posted**
▸ **Alert me if a correction is posted**
▸ **Citation Map**

Services

▸ **Email this article to a friend**
▸ **Similar articles in this journal**
▸ **Similar articles in PubMed**
▸ **Alert me to new issues of the journal**
▸ **Download to citation manager**
▸ **Request Permissions**

Google Scholar

▸ **Articles by Beckwith, J.**

PubMed

▸ **PubMed Citation**
▸ **Articles by Beckwith, J.**

Social Bookmarking

What's this?

▸   INTRODUCTION

Many of the Reflections articles in this journal have focused on particular biological problems that the authors have devoted their careers to or on particular techniques that have facilitated their biological studies. Although I began my career as a biochemist, one of the most common threads to my research has been the exploitation of a genetic approach, the analysis of suppressor mutations. Particularly in the last 20 years, the success of this approach has led me into

biological problems that required a remembering or relearning of my early training in biochemistry. This article will describe the voyage from chemistry and biochemistry to genetics and then to a recovery of remnants of my biochemical memory.

As a graduate student working under Lowell Hager at Harvard in the late 1950s, my thesis research was chemical and biochemical, as I performed the chemical synthesis and studied the biosynthesis of the chlorinated mold product caldariomycin (2,2-dichloro-1,3-cyclopentanediol) (1–3). With his postdoctoral fellow Paul Shaw, Lowell had discovered an enzyme, chloroperoxidase, that was capable of

jbc ONLINE

This Article

► Full Text
► Full Text (PDF)
► Supplemental Data
► All Versions of this Article: 284/19/13256 most recent M807095200v1
► Submit a Letter to Editor
► Alert me when this article is cited
► Alert me when eLetters are posted
► Alert me if a correction is posted
► Citation Map

Structural and Functional Similarities between a Ribulose-1,5-bisphosphate Carboxylase/Oxygenase (RuBisCO)-like Protein from *Bacillus subtilis* and Photosynthetic RuBisCO[*][S]

Yohtaro Saito[‡1], Hiroki Ashida[‡], Tomoko Sakiyama[‡],
Nicole Tandeau de Marsac[§], Antoine Danchin[¶],
Agnieszka Sekowska[¶2], and Akiho Yokota[‡3]

From the [‡] Nara Institute of Science and Technology, Graduate School of Biological Sciences, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan and the Departments of [§] Microbiology and [¶] Genetics of Bacterial Genomes, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France

Services

► Email this article to a friend
► Similar articles in this journal
► Similar articles in PubMed
► Alert me to new issues of the journal
► Download to citation manager
► Request Permissions

Google Scholar

► Articles by Saito, Y.
► Articles by Yokota, A.

PubMed

► PubMed Citation
► Articles by Saito, Y.
► Articles by Yokota, A.

Social Bookmarking

What's this?

The sequences classified as genes for various ribulose-1,5-bisphosphate (RuBP) carboxylase/oxygenase (RuBisCO)-like proteins (RLPs) are widely distributed among bacteria, archaea, and eukaryota. In the phylogenic tree constructed with these sequences, RuBisCOs and RLPs are grouped into four separate clades, forms I-IV. In RuBisCO enzymes encoded by form I, II, and III sequences, 19 conserved amino acid residues are essential for CO_2 fixation; however, 1-11 of these 19 residues are substituted with other amino acids in form IV RLPs. Among form IV RLPs, the only enzymatic activity detected to date is a 2,3-diketo-5-methylthiopentyl 1-phosphate (DK-MTP-1-P) enolase reaction catalyzed by *Bacillus subtilis*, *Microcystis aeruginosa,* and *Geobacillus kaustophilus* form IV RLPs. RLPs from *Rhodospirillum rubrum*, *Rhodopseudomonas palustris*, *Chlorobium tepidum*, and *Bordetella bronchiseptica* were inactive in the enolase reaction. DK-MTP-1-P enolase activity of *B. subtilis* RLP required Mg^{2+} for catalysis and, like RuBisCO, was stimulated by CO_2. Four residues that are essential for the enolization reaction of RuBisCO, Lys^{175}, Lys^{201}, Asp^{203}, and Glu^{204}, were conserved in RLPs and were essential for DK-MTP-1-P enolase catalysis. Lys^{123}, the residue conserved in DK-MTP-1-P enolases, was also essential for *B. subtilis* RLP enolase activity. Similarities between the active site structures of RuBisCO and *B. subtilis* RLP were examined by analyzing the effects of structural analogs of RuBP on DK-MTP-1-P enolase activity. A transition state analog for the RuBP carboxylation of RuBisCO was a competitive inhibitor in the DK-MTP-1-P enolase reaction with a K_i value of 103 μM. RuBP and D-phosphoglyceric acid, the substrate and product, respectively, of RuBisCO, were weaker competitive inhibitors. These results suggest that the amino acid residues utilized in the *B. subtilis* RLP enolase reaction are the same as those utilized in the RuBisCO RuBP enolization reaction.

jbc ONLINE

HOME HELP FEEDBACK SUBSCRIPTIONS ARCHIVE SEARCH TABLE OF CONTENTS
Institution: Harvard Libraries | Sign In via User Name/Password
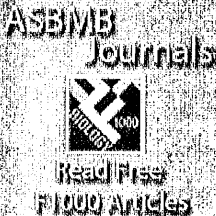
QUICK SEARCH: [advanced]
        Author: Keyword(s):
Go
Year: Vol: Page:

Inhibitory Mechanism of *Escherichia coli* RelE-RelB Toxin-Antitoxin Module Involves a Helix Displacement Near an mRNA Interferase Active Site[*][S]

This Article

▶ Full Text
▶ Full Text (PDF)
▶ Supplemental Data
▶ All Versions of this Article:
   284/21/14628 most recent
   M809656200v1
▶ Submit a Letter to Editor
▶ Alert me when this article is cited
▶ Alert me when eLetters are posted
▶ Alert me if a correction is posted
▶ Citation Map

Services

▶ Email this article to a friend
▶ Similar articles in this journal
▶ Similar articles in PubMed
▶ Alert me to new issues of the journal
▶ Download to citation manager
▶ Request Permissions

Google Scholar

▶ Articles by Li, G.-Y.
▶ Articles by Ikura, M.

PubMed

▶ PubMed Citation
▶ Articles by Li, G.-Y.
▶ Articles by Ikura, M.

Social Bookmarking

What's this?

Guang-Yao Li[‡], Yonglong Zhang[§], Masayori Inouye[§], and Mitsuhiko Ikura[‡1]

From the [‡] Division of Signaling Biology, Ontario Cancer Institute, and Department of Medical Biophysics, University of Toronto, Toronto, Ontario M5G 1L7, Canada and the [§] Department of Biochemistry, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854-5635

In *Escherichia coli*, RelE toxin participates in growth arrest and cell death by inducing mRNA degradation at the ribosomal A-site under stress conditions. The NMR structures of a mutant of *E. coli* RelE toxin, RelE$^{R81A/R83A}$, with reduced toxicity and its complex with an inhibitory peptide from RelB antitoxin, RelB$_C$ (Lys^{47}-Leu^{79}), have been determined. In the free RelE$^{R81A/R83A}$ structure, helix α4 at the C terminus adopts a closed conformation contacting with the β-sheet core and adjacent loops. In the RelE$^{R81A/R83A}$-RelB$_C$ complex, helix α3[*] of RelB$_C$ displaces α4 of RelE$^{R81A/R83A}$ from the binding site on the β-sheet core. This helix replacement results in neutralization of a conserved positively charged cluster of RelE by acidic residues from α3[*] of RelB. The released helix α4 becomes unfolded, adopting an open conformation with increased mobility. The displacement of α4 disrupts the geometry of critical residues, including Arg^{81} and Tyr^{87}, in a putative active site of RelE toxin. Our structures indicate that RelB counteracts the toxic activity of RelE by displacing α4 helix from the catalytically competent position found in the free RelE structure.

SlipChip‡

Wenbin Du , Liang Li , Kevin P. Nichols and Rustem F. Ismagilov *
*Department of Chemistry and Institute for Biophysical Dynamics, The University of
Chicago, 929 East 57th Street, Chicago, Illinois 60637, USA. E-mail: r-
ismagilov@uchicago.edu*

The SlipChip is a microfluidic device designed to perform multiplexed microfluidic reactions
without pumps or valves. The device has two plates in close contact. The bottom plate contains
wells preloaded with many reagents; in this paper plates with 48 reagents were used. These wells
are covered by the top plate that acts as a lid for the wells with reagents. The device also has a
fluidic path, composed of ducts in the bottom plate and wells in the top plate, which is connected
only when the top and bottom plate are aligned in a specific configuration. Sample can be added
into the fluidic path, filling both wells and ducts. Then, the top plate is "slipped", or moved,
relative to the bottom plate so the complementary patterns of wells in both plates overlap,
exposing the sample-containing wells of the top plate to the reagent-containing wells of the
bottom plate, and enabling diffusion and reactions. Between the two plates, a lubricating layer of
fluorocarbon was used to facilitate relative motion of the plates. This paper implements this
approach on a nanoliter scale using devices fabricated in glass. Stability of preloaded solutions,
control of loading, and lack of cross-contamination were tested using fluorescent dyes.
Functionality of the device was illustrated *via* crystallization of a model membrane protein.
Fabrication of this device is simple and does not require a bonding step. This device requires no
pumps or valves and is applicable to resource-poor settings. Overall, this device should be
valuable for multiplexed applications that require exposing one sample to many reagents in small
volumes. One may think of the SlipChip as an easy-to-use analogue of a preloaded multi-well
plate, or a preloaded liquid-phase microarray.

Introduction

This paper describes the SlipChip, a microfluidic device for carrying out multiplexed solution-
phase experiments in a simple format. Multiplexed experiments are common, especially in
applications that require screening. Miniaturization and simplification of these experiments are
attractive to minimize both sample volume and labor costs and to improve efficiency. Various
microfluidic systems have been developed to perform multiplexed experiments on the nanoliter
scale.[1-18] Sophisticated approaches have used valve-based systems[4] to route the sample to
many reaction chambers that can in turn be loaded with many different reagents. These systems
required external equipment for hydraulic control and used gas-permeable materials for dead-end
filling, which made storing reagents on-board difficult.[19,20] Compact-disc (CD)-based

Microwave dielectric heating of drops in microfluidic devices

David Issadore,[a] Katherine J. Humphry,[b] Keith A. Brown,[a] Lori Sandberg,[c] David A. Weitz[ab] and Robert M. Westervelt[*ab]

We present a technique to locally and rapidly heat water drops in microfluidic devices with microwave dielectric heating. Water absorbs microwave power more efficiently than polymers, glass, and oils due to its permanent molecular dipole moment that has large dielectric loss at GHz frequencies. The relevant heat capacity of the system is a single thermally isolated picolitre-scale drop of water, enabling very fast thermal cycling. We demonstrate microwave dielectric heating in a microfluidic device that integrates a flow-focusing drop maker, drop splitters, and metal electrodes to locally deliver microwave power from an inexpensive, commercially available 3.0 GHz source and amplifier. The temperature change of the drops is measured by observing the temperature dependent fluorescence intensity of cadmium selenide nanocrystals suspended in the water drops. We demonstrate characteristic heating times as short as 15 ms to steady-state temperature changes as large as 30 °C above the base temperature of the microfluidic device. Many common biological and chemical applications require rapid and local control of temperature and can benefit from this new technique.

Introduction

The miniaturization of the handling of liquid and biological samples has enabled great advances in fields such as drug discovery, genetic sequencing and synthesis, cell sorting, single cell gene expression studies, and low-cost portable medicine.[1-7] At the forefront of this technology are the micro-fabricated pipes, valves, pumps, and mixers of microfluidics that are leading to integrated lab-on-a-chip devices. These integrated microfluidic devices are causing a paradigm shift in fluid handling that is analogous to what integrated circuit technology did for electronics half of a century ago.[6] A growing library of elements for lab-on-a-chip systems have been developed in recent years for tasks such as the mixing of reagents, detecting and counting of cells, sorting cells, genetic analysis, and protein detection.[1-7] There is one function, however, that is crucial to many applications and which has remained a challenge: the local control of temperature. The large surface area to volume ratios found in micrometre-scale channels and the close proximity of microfluidic elements make temperature control in such systems a unique challenge.[8,9]

Much work has been done in the last decade to improve local temperature control in microfluidic systems. The most common technique uses Joule heated metal wires and thin films to conduct heat into fluid channels.[10-13] The thermal conductivity of microfluidic devices control the localization of the temperature change and tends to be in the order of centimetres.[9,11] Temporal control is limited by the heat capacity and thermal coupling of the microfluidic device to the environment and thermal relaxation times tend to be in the order of seconds.[9,11] Alternative techniques to improve the localization and response time have been developed, such as those that use non-contact infrared heating of water in glass microfluidic systems[14] and integrated micrometre size Peltier Junctions to transfer heat between two channels containing fluid at different temperatures.[15] Fluids have also been cooled on millisecond time scales with evaporative cooling by pumping gasses into the fluid channels.[16]

The focus of our research is to integrate electronics with microfluidics to bring new capabilities to lab-on-a-chip systems.[6,7] In this paper we present a technique to locally and rapidly heat water in drop based microfluidic systems with microwave dielectric heating. The devices are fabricated using soft lithography and are connected to inexpensive commercially available microwave electronics. This work builds on previous work in which microwaves have been used to heat liquid in microfluidic devices[17-20] by achieving significantly faster thermal response times and a greater temperature range. In our device, drops of water are thermally isolated from the bulk of the device and this allows exceptionally fast heating and cooling times $\tau_s = 15$ ms to be attained and the drops' temperature to be increased by 30 °C. The coupling of microwave electronics with microfluidics technology offers an inexpensive and easily integrated technique to locally and rapidly control temperature.

Theory of dielectric heating

Dielectric heating describes the phenomenon by which a material is heated with a time-varying electric field. Induced and intrinsic dipole moments within an object will align themselves with a time-varying electric field. The energy associated with this alignment is viscously dissipated as heat into the surrounding solution. The power density P absorbed by a dielectric material is given by the frequency ω of the applied electric field, the loss factor ε'' of the material, the vacuum permittivity ε_0, and the electric field strength $|E|$ with the expression:[21]

[a]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138, USA. E-mail: westervelt@seas.harvard.edu
[b]Department of Physics, Harvard University, Cambridge, MA, 02138, USA
[c]College of Engineering and Applied Sciences, University of Wyoming, Laramie, WY, 82071, USA

Microfluidic valves with integrated structured elastomeric membranes for reversible fluidic entrapment and *in situ* channel functionalization

Siva A. Vanapalli,*[acd] Daniel Wijnperle,[bc] Albert van den Berg,[bc] Frieder Mugele[ac] and Michel H. G. Duits*[ac]

We report the utility of structured elastomeric membranes (SEMs) as a multifunctional microfluidic tool. These structured membranes are part of a two-layer microfluidic device, analogous to membrane valves, with the novelty that they incorporate topographical features on the roof of the fluid channel. We demonstrate that when the topographical features are recessed into the roof of the fluid channel, actuation of the membrane leads to effective confinement of fluids down to femtolitres in preformed microfluidic containers. Thus, the SEMs in this case function as fluidic traps that could be coupled to microfluidic networks for rapid and repeated flushing of solvents. Alternatively, when the topographical features on the roof protrude into the fluid channel, we demonstrate that the SEMs can be used to pattern proteins and cells in microchannels. Thus in this case, the SEMs serve as fluidic stamps for functionalizing microchannel surfaces. In addition, we show that the trap or pattern shape and size can be manipulated simply by varying the topography on the elastomeric membrane. Since SEMs, membrane valves and pumps use similar fabrication technology, we believe that SEMs can be integrated into microfluidic large-scale circuits for biotechnological applications.

1. Introduction

Soft lithography tools such as microcontact printing, microfluidic networks and membrane valves and pumps have emerged as powerful methods for addressing the current needs in biotechnology.[1,2] Yet, these tools suffer from some limitations or need further improvement, so that they can be generically applicable to the broad range and complexity associated with biotechnological applications. For example, microcontact printing provides excellent control over material pattern shape and size on open substrates.[3] However, such patterns are not easily integrable into microfluidic devices, thereby precluding the capability to perform combinatorial assays (*e.g.* screening several biomarkers for immunoassays). Alternatively, microfluidic networks coupled with laminar flows allow patterning on microchannel surfaces,[4,5] but lack control over pattern shape and size and therefore are not amenable to fabrication of high-density bioarrays. Here we present a soft lithography-based microfluidic tool called structured elastomeric membranes (SEMs) that overcomes some of these limitations. In particular we incorporate positive relief features (that are common to microcontact printing stamps) on the elastomeric deformable membrane of microfluidic valves[6,7] and show their application as fluidic stamps. Using this approach we generate patterned areas of proteins and cells of controllable shape and size on microchannel

surfaces. We also show that by incorporating negative relief features on the membrane, we can effectively confine femtolitre volumes in preformed microfluidic containers of controllable shape and size. Ultimately we believe that SEMs are a valuable addition to the toolbox needed for building integrated microfluidic circuits capable of high throughput bioanalysis since they can be readily integrated with membrane valves and pumps.

2. Experimental

2.1 Materials

Surface blocking agents bovine serum albumin (BSA, 4 mg mL^{-1}) and TRITC labeled BSA (1 mg mL^{-1}) were procured from Sigma-Aldrich. Fibronectin (100 μg mL^{-1}, obtained from a local blood bank) was used as a representative protein to demonstrate patterning of proteins by SEMs. It was detected by using immunofluorescence staining using fibronectin antibody (0.05 mg mL^{-1}) and Alexa 488 labeled Anti-rabbit IgG (0.1 mg mL^{-1}), both of which were purchased from Abcam. FITC-phalloidin and DAPI were used to stain the intracellular actin network and the nucleus respectively, both of which were acquired from Molecular Probes. Phosphate buffer saline (PBS) solution was obtained from the local hospital in Enschede.

Human umbilical vein endothelial cells were isolated from umbilical cords by the method of Jaffe *et al.*,[8] using a PBS solution containing trypsin (0.05% (w/v) and 0.02% (w/v) EDTA. The obtained endothelial cells were cultured in fibronectin-coated culture flasks in Endothelial Growth Medium-2 (Lonza), containing 2% fetal bovine serum, until they reached confluence. When confluent, cells were detached from the surface using trypsin solution and diluted 1 : 3 in a fresh fibronectin-coated culture flask. Cells were kept in a humidified incubator (37 °C, 5% CO$_2$) up to passage 8, after which they were discarded. Cells

[a]Physics of Complex Fluids, University of Twente, 7500, AE, Enschede, The Netherlands. E-mail: siva.vanapalli@ttu.edu; m.h.g.duits@tnw.utwente.nl
[b]BIOS Lab on a Chip, University of Twente, 7500, AE, Enschede, The Netherlands
[c]MESA+ Institute of Nanotechnology, University of Twente, 7500, AE, Enschede, The Netherlands
[d]Department of Chemical Engineering, Texas Tech University, Lubbock, TX, 79409, USA

Mutational analysis of *Escherichia coli* σ²⁸ and its target promoters reveals recognition of a composite −10 region, comprised of an 'extended −10' motif and a core −10 element

Byoung-Mo Koo,[1] Virgil A. Rhodius,[1] Elizabeth A. Campbell[2] and Carol A. Gross[1,3]*
Departments of [1]Microbiology and Immunology, and [3]Cell and Tissue biology, University of California at San Francisco, San Francisco, CA 94158, USA.
[2]Laboratory of Molecular Biophysics, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA.

Summary

σ²⁸ controls the expression of flagella-related genes and is the most widely distributed alternative σ factor, present in motile Gram-positive and Gram-negative bacteria. The distinguishing feature of σ²⁸ promoters is a long −10 region (GCCGATAA). Despite the fact that the upstream GC is highly conserved, previous studies have not indicated a functional role for this motif. Here we examine the functional relevance of the GCCG motif and determine which residues in σ²⁸ participate in its recognition. We find that the GCCG motif is a functionally important composite element. The upstream GC constitutes an extended −10 motif and is recognized by R91, a residue in Domain 3 of σ²⁸. The downstream CG is the upstream edge of −10 region of the promoter; two residues in Region 2.4, D81 and R84, participate in its recognition. Consistent with their role in base-specific recognition of the promoter, R91, D81 and D84 are universally conserved in σ²⁸ orthologues. σ²⁸ is the second Group 3 σ shown to use an extended −10 region in promoter recognition, raising the possibility that other Group 3 σs will do so as well.

Introduction

Bacteria use a family of σ factors to orchestrate transcription. The housekeeping σ, called σ⁷⁰ in *Escherichia coli*, directs core RNA polymerase (α₂ββ'ω) to the vast bulk of promoters active during exponential phase,

whereas alternative σs direct core RNA polymerase to mediate the transcription of regulons required for specific tasks such as response to stress, or mediating growth transitions or development (Gross *et al.*, 1998; Paget and Helmann, 2003). The σ⁷⁰ family of proteins has conserved modular domains with discrete functions. In addition to the N-terminal regulatory domain found only in housekeeping σs, there are three additional domains, each with recognition determinants both for core RNA polymerase and for portions of the promoter (Gruber and Gross, 2003). Thus, Domain 2 recognizes the −10 region of the promoter; Domain 3 recognizes the 'extended −10 motif' immediately upstream of the −10 region; and Domain 4 recognizes the −35 region of the promoter (Gross *et al.*, 1998; Campbell *et al.*, 2002). In addition, Region 1.2 recognizes a motif downstream of the −10 region (Feklistov *et al.*, 2006; Haugen *et al.*, 2006). σs have been divided into four groups based on their phylogenetic relationships and modular structure (Lonetto *et al.*, 1992; Gruber and Gross, 2003; Paget and Helmann, 2003). Group 1 σs (housekeeping σs) are essential and have all domains; Group 2 σs are the most related to Group 1 σs but are not essential. Group 3 σs have Domains 2–4. Group 4 σs have only Domains 2 and 4 and comprise the largest group of σs.

Promoter recognition has been extensively studied in σ⁷⁰ and several other Group 1 σs (Kenney *et al.*, 1989; Siegele *et al.*, 1989; Waldburger *et al.*, 1990; Campbell *et al.*, 2002; Sanderson *et al.*, 2003). However, less attention has been devoted to promoter recognition in the alternative σs. σ²⁸, a Group 3 σ, is the most widely distributed alternative σ factor, making it an attractive candidate for study of its promoter recognition. σ²⁸ controls expression of flagella-related genes in all motile Gram-negative and Gram-positive bacteria, and plays a role in development in some non-motile bacteria (e.g. Chlamydia) (Chilcott and Hughes, 2000; Yu and Tan, 2003; Serizawa *et al.*, 2004; Shen *et al.*, 2006; Yu *et al.*, 2006b). The conservation of this σ across millions of years of evolution is indicated by the fact that SigD, the σ²⁸ orthologue in *Bacillus subtilis*, can substitute for the function of σ²⁸ in *E. coli* (Chen and Helmann, 1992). Likewise, the σ²⁸ promoter sequence is conserved across organisms

Dissection of recognition determinants of *Escherichia coli* σ³² suggests a composite −10 region with an 'extended −10' motif and a core −10 element

Byoung-Mo Koo,[1] Virgil A. Rhodius,[1]
Elizabeth A. Campbell[2] and Carol A. Gross[1,3]*
Departments of [1]Microbiology and Immunology, and
[3]Cell and Tissue biology, University of California at
San Francisco, San Francisco, CA 94158, USA.
[2]Laboratory of Molecular Biophysics, The Rockefeller
University, 1230 York Avenue, New York, NY 10065,
USA.

Summary

σ³² controls expression of heat shock genes in
Escherichia coli and is widely distributed in
proteobacteria. The distinguishing feature of σ³² pro-
moters is a long −10 region (CCCCATNT) whose
tetra-C motif is important for promoter activity. Using
alanine-scanning mutagenesis of σ³² and *in vivo* and
in vitro assays, we identified promoter recognition
determinants of this motif. The most downstream C
(−13) is part of the −10 motif; our work confirms and
extends recognition determinants of −13C. Most
importantly, our work suggests that the two upstream
Cs (−16, −15) constitute an 'extended −10' recognition
motif that is recognized by K130, a residue universally
conserved in β- and γ-proteobacteria. This residue is
located in the α-helix of σDomain 3 that mediates
recognition of the extended −10 promoter motif in
other σs. K130 is not conserved in α- and δ-/ε-
proteobacteria and we found that σ³² from the
α-proteobacterium *Caulobacter crescentus* does not
need the extended −10 motif for high promoter
activity. This result supports the idea that K130 medi-
ates extended −10 recognition. σ³² is the first Group 3
σ shown to use the 'extended −10' recognition motif.

Introduction

Bacterial transcription initiation is mediated by RNA poly-
merase holoenzyme, consisting of a 5-subunit core RNA
polymerase (α₂ββ'ω) and a dissociable subunit, σ. σs
contain many of the promoter recognition determinants

(Gross *et al.*, 1998). Most transcription in exponentially
growing cells is initiated by RNA polymerase holoenzyme
carrying the housekeeping σ, called σ⁷⁰ in *Escherichia
coli*. Alternative σs mediate transcription of regulons acti-
vated by specific stress conditions, or by growth or devel-
opmental transitions. The σ⁷⁰ family of proteins has been
divided into four groups, based on their phylogenetic rela-
tionships and modular structure (Lonetto *et al.*, 1992;
Gruber and Gross, 2003). Group 1 σs are the essential
housekeeping σs and contain four domains, of which the
first is unique to the Group 1 σs. Group 2 σs are most
closely related to Group 1 σs, but are not essential. Most
Group 3 σs contain Domains 2–4, and the Group 4 σs
contain only Domains 2 and 4 (Gruber and Gross, 2003).

Promoters recognized by σ⁷⁰ and other Group 1 σs
have been extensively studied. Promoters are multipar-
tite, and can have the following potential elements to
mediate recognition: (i) the 'UP element', generally
located from −61 to −41 relative to the transcription start
point (Ross *et al.*, 1993; Gourse *et al.*, 2000), (ii) the −35
region, a hexamer (TTGACA) centred about 35 bp
upstream of the transcription start point (Gross *et al.*,
1998), (iii) the 'extended −10 motif' (TG) located immedi-
ately upstream of the −10 region (Mitchell *et al.*, 2003), (iv)
the −10 region, a hexamer (TATAAT) centred about 10 bp
upstream of the transcription start (Helmann and
deHaseth, 1999) and (v) the discriminator region (GGGa)
located downstream of the −10 region of the promoter
(Feklistov *et al.*, 2006; Haugen *et al.*, 2006). With the
exception of the UP element, recognized by the
C-terminal domain of α, promoter elements are recog-
nized by the σ subunit. The −35 element is recognized by
Region 4.2 in σDomain 4; the extended −10 element by an
α-helix in σDomain 3; the −10 by Region 2.4 in σDomain
2 and the discriminator by σRegion 1.2 (Campbell *et al.*,
2002; Feklistov *et al.*, 2006; Haugen *et al.*, 2006). The
number of elements in promoters is variable.

Promoters recognized by the housekeeping σ in *E. coli*
can be roughly divided into those with an extended −10
motif ('extended −10 promoters') and those without this
motif. Extended −10 promoters have a lower match to the
consensus −35 motif than do those lacking the extended
−10 motif (Mitchell *et al.*, 2003). Moreover, at least *in vitro*,

LETTERS

Irreversibility of mitotic exit is the consequence of systems-level feedback

Sandra López-Avilés[1], Orsolya Kapuy[2,3], Béla Novák[2] & Frank Uhlmann[1]

The eukaryotic cell cycle comprises an ordered series of events, orchestrated by the activity of cyclin-dependent kinases (Cdks), leading from chromosome replication during S phase to their segregation in mitosis. The unidirectionality of cell-cycle transitions is fundamental for the successful completion of this cycle. It is thought that irrevocable proteolytic degradation of key cell-cycle regulators makes cell-cycle transitions irreversible, thereby enforcing directionality[1–3]. Here we have experimentally examined the contribution of cyclin proteolysis to the irreversibility of mitotic exit, the transition from high mitotic Cdk activity back to low activity in G1. We show that forced cyclin destruction in mitotic budding yeast cells efficiently drives mitotic exit events. However, these remain reversible after termination of cyclin proteolysis, with recovery of the mitotic state and cyclin levels. Mitotic exit becomes irreversible only after longer periods of cyclin degradation, owing to activation of a double-negative feedback loop involving the Cdk inhibitor Sic1 (refs 4, 5). Quantitative modelling suggests that feedback is required to maintain low Cdk activity and to prevent cyclin resynthesis. Our findings demonstrate that the unidirectionality of mitotic exit is not the consequence of proteolysis but of systems-level feedback required to maintain the cell cycle in a new stable state.

After completion of chromosome segregation during mitosis, the activity of the key cell-cycle kinase Cdk is downregulated to promote mitotic exit and the return of cells to G1. This involves ubiquitin-mediated degradation of mitotic cyclins under control of the anaphase promoting complex (APC), a multisubunit ubiquitin ligase. Mitotic cyclins are initially targeted for degradation by the APC in association with its activating subunit Cdc20 (APC$^{\text{Cdc20}}$). Later, declining Cdk levels and activation of the Cdk-counteracting phosphatase Cdc14 allow a second APC activator, Cdh1, to associate with the APC (APC$^{\text{Cdh1}}$)[6–8]. Cyclin proteolysis, a thermodynamically irreversible reaction, is thought to be responsible for the irreversibility of mitotic exit[1–3]. However, de novo protein synthesis can counteract degradation and constitutes a similar thermodynamically irreversible process, driven by ATP hydrolysis. In a cellular setting, therefore, protein levels are defined by reversible changes to the rates of two individually irreversible reactions: protein synthesis and degradation. These considerations have led to the hypothesis that it is not proteolysis itself, but systems-level feedback that affects synthesis and degradation rates, making cell-cycle transitions irreversible[9].

To test this hypothesis, we investigated the contribution of cyclin proteolysis to the irreversibility of budding yeast mitotic exit (Fig. 1a). We arrested budding yeast cells in mitosis with high levels of mitotic cyclins by depleting Cdc20 under control of the MET3 promoter. In these cells, we induced Cdh1 expression from the galactose-inducible GALL promoter (an attenuated version of the GAL1 promoter) for 30 min[6]. We expressed a Cdh1 variant, Cdh1(m11), that activates the

APC even in the presence of high Cdk activity owing to mutation of 11 Cdk phosphorylation sites. This led to efficient degradation of the major budding yeast mitotic cyclin Clb2 (Fig. 1b). The mitotic Polo-like kinase, another APC$^{\text{Cdh1}}$ target[10], was also efficiently degraded, whereas levels of the S phase cyclin Clb5, a preferential substrate for APC$^{\text{Cdc20}}$ (ref. 11), remained largely unaffected (Supplementary Fig. 1). Clb2 destruction was accompanied by dephosphorylation of known mitotic Cdk substrates, seen by their change in electrophoretic mobility (Fig. 1b). These included three proteins whose dephosphorylation contributes to spindle elongation and chromosome segregation, Sli15, Ase1 and Ask1 (refs 12–14). Their dephosphorylation depended on the activity of the mitotic exit phosphatase Cdc14 (Supplementary Fig. 2). Mitotic spindles that were present in the metaphase-arrested cells disassembled as Clb2 levels declined, accompanied by outgrowth of pronounced astral microtubules (Fig. 1c and Supplementary Fig. 3), reminiscent of spindle breakdown at the end of mitosis. APC$^{\text{Cdh1(m11)}}$-mediated destruction of the spindle stabilizing factor Ase1 (Fig. 1b), in addition to its dephosphorylation, may contribute to this phenotype[15].

After 50 min, when Clb2 levels became almost undetectable, we turned off APC$^{\text{Cdh1(m11)}}$ by inactivating a temperature-sensitive APC core subunit encoded by the cdc16-123 allele[16]. Notably, in response to APC$^{\text{cdc16-123}}$ inactivation, Clb2 levels recovered and Sli15, Ase1 and Ask1 reappeared in their mitotic hyperphosphorylated forms. Mitotic spindles formed again, suggesting that cells had re-entered a mitotic state. Spindles appeared morphologically intact, but were longer after cyclin reaccumulation (3.9 ± 0.8 µm; mean ± s.d.) compared to metaphase spindles at the beginning of the experiment (2.1 ± 0.6 µm). A probable reason for this lies in compromised sister chromatid cohesion after some, albeit inefficient, inactivation of the anaphase inhibitor securin by APC$^{\text{Cdh1(m11)}}$ (Supplementary Fig. 4)[17]. Fluorescence-activated cell sorting (FACS) analysis of DNA content confirmed that cells maintained a 2C DNA content throughout the experiment (Fig. 1d). This demonstrates that cyclin destruction promotes mitotic exit events, but it is not sufficient to render them irreversible. Cyclin resynthesis can reverse mitotic exit. Note that reversibility of mitotic exit under these conditions did not depend on the persistence of the S phase cyclin Clb5 (Supplementary Fig. 1).

We next addressed what makes mitotic exit irreversible, if not cyclin destruction. When we repeated the experiment, but continued Cdh1(m11) induction and inactivated APC$^{\text{cdc16-123}}$ only after 60 min, Clb2 did not reaccumulate and over half of the cells subsequently completed cytokinesis and entered G1 (Fig. 2a). This suggests that after longer periods of Clb2 destruction mitotic exit becomes irreversible. Western blotting showed that around the time when Clb2 loss turned irreversible, the Cdk inhibitor Sic1 accumulated[4,5]. We therefore asked whether Sic1 accumulation was responsible for the irreversibility of mitotic exit. When we repeated the experiment using a sic1Δ strain, Clb2 reappeared after APC$^{\text{cdc16-123}}$ inactivation, and only a minority of

LETTERS

Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis

Sabrina L. Spencer[1,2]*, Suzanne Gaudet[1]†*, John G. Albeck[1], John M. Burke[1] & Peter K. Sorger[1]

In microorganisms, noise in gene expression gives rise to cell-to-cell variability in protein concentrations[1-7]. In mammalian cells, protein levels also vary[8-10] and individual cells differ widely in their responsiveness to uniform physiological stimuli[11-15]. In the case of apoptosis mediated by TRAIL (tumour necrosis factor (TNF)-related apoptosis-inducing ligand) it is common for some cells in a clonal population to die while others survive—a striking divergence in cell fate. Among cells that die, the time between TRAIL exposure and caspase activation is highly variable. Here we image sister cells expressing reporters of caspase activation and mitochondrial outer membrane permeabilization after exposure to TRAIL. We show that naturally occurring differences in the levels or states of proteins regulating receptor-mediated apoptosis are the primary causes of cell-to-cell variability in the timing and probability of death in human cell lines. Protein state is transmitted from mother to daughter, giving rise to transient heritability in fate, but protein synthesis promotes rapid divergence so that sister cells soon become no more similar to each other than pairs of cells chosen at random. Our results have implications for understanding 'fractional killing' of tumour cells after exposure to chemotherapy, and for variability in mammalian signal transduction in general.

TRAIL elicits a heterogeneous phenotypic response in both sensitive and relatively resistant cell lines: some cells die within 45 min, others 8–12 h later, and yet others live indefinitely (Supplementary Fig. 1). During the variable delay between TRAIL addition and mitochondrial outer membrane permeabilization (MOMP), upstream initiator caspases are active but downstream effector caspases are not[11,12]. Possible sources of cell-to-cell variability in response to TRAIL include genetic or epigenetic differences, stochastic fluctuations in biochemical reactions involving low copy number components ('intrinsic noise'[3]), differences in cell cycle phase, and natural variation in the concentrations of important reactants. To distinguish between these and other possibilities, we used live-cell microscopy to compare the timing and probability of death in sister cells exposed to TRAIL. If phenotypic variability is caused by genetic or epigenetic differences, sister cells should behave identically. In contrast, if stochastic fluctuations in reactions triggered by TRAIL predominate, sister cells should be no more similar to each other than pairs of cells selected at random. The influence of cell cycle state on apoptosis should be readily observable from time-lapse imaging of asynchronous cultures. Furthermore, variability arising from differences in protein levels (or in their activity or modification state) should produce a highly distinctive form of inheritance in which newly born sister cells are very similar, because they inherit similar numbers of abundant factors from their mother[4,7], but then diverge as new proteins are made and levels drift[10,16]. With this in mind, we examined apoptosis in HeLa cells and in non-transformed

MCF10A mammary epithelial cells in the presence and absence of protein synthesis inhibitors.

Pairs of sister cells expressing a fluorescent reporter of MOMP (mitochondrial intermembrane space reporter protein, IMS-RP[11]) born during a 20–30 h period were identified by time-lapse microscopy. TRAIL and the protein synthesis inhibitor cycloheximide were then added and imaging continued for another 8 h. The TRAIL-to-MOMP interval (T_d) was calculated for each cell (Fig. 1a). Among the recently divided sisters (<7 h between division and death), T_d was highly correlated ($R^2 = 0.93$, Fig. 1b), whereas T_d was uncorrelated ($R^2 = 0.04$) for recently divided cells chosen at random. The time since division (Fig. 1c) and the position in the dish (data not shown) did not correlate with T_d, ruling out a role for cell cycle state and cell–cell interactions under our experimental conditions. However, as the time since division increased, sister-to-sister correlation in T_d decayed exponentially with a half-life of ~11 h, so that sisters lost memory of shared ancestry within ~50 h or about two cell generations ($R^2 \leq 0.05$, the same as random pairs of cells; Fig. 1d, e). Similar results were obtained with MCF10A cells (Supplementary Fig. 2).

High correlation between recently born sisters shows that the variability in T_d arises from differences that exist before TRAIL exposure, and rules out stochastic fluctuations in signalling reactions. Rapid decorrelation also rules out genetic mutation or conventional epigenetic differences (which typically last $10-10^5$ cell divisions[17]). However, transient heritability is precisely what we expected for cell-to-cell differences arising from variations in the concentrations or states of proteins that are partitioned binomially at cell division.

Whereas all TRAIL-treated HeLa cells eventually died in the presence of cycloheximide, in its absence a fraction always survived (presumably owing to induction of survival pathways[18]). When the fates of sister cells were compared, both lived or both died in most cases (chi-squared test, $P = 7 \times 10^{-19}$, Supplementary Fig. 3). Variability in T_d across the population was large (Fig. 1g and Supplementary Fig. 4), but recently born sisters were nevertheless correlated in T_d ($R^2 = 0.75$, Fig. 1f). Again, cell cycle phase was not correlated with fate or the time-to-death (Fig. 1g). Decorrelation in T_d among sisters was an order of magnitude more rapid in the presence of protein synthesis than in its absence (~1.5 h half-life, Fig. 1h, i and Supplementary Fig. 5). Thus, the length of time that T_d is heritable is very sensitive to rates of protein synthesis, both basal and TRAIL-induced.

We then asked whether the concentrations of proteins regulating TRAIL-induced apoptosis are sufficiently different from cell to cell to account for the variability in T_d. Using flow cytometry, we measured the distributions of five apoptotic regulators for which specific antibodies are available. All five proteins were log-normally distributed across the population with coefficients of variation ranging between 0.21 and 0.28 for cells of similar size (Fig. 2a), consistent with data on

[1]Center for Cell Decision Processes, Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. †Present address: Department of Cancer Biology, Dana-Farber Cancer Institute, and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.
*These authors contributed equally to this work.

428

LETTERS

Snowdrift game dynamics and facultative cheating in yeast

Jeff Gore[1], Hyun Youk[1] & Alexander van Oudenaarden[1]

The origin of cooperation is a central challenge to our understanding of evolution[1–3]. The fact that microbial interactions can be manipulated in ways that animal interactions cannot has led to a growing interest in microbial models of cooperation[4–10] and competition[11,12]. For the budding yeast *Saccharomyces cerevisiae* to grow on sucrose, the disaccharide must first be hydrolysed by the enzyme invertase[13,14]. This hydrolysis reaction is performed outside the cytoplasm in the periplasmic space between the plasma membrane and the cell wall. Here we demonstrate that the vast majority (~99 per cent) of the monosaccharides created by sucrose hydrolysis diffuse away before they can be imported into the cell, serving to make invertase production and secretion a cooperative behaviour[15,16]. A mutant cheater strain that does not produce invertase is able to take advantage of and invade a population of wild-type cooperator cells. However, over a wide range of conditions, the wild-type cooperator can also invade a population of cheater cells. Therefore, we observe steady-state coexistence between the two strains in well-mixed culture resulting from the fact that rare strategies outperform common strategies—the defining features of what game theorists call the snowdrift game[17]. A model of the cooperative interaction incorporating nonlinear benefits explains the origin of this coexistence. We are able to alter the outcome of the competition by varying either the cost of cooperation or the glucose concentration in the media. Finally, we note that glucose repression of invertase expression in wild-type cells produces a strategy that is optimal for the snowdrift game—wild-type cells cooperate only when competing against cheater cells.

Yeast prefers to use the monosaccharides glucose and fructose as carbon sources. However, when these sugars are not available, yeast can metabolize alternative carbon sources such as the disaccharide sucrose[18]. After sucrose is hydrolysed by invertase, the resulting monosaccharides are imported[13,14], yet some of the glucose and fructose may diffuse away from the cell before it is able to import them into the cytoplasm (Supplementary Fig. 1). If such sugar loss by diffusion is significant then we might expect high-density cultures to grow more quickly than low-density cultures, because cells at high density benefit from their hydrolysis products and those of their abundant neighbours. Indeed, we find that cells grown in media supplemented with sucrose—but not glucose—grow much faster at high cell density than at low cell density. The growth rate at high cell density in 5% sucrose is similar to the growth rate at saturating (2%) glucose concentrations. However, the growth rate at low cell density is ~40% lower, equivalent to the growth rate in only 0.003% glucose (Supplementary Fig. 2). The fraction of invertase-created glucose that is captured can be estimated by dividing the rate of glucose uptake of cells growing in 0.003% glucose by the measured rate of invertase activity, yielding an estimated glucose capture efficiency of only ~1% (Supplementary Fig. 3). Analytic calculations of glucose diffusion suggest that this low capture efficiency is an expected consequence of diffusion and the known properties of the sugar importers (Supplementary Fig. 4).

Given that 99% of glucose created by a cell is lost to neighbouring cells, it may be possible for a 'cheater' strain to take advantage of the cooperators by not secreting invertase and instead simply consuming the glucose created by other cells[15]. If cooperative cells shared all of the glucose that they created (that is, if 100% of hydrolysed glucose and fructose diffused away from the hydrolysing cell), then both the cooperators and the cheaters would have the same access to sugar, yet only the cooperators would bear the metabolic cost of invertase production and secretion. In this case, the cheaters would always outgrow the cooperators, and the interaction would be what is called a prisoner's dilemma, in which cooperation is not sustainable in a well-mixed environment[1,17]. However, we found that yeast retains a small fraction of the glucose created by sucrose hydrolysis, which may be sufficient to allow cooperative strategies to survive.

To explore this problem, we performed a set of competition experiments between the wild-type strain ('cooperator') and a mutant strain lacking the invertase gene ('cheater' or 'defector'; see Supplementary Fig. 1). Consistent with there being a metabolic cost associated with invertase production, we find that in glucose-supplemented media, cooperators grow more slowly than cheaters only when invertase is being expressed (Supplementary Fig. 5)[15]. In addition, the cooperator strain in our experiments is a histidine auxotroph; therefore, limiting the histidine concentration in the media slows the growth of the cooperator relative to the cheater, allowing us to experimentally increase the 'cost of cooperation' (Supplementary Fig. 6). We can measure the relative abundance ('fractions') of the two strains in a mixed culture by flow cytometry because they express different fluorescent proteins (Supplementary Fig. 7).

We began by monitoring the change over time in the fractions of cooperators and cheaters co-cultured in sucrose media. Each co-culture started from a different initial fraction of cooperators, and each day we performed serial dilutions into fresh media and measured the cell density and relative abundance of the two strains. In cultures starting with a small fraction of cheaters, the cheaters increased in frequency, consistent with the cheaters 'taking advantage' of the cooperators (Fig. 1a). However, when the initial fraction of cooperators was low, we found that the frequency of cooperators increased, suggesting that in the steady state there will be coexistence between the two strains. Indeed, the equilibrium fraction is independent of the starting fraction but depends upon the histidine concentration (Fig. 1b; the equilibrium fraction in saturating histidine was $f \approx 0.3$). As the cost of cooperation increased, we observed a decrease in both the equilibrium fraction of cooperators and the mean growth rate of the culture at equilibrium (Fig. 1c). A large cost of cooperation therefore allows the cheaters to dominate the population but also results in a low growth rate of both strains. Coexistence was also observed in continuous culture, meaning that the 'seasonality'

[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

LETTERS

Serial time-encoded amplified imaging for real-time observation of fast dynamic phenomena

K. Goda[1]*, K. K. Tsia[1]* & B. Jalali[1]*

Ultrafast real-time optical imaging is an indispensable tool for studying dynamical events such as shock waves[1,2], chemical dynamics in living cells[3,4], neural activity[5,6], laser surgery[7–9] and microfluidics[10,11]. However, conventional CCDs (charge-coupled devices) and their complementary metal–oxide–semiconductor (CMOS) counterparts are incapable of capturing fast dynamical processes with high sensitivity and resolution. This is due in part to a technological limitation—it takes time to read out the data from sensor arrays. Also, there is the fundamental compromise between sensitivity and frame rate; at high frame rates, fewer photons are collected during each frame—a problem that affects nearly all optical imaging systems. Here we report an imaging method that overcomes these limitations and offers frame rates that are at least 1,000 times faster than those of conventional CCDs. Our technique maps a two-dimensional (2D) image into a serial time-domain data stream and simultaneously amplifies the image in the optical domain. We capture an entire 2D image using a single-pixel photodetector and achieve a net image amplification of 25 dB (a factor of 316). This overcomes the compromise between sensitivity and frame rate without resorting to cooling and high-intensity illumination. As a proof of concept, we perform continuous real-time imaging at a frame speed of 163 ns (a frame rate of 6.1 MHz) and a shutter speed of 440 ps. We also demonstrate real-time imaging of microfluidic flow and phase-explosion effects that occur during laser ablation.

Optical imaging is a widespread and versatile diagnostics and inspection tool in use today. Although most of the current research in imaging is aimed at improving the spatial resolution to below the diffraction limit[12,13], there are numerous applications that demand improvement in temporal resolution. Imaging systems with high temporal resolution are needed to study rapid physical phenomena ranging from shock waves, including extracorporeal shock waves used for surgery[1], to diagnostics of laser fusion[2] and fuel injection in internal combustion engines. High-speed imaging is becoming increasingly important in microscopy because on the micrometre scale even slow-moving phenomena require high temporal resolution. One example is the spatiotemporal study of biochemical waves in cells and tissues, which requires imaging with a micro- to nanosecond response time and is important for the study of cell signalling and drug transport[14]. Another key application is in the field of flow cytometry[11], where high-speed cameras are needed to provide high-throughput cell characterization.

The CCD or CMOS imager is by far the most widely deployed optical imaging technology. It offers a spatial resolution of a few micrometres, a large number of pixels and relatively low cost. Typical imagers used in consumer electronics have frame rates of 30 Hz, although high-end versions can operate at rates on the order of 1 kHz by reducing the number of pixels that are read out from the arrays. A CCD imager

with a 1-MHz frame rate—the world's fastest CCD camera—has been developed by Shimadzu Corporation (ref. 15; see also http://www.shimadzu.com/products/test/hsvc/oh80jt0000001d6t-att/booe13-02.pdf). Its impressive performance stems from several technological modifications aimed at overcoming the loss of sensitivity that results from the shorter integration time during high-speed imaging, and also at reducing the time it takes to read out the data from the 2D pixel arrays. These modifications include a larger pixel size for the purpose of adding pixel-level *in situ* memory (at the expense of reduced spatial resolution), and cooling of the camera to reduce noise (at the expense of the added complexity of refrigeration). The Shimadzu device also relies on a high-power illuminator to ensure adequate signal-to-noise ratio and to prevent a drop in sensitivity—a requirement that renders it unsuitable for microscopy, where focusing of the high-power illumination over a small field of view can cause damage to a biological sample. This compromise between sensitivity and frame rate is not unique to the CCD—it impacts almost all imaging and detection systems.

In scientific applications, high-speed imaging is often achieved using the time-resolved pump–probe technique[16–18]. Pump–probe techniques can capture the dynamics of fast events, but only if the event is repetitive. Because they do not operate in real time, they are unable to capture non-repetitive random events that occur only once or do not occur at regular intervals, such as rogue events[19]. Detection of such events requires an imaging technology with fast, continuous, real-time capability.

Another type of high-speed image sensor is the framing streak camera, which has been used for diagnostics in laser fusion, plasma radiation and combustion. This device operates in burst mode only (providing only several frames; see http://learn.hamamatsu.com/tutorials/streakcamera/) and requires synchronization of the camera with the event to be captured, rendering streak cameras also unable to capture unique or random events. This, along with the high cost of the camera, limits its use in practical applications.

Although CCD imagers will continue to be the most widely used imaging modality and pump–probe experiments will remain a powerful tool for studying fast repetitive events, a new and complementary imaging modality that can capture the dynamics of fast single-shot or random events is clearly needed. Serial time-encoded amplified microscopy (STEAM) technology is a new approach to imaging that provides such a capability. The 2D image is encoded into a serial time-domain waveform that is amplified and captured, not by a CCD camera, but instead by a single-pixel photodiode and an oscilloscope. The main attributes of the new imager are the image amplification in the optical domain and the elimination of the CCD—when combined, they enable continuous real-time operation at a frame rate of 6.1 MHz and a shutter speed of 440 ps. The STEAM camera operates continuously and can capture ultrafast events

[1]Department of Electrical Engineering, University of California, Los Angeles, California 90095, USA.
*These authors contributed equally to this work.

REVIEWS

Using movies to analyse gene circuit dynamics in single cells

James C. W. Locke and Michael B. Elowitz

Abstract | Many bacterial systems rely on dynamic genetic circuits to control crucial biological processes. A major goal of systems biology is to understand these behaviours in terms of individual genes and their interactions. However, traditional techniques based on population averages 'wash out' crucial dynamics that are either unsynchronized between cells or are driven by fluctuations, or 'noise', in cellular components. Recently, the combination of time-lapse microscopy, quantitative image analysis and fluorescent protein reporters has enabled direct observation of multiple cellular components over time in individual cells. In conjunction with mathematical modelling, these techniques are now providing powerful insights into genetic circuit behaviour in diverse microbial systems.

As biologists, we must grapple with, and reconcile, two very different views of cellular behaviour. On the one hand, we frequently think of cellular functions as being determined by 'circuits' of interacting genes and proteins. In a loosely analogous way to electronic circuits, these chemical circuits encode genetic programmes that underlie differentiation, the cell cycle and other behaviours (FIG. 1a). They accurately respond to stimuli and generate precise behavioural programmes in individual cells. On the other hand, there is the 'noisy' view of the cell we get when we actually look at cells: they exist in squishy, dynamic and heterogeneous populations, the morphologies, gene-expression patterns and differentiated states of which differ from one another, even when environment and genotype are fixed (FIG. 1b). How can precisely defined genetic circuits give rise to heterogeneity and, conversely, how does heterogeneity affect the behaviour of biological circuits?

Movies offer a powerful way to address these questions (FIG. 1c). By engineering microbial strains to express fluorescent protein reporters for key genes, researchers can follow the changing characteristics of individual cells over time. Quantitative detection methods, improved microscope automation and software, and the range of fluorescent reporter genes that are now available, in conjunction with mathematical modelling, can be combined to analyse gene circuit dynamics. Together, these techniques allow researchers to characterize epigenetic states, identify new dynamic phenomena, analyse biochemical interactions within circuits and elucidate the physiological function of genetic circuits, all at the single-cell level. Finally, movies provide an aesthetically compelling view of cellular

function that is often fascinating to watch. With movies, the eye often picks out subtle patterns in individual living cells that would be difficult to notice with less direct techniques. Few techniques are more fun.

How does quantitative movie analysis compare with alternative techniques for analysing gene circuits? Time-lapse microscopy follows a few genes over time in individual living cells. It complements approaches such as microarrays (which provide genome-scale expression data averaged over populations, but do not allow analysis of variability) and flow cytometry (which allows high-throughput acquisition of single-cell fluorescence values, but does not allow the same cell to be tracked over time). Movies also complement new single-cell quantitative PCR approaches, which enable analysis of expression of multiple genes in individual cells, but, because they require lysis of the cell, do not permit tracking of expression dynamics[1]. Movies enable researchers to determine the 'trajectories' of gene expression levels in individual living cells. One potential drawback of movies is that although many genes, and their expression levels, may be important for a particular process under study, most studies currently follow the dynamics of only a few genes at a time owing to the lack of distinguishable reporters. In the future, multi-spectral techniques may expand the number of simultaneous reporters[2]. However, studies that follow the dynamics of only two or three genes at a time can still be extremely informative.

Here we review work in which movies provide new insights into the dynamic behaviour of genetic components and circuits. For this Review, we have confined ourselves to microbial systems and have therefore

Howard Hughes Medical Institute, Division of Biology and Department of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA.
Correspondence to M.B.E.
e-mail: melowitz@caltech.edu

Unraveling infectious structures, strain variants and species barriers for the yeast prion [*PSI*⁺]

Peter M Tessier & Susan Lindquist

Prions are proteins that can access multiple conformations, at least one of which is β-sheet rich, infectious and self-perpetuating in nature. These infectious proteins show several remarkable biological activities, including the ability to form multiple infectious prion conformations, also known as strains or variants, encoding unique biological phenotypes, and to establish and overcome prion species (transmission) barriers. In this Perspective, we highlight recent studies of the yeast prion [*PSI*⁺], using various biochemical and structural methods, that have begun to illuminate the molecular mechanisms by which self-perpetuating prions encipher such biological activities. We also discuss several aspects of prion conformational change and structure that remain either unknown or controversial, and we propose approaches to accelerate the understanding of these enigmatic, infectious conformers.

The 'misfolding' and assembly of proteins into β-sheet–rich amyloid fibers is important in both disease[1] and normal biological function[2,3]. Although many proteins form amyloid fibers *in vitro* (see Table 1 for definitions), understanding the biological relevance and consequences of this process *in vivo* is difficult. Prions are one class of naturally occurring amyloid-forming proteins that have received much attention[3–10]. The first prion protein, PrP, was identified in mammals as an infectious agent responsible for several related neurodegenerative diseases, known collectively as the spongiform encephalopathies[8,10]. How a protein could be infectious was a complete mystery until the protein in question was identified as a normal constituent of the brain that simply changed its conformation from an α-helical to a β-sheet form to become infectious[8–10]. Once this alternative conformation appears in the brain—via contamination by infectious material, spontaneous conversion or mutation-induced misfolding—it is self-templating, converting more and more PrP to the infectious form and wrecking havoc in the brain as it does so[8–10].

Despite a wealth of evidence, it took many years for the 'protein-only' mechanism of prion transmission to be accepted. The discovery of a similar process operating in yeast cells, where it could be investigated more

readily owing to the ease of genetic manipulation, was an important factor in winning this battle[11–13]. The prions of yeast and other fungi consist of completely different proteins whose sequences are unrelated to their mammalian counterparts[3,4,6,11]. Moreover, fungal prions are generally not deleterious and can even be beneficial[3–7]. They serve as heritable elements, producing stable new phenotypes due to a profound change in protein conformation that is self-templating and transmissible from mother to daughter cells[3,4,6,11]. Indeed, the recent proposal of a prion-like mechanism for the perpetuation of synapses and neuronal memories[14], as well as the discovery of a host of new prions with diverse functions in yeast (for example, see refs. 15, 16), indicates that prions will prove vitally important in many organisms.

An important similarity between mammalian and yeast prions is that they form not just one prion conformation, but a collection of structurally related yet distinct conformations, known as prion strains[17–23]. For example, mice infected with prions from diverse animal origins manifested different patterns of disease, and these could be stably passed from mouse to mouse[24–28]. Although a seemingly obvious explanation was distinct viral strains, an explanation independent of nucleic acid emerged as evidence mounted that these different diseases traced to different (yet related) self-templating folds of the same protein, PrP[24–28]. Similarly, for yeast prions, unique protein folds produce a suite of distinct (yet related) prion phenotypes[17–19].

Another crucially important feature shared by mammalian and fungal prions is the species barrier[9,24,25,29–38]. The aforementioned prion strains show extremely low prion infectivity when introduced into mice; yet, once these mice succumbed to disease, mouse-to-mouse transmission was extremely efficient. Yeast prions also show strong species barriers that can be crossed, but with difficulty[29–32,34,35,39–41]. Remarkably, for both mammals and yeast, prion strains and species barriers are interrelated[4,8,9,24,26,27,29,37,40].

To decipher the complexities of these problems *in vivo*, it is necessary to analyze the biochemical properties of these proteins. Unfortunately, forming highly infectious mammalian prion conformers *in vitro* from recombinant protein has been difficult (for recent progress, see refs. 42,43). In contrast, bona fide highly infectious fungal prion conformers can be readily formed *in vitro*[18,19,44–46], allowing a more thorough characterization of their assembly process and amyloid structure, which will be reviewed here.

Known and potential fungal prions

The most well-studied fungal prion proteins are Sup35, Ure2, Rnq1 in *Saccharomyces cerevisiae* and HET-s in *Podospora anserina*[3–7]. Sup35 is a protein involved in recognition of stop codons during protein

Peter M. Tessier is at the Center of Biotechnology and Interdisciplinary Studies, Department of Chemical & Biological Engineering, Rensselaer Polytechnic Institute, Troy, New York, USA. Susan Lindquist is at the Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts, USA.
e-mail: tessier@rpi.edu or lindquist_admin@wi.mit.edu

Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover

Olivia S Rissland[1,2] & Chris J Norbury[1]

Both end structures of eukaryotic mRNAs, namely the 5' cap and 3' poly(A) tail, are necessary for transcript stability, and loss of either is sufficient to stimulate decay. mRNA turnover is classically thought to be initiated by deadenylation, as has been particularly well described in *Saccharomyces cerevisiae*. Here we describe two additional, parallel decay pathways in the fission yeast *Schizosaccharomyces pombe*. First, in fission yeast mRNA decapping is frequently independent of deadenylation. Second, Cid1-dependent uridylation of polyadenylated mRNAs, such as *act1*, *hcn1* and *urg1*, seems to stimulate decapping as part of a novel mRNA turnover pathway. Accordingly, *urg1* mRNA is stabilized in *cid1Δ* cells. Uridylation and deadenylation act redundantly to stimulate decapping, and our data suggest that uridylation-dependent decapping is mediated by the Lsm1–7 complex. As human cells contain Cid1 orthologs, uridylation may form the basis of a widespread, conserved mechanism of mRNA decay.

In budding yeast, most cytoplasmic mRNA turnover is initiated by deadenylation[1]. These messages are then either decapped and subject to 5'→3' decay[2,3] or degraded by the cytoplasmic exosome[4]. Although these decay pathways are conserved[5,6], metazoans and fission yeast contain additional cytoplasmic RNA-processing enzymes; the roles of many of these have not yet been determined. *Schizosaccharomyces pombe* Cid1 is one such enzyme, a cytoplasmic member of a family of RNA nucleotidyl transferases[7,8].

Cid1, identified through its involvement in the S-M checkpoint[7], is now known to be one of a subgroup of this family possessing either poly(U) polymerase (PUP) and/or terminal uridyl transferase (TUTase) activity[9]. This subgroup also includes the human enzymes U6 TUTase, Hs2 and Hs3 (refs. 9–11), although no member of this subgroup is present in budding yeast[12].

Uridylation of mRNAs and noncoding RNAs has been described in fission yeast and metazoans. Known substrates include miRNA-directed cleavage products[13] and replication-dependent histone mRNAs, which in metazoans contain a 3' stem-loop structure rather than a poly(A) tail, and decay of which is stimulated by oligouridylation[14–16]. In addition, we previously observed terminal uridyl residues on polyadenylated *S. pombe act1* mRNA during S-phase arrest[9]. Although it seemed likely that Cid1-mediated uridylation was generally involved in mRNA metabolism, the effect of such a modification on polyadenylated messages was unclear.

Here we have used a circularized rapid amplification of cDNA ends (cRACE) technique to capture mRNA decay intermediates in *S. pombe*. Unexpectedly, in contrast to the situation in *Saccharomyces cerevisiae*, we find that decapped intermediates often contain substantial poly(A) tails, indicative of a novel deadenylation-independent decapping pathway for bulk mRNA in fission yeast. We also show that

uridylation of polyadenylated mRNAs forms the basis for an additional 5'→3' decay pathway, probably conserved in higher eukaryotes, that elicits decapping and seems to be mediated by the Lsm1–7 complex. Uridylation and deadenylation have overlapping and distinct stimulatory effects on decapping of polyadenylated mRNAs.

RESULTS

cRACE captures *act1* mRNA degradation intermediates

To dissect RNA decay pathways in fission yeast, we used the cRACE technique[6]. A tail-independent method of capturing 3' and 5' ends, this procedure allows distinction between decapped and capped mRNAs (**Fig. 1a**). We first examined decapped and capped *act1* transcripts from exponentially growing wild-type cells.

We initially wished to determine whether those products isolated from decapped cRACE analysis were derived from decay intermediates. To do this, we compared the 5' ends isolated from capped and decapped mRNAs (**Fig. 1b**). The 5' ends of products from capped transcripts generally lay 57 nucleotides (nt) or 58 nt upstream from the start codon (**Fig. 1b** and **Supplementary Fig. 1** online). These nucleotides presumably represent the major transcriptional start site for *act1*. In contrast, the 5' ends of decapped products were heterogeneous, always downstream from the major transcriptional start site and distributed significantly differently from the capped species ($P < 0.0001$, two-tailed Mann-Whitney test). Thus, we conclude that these products represent mRNAs that have been subject to decapping and subsequent partial 5'→3' decay *in vivo*.

We next compared the 3' ends of adenylated and non-adenylated transcripts (**Fig. 1c**). Similarly to previous observations[17], we detected three cleavage and polyadenylation sites in the *act1* gene: the first is approximately 1,190 nt downstream from the start codon (and 60 nt

[1]Sir William Dunn School of Pathology, University of Oxford, UK. [2]Present address: Whitehead Institute, Cambridge, Massachusetts, USA. Correspondence should be addressed to C.J.N. (chris.norbury@path.ox.ac.uk).

Ten Simple Rules To Combine Teaching and Research

Quentin Vicens[1], Philip E. Bourne[2]*

1 University of Colorado, Boulder, Colorado, United States of America, 2 Skaggs School of Pharmacy and Pharmaceutical Science, University of California San Diego, La Jolla, California, United States of America

The late Lindley J. Stiles famously made himself an advocate for teaching during his professorship at the University of Colorado: "If a better world is your aim, all must agree: The best should teach" (http://thebestshouldteach.org/). In fact, dispensing high-quality teaching and professional education is the primary goal of any university [1]. Thus, for most faculty positions in academia, teaching is a significant requirement of the job. Yet, the higher education programs offered to Ph.D. students do not necessarily incorporate any form of teaching exposure. We offer 10 simple rules that should help you to get prepared for the challenge of teaching while keeping some composure.

Rule 1: Strictly Budget Your Time for Teaching and for Doing Research

This rule may seem straightforward, but respecting it actually requires more discipline and skill than it first appears to. The key is to set aside time for both teaching and research from the beginning, with a well-marked separation (e.g., mornings will be devoted to course preparation, afternoons to experiments and manuscript writing). Firmly stick to this agenda, particularly if this is your first time teaching. Failure to do so would eventually affect the quality of your teaching or the progress of your research (or both). Over time, you will become more skilled at jumping from one commitment to the other, and therefore allowing the boundaries to fluctuate somewhat. Avoid underestimating the time necessary to fulfill teaching-related obligations (e.g., office hours, test preparation, grading, etc.) by consulting with your colleagues.

Rule 2: Set Specific Teaching and Research Goals

In order not to have one occupation overpower the other one—which would transgress Rule #1—it is a good idea to decide on specific aims for each enterprise. Compile a list of reasonable but specific long-term goals (for the month or the semester) and short-term ones (for the week) for both your teaching (e.g., finish Chapter 3 by Nov. 1; this week propose a discussion to engage students to brainstorm about the risks of GMOs) and your research (e.g., finish experiments for this project and start writing before Easter; this week do the control for my primer binding assay). Make sure you achieve them. If you don't—this is likely to happen at first—ask yourself how legitimate your reason is. Then review and adjust the goals accordingly.

Rule 3: "Don't Reinvent the Wheel"

We borrowed the title for this rule from excellent suggestions on *How To Prepare New Courses While Keeping Your Sanity* [2]. Most likely, you will not be the first one ever to teach a particular topic. So get in touch with the colleagues in your department who have taught the class you are going to teach, or who teach similar topics. You can also use your network and contact former colleagues or friends at other institutions. They will usually be happy to share their course material, and along the way you might also glean precious tips from their teaching experience (e.g., a list of do's and don'ts on how to approach a notoriously difficult topic). You will also learn a lot from sitting in one of their classes and watching how they handle their topic and their students. Here are more examples of precious time-savers:

(1) Choose a textbook that is accompanied by rich online resources such as annotated figures, pre-made PowerPoint slides, animations, and videos. Students will thank you for showing movies, for example, as they often are a better option to break down complex mechanisms or sequences of events into distinct steps.

(2) Administer a Web site for your course. Many universities and some textbooks now offer you the possibility of hosting a Web site with course-related materials, including automatically graded assessments. See, for example, the CULearn suite used at the University of Colorado (http://www.colorado.edu/its/culearn/), or more general automatic grading tools presented at http://ctl.stanford.edu/Tomprof/postings/227.html.

(3) Gather a solid team of motivated teaching or learning assistants, who will both serve as an intermediary between you and your students and help you grade. In short, don't be afraid to ask for help!

Rule 4: Don't Try To Explain Everything

Class time should be spent guiding students to create their own explanation of the material and to develop cognitive abilities that will help them become critical thinkers. In other words, you don't want to present all aspects related to a certain topic or to lay out all the explanations for them. Thus, an effective way to teach is to get students to learn by transformative learning: beyond memorizing and comprehending basic concepts, they will learn to reflect on what they learn and how they learn it (see, for example, http://en.wikipedia.org/wiki/Transformative_learning and references within). Such teaching practices require that a significant part of the learning process happens outside the classroom, through reading assignments, homework, writing essays, etc. So make sure you budget time to organize these, as specified

What Is Stochastic Resonance? Definitions, Misconceptions, Debates, and Its Relevance to Biology

Mark D. McDonnell[1]*, Derek Abbott[2]

1 Institute for Telecommunications Research, University of South Australia, Mawson Lakes, South Australia, Australia, 2 Centre for Biomedical Engineering and School of Electrical & Electronic Engineering, University of Adelaide, Adelaide, South Australia, Australia

Abstract: Stochastic resonance is said to be observed when increases in levels of unpredictable fluctuations—e.g., random noise—cause an increase in a metric of the quality of signal transmission or detection performance, rather than a decrease. This counterintuitive effect relies on system nonlinearities and on some parameter ranges being "suboptimal". Stochastic resonance has been observed, quantified, and described in a physical and biological systems, including neurons. Being a topic of widespread multidisciplinary interest, the definition of stochastic resonance has evolved significantly over the last decade or so, leading to a number of debates, misunderstandings, and controversies. Perhaps the most important debate is whether the brain has evolved to utilize random noise in vivo, as part of the "neural code". Surprisingly, this debate has been for the most part ignored by neuroscientists, despite much indirect evidence of a positive role for noise in the brain. We explore some of the reasons for this and argue why it would be more surprising if the brain did not exploit randomness provided by noise—via stochastic resonance or otherwise—than if it did. We also challenge neuroscientists and biologists, both computational and experimental, to embrace a very broad definition of stochastic resonance in terms of signal-processing "noise benefits", and to devise experiments aimed at verifying that random variability can play a functional role in the brain, nervous system, or other areas of biology.

Introduction

Noise is an all-encompassing term that usually describes undesirable disturbances or fluctuations. In biology, "noise" typically refers to variability in measured data when identical experiments are repeated, or when biosignals cannot be measured without background fluctuations distorting the desired measurement.

Noise is also the fundamental enemy for communications engineers, whose goal is to ensure messages can be transmitted error-free and efficiently from one place to another, at the fastest possible rate. When *random noise* in the form of electronic fluctuations or electromagnetic interference corrupts transmitted messages, this places limits on the rate at which error-free communication can be achieved. If everything else is ideal, then noise is the enemy.

But what if not everything is ideal? Can an ideal system always be implemented in practice? The answer is of course *no*; engineering is about designing systems with tradeoffs between different conflicting objectives. The same could be said of evolution. Given this, there are circumstances—see below—where unavoidably present noise or unpredictable fluctuations can be used purposely, or deliberately introduced to lead to a benefit.

Stochastic Resonance (SR) is the name for a phenomenon that is a flagship example of this idea. It has mostly been studied by physicists—for more than 25 years—but may also be familiar to some biologists, as well as to those in many other disciplines. Research into SR has had a colorful evolutionary journey, and extracting important principles and results from the literature can be confusing.

In particular, the paradoxical notion of "good noise" is a double-edged sword for SR researchers. To some, working to understand paradoxes and counterintuitive ideas is a significant curiosity-driven challenge. This has drawn many scientists and engineers to study SR, leading to many interesting and useful theoretical and experimental published results. Others naturally focus only on the ingrained principle of great utility in engineering, where noise needs to be eradicated, and, the more it is present, the more diminished is performance. This preconception can be a sufficient reason for many scientists to ignore or dismiss SR.

The purpose of this essay is to discuss issues that have sometimes clouded the topic. Our first aim is to bring some clarity to the debate and to illustrate the pitfalls and controversies for biologists unfamiliar with stochastic resonance, or who have held only a peripheral interest in the area.

The second aim is to advocate to readers that when they come across studies of SR, they should focus less on the counterintuitive idea of "good noise", and instead understand SR in terms of "randomness that makes a nonlinearity less detrimental to a signal". Such a change of focus away from "noise" and on to "helpful randomness" may shift the balance away from skepticism of the form "how can *noise* be good?" toward thinking "does this variability have a useful function?"

Provoking a discussion on this topic is especially timely, given recent increasing interest in the topic of neuronal variability. For example, SR is mentioned in several recent PLoS articles [1–4], while a symposium on "Neuronal Variability and Its Functional Significance" was held in conjunction with the 2008 Society for Neuroscience meeting. Furthermore, a recent review on noise in

PLoS one

Protein Dynamics in Individual Human Cells: Experiment and Theory

Ariel Aharon Cohen[1,◕], Tomer Kalisky[2,◕], Avi Mayo[1,◕], Naama Geva-Zatorsky[1], Tamar Danon[1], Irina Issaeva[1], Ronen Benjamine Kopito[3], Natalie Perzov[1], Ron Milo[4], Alex Sigal[5], Uri Alon[1]*

1 Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel, 2 Department of Bioengineering, Stanford University and Howard Hughes Medical Institute, Stanford, California, United States of America, 3 Department of Materials and Interfaces, Weizmann Institute of Science, Rehovot, Israel, 4 Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, United States of America, 5 Division of Biology, California Institute of Technology, Pasadena, California, United States of America

Abstract

A current challenge in biology is to understand the dynamics of protein circuits in living human cells. Can one define and test equations for the dynamics and variability of a protein over time? Here, we address this experimentally and theoretically, by means of accurate time-resolved measurements of endogenously tagged proteins in individual human cells. As a model system, we choose three stable proteins displaying cell-cycle-dependant dynamics. We find that protein accumulation with time per cell is quadratic for proteins with long mRNA life times and approximately linear for a protein with short mRNA lifetime. Both behaviors correspond to a classical model of transcription and translation. A stochastic model, in which genes slowly switch between ON and OFF states, captures measured cell–cell variability. The data suggests, in accordance with the model, that switching to the gene ON state is exponentially distributed and that the cell–cell distribution of protein levels can be approximated by a Gamma distribution throughout the cell cycle. These results suggest that relatively simple models may describe protein dynamics in individual human cells.

Introduction

A goal of systems biology is to understand the dynamics of protein levels, as proteins are produced and degraded. One aims for a mathematical description of these processes that captures the essentials and that can be understood intuitively.

Most current models of protein dynamics have been established and tested in micro-organisms. Proteins in bacterial cells, for example, are well described by production-degradation equations. These show that protein mass increases exponentially over time in growing cells, and that protein concentration approaches steady-state with exponential decays [1–5]. Such models have been shown to be useful for describing the measured dynamics of protein circuits with several interactions, such as feed-forward loop network motifs [6] and auto-regulatory loops [4,7–9]. Similarly, experiment and theory has established an understanding of cell-cell variability of protein levels in bacterial and yeast cells [10–15]. Protein distributions in micro-organisms have been measured at steady-state and are well described by Gamma distributions [16,17].

Much less is known about the cell-cell variability of protein dynamics in human cells [18]. Differences in expression levels between individual cells of a clonal population were previously observed [19–22] and related to stochastic processes. More recently, Raj et al. [23] have followed mRNA and protein to show bursts of mRNA production. Sigal et al. [24] measured variability and memory in nuclear proteins, finding that the time it

takes for cells expressing a protein above or below the population mean to 'relax' towards the average is longer than one cell generation. More work is needed to understand the basic dynamics and differences of human protein levels and its variation across the cell cycle in individual cells.

To address this, we experimentally followed the dynamics of selected human proteins in human cancer cells and test simple models to describe these dynamics. The experiments were made possible by recent advances in obtaining dynamics of endogenously expressed proteins in individual human cells at very high resolution and accuracy [25]. This is based on a library of cell clones, each expressing a different fluorescently tagged protein expressed from its natural chromosomal location and under its native regulation. From this library, we chose three proteins that accumulate throughout the cell cycle, but are degraded every time that the cell divides. These proteins show little memory of previous cell cycles, and thus are a good starting point for testing models based on current cell properties. In the present study, we calibrated this assay to provide units of protein concentration.

We find that the proteins are found to follow either an approximately linear or a quadratic dependence on time. They also differ in their onset time of production and in their variability across the cell cycle. We find that these features can be understood in terms of a single unified model. Approximately linear accumulation results from short mRNA half-life, and quadratic accumulation from long mRNA half-life, as suggested by

Stochasticity in Protein Levels Drives Colinearity of Gene Order in Metabolic Operons of *Escherichia coli*

Károly Kovács[1], Laurence D. Hurst[2]*, Balázs Papp[1]*

1 Institute of Biochemistry, Biological Research Center, Szeged, Hungary, 2 Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom

Abstract

In bacterial genomes, gene order is not random. This is most evident when looking at operons, these often encoding enzymes involved in the same metabolic pathway or proteins from the same complex. Is gene order within operons nonrandom, however, and if so why? We examine this issue using metabolic operons as a case study. Using the metabolic network of *Escherichia coli*, we define the temporal order of reactions. We find a pronounced trend for genes to appear in operons in the same order as they are needed in metabolism (colinearity). This is paradoxical as, at steady state, enzymes abundance should be independent of order within the operon. We consider three extensions of the steady-state model that could potentially account for colinearity: (1) increased productivity associated with higher expression levels of the most 5' genes, (2) a faster metabolic processing immediately after up-regulation, and (3) metabolic stalling owing to stochastic protein loss. We establish the validity of these hypotheses by employing deterministic and stochastic models of enzyme kinetics. The stochastic stalling hypothesis correctly and uniquely predicts that colinearity is more pronounced both for lowly expressed operons and for genes that are not physically adjacent. The alternative models fail to find any support. These results support the view that stochasticity is a pervasive problem to a cell and that gene order evolution can be driven by the selective consequences of fluctuations in protein levels.

Introduction

It is well established that the chromosomal distribution of genes is not random, and in many genomes, genes that need to be coexpressed tend to cluster [1–4]. The coexpression of adjacent genes can be enforced by the action of bidirectional promoters [5], simultaneous opening and closing of chromatin [6,7], transcriptional spill-over [8], or by inclusion within the same operon [9,10]. Given that prokaryotic operons generally contain functionally related genes that need to be expressed together [3], gene order evolution in bacteria is often considered to be driven by coexpression (although other scenarios have also been proposed to explain the origin of operons [11,12]). Such coexpression models do not obviously predict that within an operon there need be selection on gene order. However, a relationship between bacterial morphology and the relative order of genes in a cluster involved in cell division [13] provides some evidence for adaptive gene organization within an operon. Furthermore, a prior comparative genomics study found that horizontal transfer of operonic genes often involves in situ gene displacement by an ortholog from a distant organism without change of the local gene organization [14], hinting at the presence of selection on intraoperonic gene order per se. Nevertheless, it remains unclear whether these phenomena are restricted to certain gene clusters only or whether they could be a more general property of bacterial operons, and most importantly, what selective forces might be responsible for these genomic patterns. Here, then, we ask whether gene order within operons is under selection and if so why? In particular, we investigate whether gene order within metabolic operons of *E. coli* reflect the functional order of the encoded enzymes, i.e., colinearity (Figure 1).

Results

Excess of Colinearity in Metabolic Operons

To examine whether gene order within metabolic operons reflect the functional order of the encoded enzymes, we focused on *E. coli*, where high-quality and high-coverage data are available on both biochemical pathways and operon structures. We compiled data on operons encoding at least two enzymes in the same biochemical pathway according to EcoCyc [15], resulting in a list of 70 operons and 321 intraoperonic gene pairs (Methods). For each intraoperonic enzymatic gene pair, we recorded whether their relative position in the operon corresponds to their functional order and hence displays colinearity. Approximately 60% of the 321 gene pairs showed colinearity, compared to 50% expected if intraoperonic gene order was random ($p = 0.0011$, from randomisation, see Methods).

Theories to Explain Colinearity

At first sight, the above result is unexpected as gene order should not affect the steady-state pathway productivity under the

Stochasticity of gene products from transcriptional pulsing

Srividya Iyer-Biswas,[1,*] F. Hayot,[2] and C. Jayaprakash[1]

[1]*Department of Physics, Ohio State University, Woodruff Avenue, Columbus, Ohio 43210, USA*
[2]*Department of Neurology, Mount Sinai School of Medicine, Levy Place, New York, New York 10029, USA*
(Received 28 October 2008; published 23 March 2009)

Transcriptional pulsing has been observed in both prokaryotes and eukaryotes and plays a crucial role in cell-to-cell variability of protein and mRNA numbers. An important issue is how the time constants associated with episodes of transcriptional bursting and mRNA and protein degradation rates lead to different cellular mRNA and protein distributions, starting from the transient regime leading to the steady state. We address this by deriving and then investigating the exact time-dependent solution of the master equation for a transcriptional pulsing model of mRNA distributions. We find a plethora of results. We show that, among others, bimodal and long-tailed (power-law) distributions occur in the steady state as the rate constants are varied over biologically significant time scales. Since steady state may not be reached experimentally we present results for the time evolution of the distributions. Because cellular behavior is determined by proteins, we also investigate the effect of the different mRNA distributions on the corresponding protein distributions using numerical simulations.

I. INTRODUCTION

Cell-to-cell variability in mRNA and protein numbers is now recognized as a major aspect of cellular response to stimuli, a variability which is hidden in cell population studies. The most egregious example of the latter is provided in cases where a graded average response hides the all-or-nothing behavior of single cells [1–3]. Variability of cellular response can have many origins, which are generally classified as extrinsic and intrinsic noise or fluctuations [4]. Many studies, both experimental and theoretical from bacteria to eukaryotes, have been undertaken to characterize intrinsic and extrinsic fluctuations [4–11]. Extrinsic fluctuations can have multiple origins, such as variations from cell to cell in the number of regulatory molecules or signaling cascade components. The source of intrinsic fluctuations is the random occurrence of reactions that can lead to variability for genetically identical cells in identical fixed environments. When the number of molecules is small, intrinsic noise can play a significant role in determining the behavior of individual cells. The difficulty in determining rate constants experimentally and the possibility of experiments being either in the steady state or in the transient regime make exact time-dependent solutions to models especially useful in the interpreting observed behavior of intrinsic noise. In this paper we present the exact time-dependent solution to a widely used model for transcriptional noise and discuss its implications.

II. EXPERIMENTAL AND THEORETICAL BACKGROUNDS

Intrinsic fluctuations arise from either noisy transcription, noisy translation, or both, the effects of which can be measured in single cell mRNA and protein experiments. The sim-

plest model of mRNA and protein number distributions is to consider both transcription and translation as Poisson processes [7]. Recent experimental studies of mRNA distributions have shown strong evidence for transcriptional noise beyond what can be described by a simple Poisson process, i.e., distributions with variances significantly larger than the mean have been observed. In particular, transcriptional pulsing, where bursts of transcription alternate with quiescent periods, has been observed in both prokaryotes and eukaryotes. Raser and O'Shea [5], who studied intrinsic and extrinsic noise in *Saccharomyces cerevisiae*, showed that the noise associated with a particular promoter could be explained in a transcriptional pulsing model and confirmed it by mutational analysis. Transcriptional bursts were recorded in *E. coli* [12] by following mRNA production in time and their statistics computed. Evidence for a pulsing model of transcription, obtained from fluorescent microscopy, has also been presented for the expression of the discoidin Ia gene of *Dictyostelium* [13]. Transcriptional bursts have as well been detected in Chinese hamster ovary cells [10]. In these experiments the production of mRNA occurs in a sequence of bursts of transcriptional activity separated by quiescent periods. Transcriptional bursting, an intrinsically random phenomenon, thus becomes an important element to consider when evaluating cell-to-cell variability. One can predict that in many cases it will be a significant part of overall noise and most certainly of intrinsic noise.

Our study focuses on the consequences of transcriptional bursting in a simplified model of transcription that has been the subject of many studies and is believed to encapsulate the key features of bursting [2,5,10,12,14–16]. Related models that include feedback have been studied theoretically [17–19]. The complex phenomena that can occur in transcription (chromatin remodeling, enhanceosome formation, preinitiation assembly, etc.) are modeled through positing two states of gene activity: an inactive state, where no transcription occurs, and an active one, in which transcription occurs according to a Poisson process. The production of mRNA is thus pulsatile: temporally there are periods of in-

*srividya@mps.ohio-state.edu

Studying membrane proteins through the eyes of the genetic code revealed a strong uracil bias in their coding mRNAs

Jaime Prilusky[a] and Eitan Bibi[b,1]

[a]Bioinformatics Unit, Department of Biological Services, and [b]Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel

Posttranscriptional processes often involve specific signals in mRNAs. Because mRNAs of integral membrane proteins across evolution are usually translated at distinct locations, we searched for universally conserved specific features in this group of mRNAs. Our analysis revealed that codons of very hydrophobic amino acids, highly represented in integral membrane proteins, are composed of 50% uracils (U). As expected from such a strong U bias, the calculated U profiles of mRNAs closely resemble the hydrophobicity profiles of their encoded proteins and may designate genes encoding integral membrane proteins, even in the absence of information on ORFs. We also show that, unexpectedly, the U-richness phenomenon is not merely a consequence of the codon composition of very hydrophobic amino acids, because counterintuitively, the relatively hydrophilic serine and tyrosine, also encoded by U-rich codons, are overrepresented in integral membrane proteins. Interestingly, although the U-richness phenomenon is conserved, there is an evolutionary trend that minimizes usage of U-rich codons. Taken together, the results suggest that U-richness is an evolutionarily ancient feature of mRNAs encoding integral membrane proteins, which might serve as a physiologically relevant distinctive signature to this group of mRNAs.

evolution | hydrophobicity scale | mRNA targeting | U-rich mRNA

In addition to protein-coding information, mRNAs sometimes harbor signals required for posttranscriptional regulatory pathways, such as processing, translation, degradation, and localization (1, 2). For selective targeting, mRNAs use various protein-interaction determinants (structural, sequence specific, or nonspecific) (3), mostly in untranslated regions, although unique exceptions have been described (e.g., ref. 4). mRNAs of integral membrane proteins across evolution are usually translated at distinct locations, and our studies in *Escherichia coli* have suggested a step through which these mRNAs might be selectively targeted to membrane-bound ribosomes (5–7). Therefore, we investigated the possibility that mRNAs encoding integral membrane proteins have species-independent characteristic features, which might provide an evolutionarily conserved means for their selective recognition and targeting to the membrane. As a basis for our analysis, we reasoned that because prokaryotes express polycistronic transcripts, sometimes encoding a mixture of membrane and cytosolic proteins, targeting signals might be located inside ORFs in addition to untranslated regions.

Results and Discussion

Analysis of the Genetic Code in mRNAs Encoding Integral Membrane Proteins.
A unique property of integral membrane proteins is that they have stretches containing very hydrophobic amino acid residues (≈20). Therefore, we analyzed the nucleotide composition of very hydrophobic codons [according to the Goldman, Engelman, Steitz (GES) scale] (8), in the context of the entire genetic code. Since the first description of the nearly universal genetic code (for a review see ref. 9), various explanations of its organization and the assignment of the 64 triplets have been

offered (10, 11). It was soon realized that chemically similar amino acids are often encoded by relatively similar codons (12, 13) and that very hydrophobic amino acids are encoded by codons having uracil (U) in the second position (14). Our analysis revealed that, in addition to their second position, codons of very hydrophobic amino acids have a remarkably high U content in general (Fig. 1). Specifically, 50% of the combined numbers of nucleotides in these codons are Us. In contrast, the U content in codons of all other groups of amino acids is ≤22%, and the total U content in all of the 61-aa coding triplets is 24.6%, suggesting a strong U bias in mRNAs encoding integral membrane proteins. Next, we investigated whether the proposed U bias is an inherent requirement for mRNAs encoding integral membrane proteins or merely a trivial consequence of the high U content in codons of very hydrophobic amino acids. As shown in Fig. 2A, there are 2 relatively hydrophilic amino acids, serine and tyrosine, both encoded by U-rich codons (33% and 50% U, respectively). On the basis of the chemical nature of serine and tyrosine, we predicted that both of them should be more abundant in soluble proteins. In contrast, analysis of their usage in multipass membrane and cytoplasmic proteins from various organisms (supporting information Tables S1 and S2) revealed higher content of serine and tyrosine in the membrane protein group (≈20% more) (Fig. 2B). These results strongly suggest that integral membrane protein transcripts might have been programmed or had evolved to contain high contents of U.

Analysis of the Distribution of U in Membrane Protein mRNAs.
Traditionally, integral membrane proteins are analyzed for their hydrophobicity profiles (15), using algorithms that help identify their transmembrane helices. We wondered whether our discovery of the high U content of very hydrophobic codons might be helpful in identifying integral membrane proteins through the analysis of U profiles of genes. Initially, we compared the hydrophobicity profiles of several integral membrane proteins with the U profiles of their mRNAs. Fig. 3 shows several examples in which both curves are strikingly similar. MalF is a complex integral membrane protein with 8 transmembrane helices and a large external hydrophilic domain (Fig. 3A, Top) (16). The calculated Kyte-Doolittle-based hydrophobicity profile of MalF (Fig. 3A, Middle) supports the proposed secondary structure model. Remarkably, the calculated U profile of the malF mRNA also supports this model, and the 2 profiles are very similar. Notably, although similar, there are subtle differences that might indicate the importance of features other than the identified relationship between the protein hy-

Complexity in bacterial cell–cell communication: Quorum signal integration and subpopulation signaling in the *Bacillus subtilis* phosphorelay

Ilka B. Bischofs[a,b,1], Joshua A. Hug[c], Aiwen W. Liu[a], Denise M. Wolf[b], and Adam P. Arkin[a,b,1]

Departments of [a]Bioengineering and [c]Electrical Engineering, University of California, 311 Hildebrand Hall, MC 5230, Berkeley, CA 94704-3224; and [b]Physical Biosciences Division, Lawrence Berkeley Laboratory, 1 Cyclotron Road, MS Calvin, Berkeley, CA 94720

A common form of quorum sensing in Gram-positive bacteria is mediated by peptides that act as phosphatase regulators (Phr) of receptor aspartyl phosphatases (Raps). In *Bacillus subtilis*, several Phr signals are integrated in sporulation phosphorelay signal transduction. We theoretically demonstrate that the phosphorelay can act as a computational machine performing a sensitive division operation of kinase-encoded signals by quorum-modulated Rap signals, indicative of cells computing a "food per cell" estimate to decide whether to enter sporulation. We predict expression from the *rapA-phrA* operon to bifurcate as relative environmental signals change in a developing population. We experimentally observe that the *rapA-phrA* operon is heterogeneously induced in sporulating microcolonies. Uninduced cells sporulate rather synchronously early on, whereas the RapA/PhrA subpopulation sporulates less synchronously throughout later stationary phase. Moreover, we show that cells sustain PhrA expression during periods of active growth. Together with the model, these findings suggest that the phosphorelay may normalize environmental signals by the size of the (sub)population actively competing for nutrients (as signaled by PhrA). Generalizing this concept, the various Phrs could facilitate subpopulation communication in dense isogenic communities to control the physiological strategies followed by differentiated subpopulations by interpreting (environmental) signals based on the spatiotemporal community structure.

heterogeneity | Phr | quorum sensing | sporulation | model

The 11 homologous *rap-phr* genes in *Bacillus subtilis* code for a family of proteins in which Rap activity is regulated by small peptides derived from the cognate *phr* gene product (Fig. 1A) (1). The small Phr precursor peptides are cleaved and secreted from the cell, and at least some of them accumulate in the culture supernatant (2). They are subsequently imported by Opp oligopeptide permeases into the cytoplasm where the Phr derived pentapeptides (PEP5) intracellularly inhibit the activity of their cognate Raps (3–5). Thus, PEP5s can be indicators of cell density and have been implied in facilitating cell–cell communication (2, 5–8). A subset of Raps (RapA,B,E,H) act as phosphatases that inhibit signaling through a central phosphorelay, and their activity is counteracted by their cognate PEP5s (PhrA,C,E,H). The phosphorelay plays a central role in *B. subtilis* stress response induction during stationary phase and ultimately controls sporulation induction (Fig. 1B) (9).

Unlike autoinducer quorum signals, where autoinducer molecules bind to a transcription factor, peptide based quorum signaling indirectly regulates transcription by controlling phosphoryl signaling (10). The presence of an autoinducer signal is typically sufficient to activate a pathway and induce a cell response (Fig. 2A *Left*). In contrast, because the Phr signal acts on a phosphatase, it instructs cells by regulating the phosphoryl flux originating from the activation of kinases that are under control of the environment (Fig. 2A *Right*). For example, it is well known that sporulation is affected by cell density. However, a high cell density alone is not sufficient to trigger spore

formation unless there is a starvation signal in addition, whereas starvation alone is able to trigger sporulation, albeit at a reduced efficiency (5, 11). It therefore appears that the Rap-Phr-systems have evolved toward a mode of quorum signaling to which signal integration is essential.

It has recently been shown that even monocultures show a large degree of diversity by heterogeneous cell differentiation and population structuring (12). Stochastic population splitting has been demonstrated for a variety of stress responses including sporulation, competence, motility, extracellular matrix production, and exoenzyme synthesis (13–16). The coordinated development of specialized differentiated subpopulations implies a potential need for cell–cell signaling systems that facilitate communication among and across differentiating subpopulations. The Rap-Phr genes comprise a promising family to mediate such subpopulation signaling.

Here, we begin to explore these two aspects of Rap-Phrs in phosphorelay signaling: signal integration and subpopulation signaling. First, we develop a computational model to address the interplay of quorum and environmental signals in phosphorelay signaling and show that the signal transduction architecture supports a sensitive ratiometric integration of kinase-encoded environmental signals with respect to quorum-modulated phosphatase signals. This is indicative of a cell-density-dependent normalization of the environmental stressors such as starvation to drive downstream decision making. Within this theoretical framework, we predict a bifurcation of RapA/PhrA expression as relative environmental conditions change. We then experimentally elucidate the dynamics and fate of the previously observed subpopulation of RapA/PhrA-expressing cells (17) late into stationary phase and suggest the PhrA/RapA may be a subpopulation communication system to measure its own population size. Although the experiments do not directly test the theory, the theory aids in interpreting the importance of the experimental observations. Together, our computational and experimental results suggest a model in which phosphorelay-based decision making during starvation is based on normalizing environmental signals (such as "food") with respect to the size of a growing subpopulation actively competing for nutrients ("food per growing cell"). Generalizing this concept, we suggest that phosphorelay-integrated Rap-Phrs play a central role in determining and maintaining cell differentiation development by integrating (environmental) signals based on the local community structure.

Transcriptome transfer produces a predictable cellular phenotype

Jai-Yoon Sul[a,1], Chia-wen K. Wu[a,1], Fanyi Zeng[a,b,1], Jeanine Jochems[a], Miler T. Lee[c,d], Tae Kyung Kim[a], Tiina Peritz[a], Peter Buckley[a,c], David J. Cappelleri[e], Margaret Maronski[f], Minsun Kim[g], Vijay Kumar[c,e], David Meaney[c,h], Junhyong Kim[c,d,1], and James Eberwine[a,c,i,1,2]

Departments of [a]Pharmacology, [c]Penn Genome Frontiers Institute, [d]Biology, [e]Mechanical Engineering and Applied Mechanics, [f]Neuroscience, [h]Bioengineering, and [i]Psychiatry, University of Pennsylvania, Philadelphia, PA 19104; [b]Shanghai Institute of Medical Genetics and Institute of Medical Sciences, Shanghai Jiao Tong University School of Medicine, Shanghai 200040, Peoples Republic of China; and [g]Department of Physiology, Wonkwang University School of Medicine, Iksan 570-749, South Korea

Cellular phenotype is the conglomerate of multiple cellular processes involving gene and protein expression that result in the elaboration of a cell's particular morphology and function. It has been thought that differentiated postmitotic cells have their genomes hard wired, with little ability for phenotypic plasticity. Here we show that transfer of the transcriptome from differentiated rat astrocytes into a nondividing differentiated rat neuron resulted in the conversion of the neuron into a functional astrocyte-like cell in a time-dependent manner. This single-cell study permits high resolution of molecular and functional components that underlie phenotype identity. The RNA population from astrocytes contains RNAs in the appropriate relative abundances that give rise to regulatory RNAs and translated proteins that enable astrocyte identity. When transferred into the postmitotic neuron, the astrocyte RNA population converts 44% of the neuronal host cells into the destination astrocyte-like phenotype. In support of this observation, quantitative measures of cellular morphology, single-cell PCR, single-cell microarray, and single-cell functional analyses have been performed. The host-cell phenotypic changes develop over many weeks and are persistent. We call this process of RNA-induced phenotype changes, transcriptome-induced phenotype remodeling.

neuron | transcriptome-induced phenotype remodeling | single cell | Waddington

In multicellular organisms, all cells contain nearly identical copies of the genome but exhibit drastically different phenotypes. Even a single neuron has a set of phenotypic characteristics that distinguish it from other neurons as well as other cell types, such as the nearby astrocytes. Indeed, as Waddington proposed in his classical epigenetic landscape model, genetically predetermined cells can follow any specific permitted trajectories that eventually lead to different cellular phenotypes (1). From this point of view, the genome serves as a repository of dynamic control information whose state can be reprogrammed to match the stable phenotypic states.

Emerging evidence has demonstrated the reversibility and flexibility of the cellular phenotype. Gurdon et al. first showed that the ability to obtain fertile adult male and female frogs by injecting endoderm nuclei into enucleated eggs (2). This result not only forms the foundation of the field in nuclear transplantation but also provides evidence that the cytoplasmic components of a differentiated cell can support nuclear reprograming. Generation of induced pluripotent stem (iPS) cells by transfection of transcription factors into dividing fibroblasts (3), followed by cell selection, represents a new strategy to globally revert a mature cell into a different cell type (4–9). The need for redifferentiation of these embryonic stem cell-like-iPS cells into desired cell types adds a layer of complexity that is difficult to control (10, 11). Nevertheless, studies of nuclear reprogramming from genomic and epigenetic modification, as seen from somatic-cell nuclear-transfer-cloned animals and iPS cells, strongly demonstrate the flexibility of a

differentiated phenotype, as well as the dynamic changes of a genome (7). A key question is whether there is a general strategy to efficiently reprogram a cell to any target-cell type.

cDNA microarray analysis has shown that phenotypic differences at the cellular level are associated with differences in the presence, absence, and abundance of particular RNAs. As the transcriptome contains the RNAs that encode the proteins necessary to generate and maintain phenotype, such as kinases, phosphatase, histone acetylases and deacetylases, and other proteins, including transcription factors that collaborate in the right proportions to generate the phenotype, it is reasonable to assume it has a dominant role in determining phenotype. To test this hypothesis using single-cell approaches, we introduced the transcriptome from donor cells into individual host cells and quantitatively assessed conversion from the host to the donor cell phenotype (12, 13).

Results

Modeling of the Phototransfection of Complex RNA Populations. To evaluate whether the astrocyte transcriptome can directly convert host neurons into astrocytes, we first characterized our cell cultures to ensure the purity of the neurons that are to be transfected [supporting information (SI) Fig. S1] and then established a working transcriptome-induced phenotype remodeling (TIPeR) protocol, which consists of multiple RNA phototransfections over a 10-day period (Fig. 1A). Phototransfection was selected as the means to transfect the astrocyte transcriptome into neurons, as it can transfect RNA into neurons with high efficiency (12). To optimize for the amount of RNA that can be introduced into the host cells, we modeled the phototransfection of RNA transcripts that diffuse into the photoinduced pores on the cell membranes (Fig. S2). This process was modeled with an average transcript size of 1.5 kb $\pm$ 0.2 kb with the effective radius of the transcript determined by the Flory approximation ($R_{transcript} \approx 5.5N^{1/3}$, where N is the size of the transcript in bases). Pore sizes were systematically varied to test the relative flux of transcriptome in a typical phototransfection experiment. Simulations with the approximated transcriptome cargo showed that a single sequence of 16 pulses across the cell membrane would be sufficient to deliver a large number of transcripts while retaining their relative abundances (Fig. 1 B and

The cost of gene expression underlies a fitness trade-off in yeast

Gregory I. Lang[a,1], Andrew W. Murray[b,1,2], and David Botstein[a,1,2]

[a]Lewis-Sigler Institute for Integrative Genomics and the Department of Molecular Biology, Princeton University, Princeton, NJ 08544; and [b]FAS Center for Systems Biology and the Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138

Natural selection optimizes an organism's genotype within the context of its environment. Adaptations to one environment can decrease fitness in another, revealing evolutionary trade-offs. Here, we show that the cost of gene expression underlies a trade-off between growth rate and mating efficiency in the yeast *Saccharomyces cerevisiae*. During asexual growth, mutations that eliminate the ability to mate provide an ≈2% per-generation growth-rate advantage. Some strains, including most laboratory strains, carry an allele of *GPA1* (an upstream component of the mating pathway) that increases mating efficiency by ≈30% per round of mating at the cost of an ≈1% per-generation growth-rate disadvantage. In addition to demonstrating a trade-off between growth rate and mating efficiency, our results illustrate differences in the selective pressures defining fitness in the laboratory versus the natural environment and show that selection, acting on the cost of gene expression, can optimize expression levels and promote gene loss.

evolution | GPA1 | mating pathway | *Saccharomyces cerevisiae*

A frequent observation in evolution is that traits not maintained by selection will be lost—this holds true at the morphological level and at the genetic level. Examples of gene loss include the loss of olfactory receptors in primates (1), the loss of pigmentation and vision in *Astyanax* cavefish (2), the loss of the galactose utilization pathway in yeast (3), and the degeneration of genes involved in carbon utilization during domestication of *Streptococcus thermophilus* (4). Such regressive evolution also occurs in laboratory populations; reduction in catabolic breadth and thermal tolerance is observed during long-term evolution in *Escherichia coli* (5–9), and sterility frequently arises during long-term asexual propagation of *Saccharomyces cerevisiae* (10).

Two mechanisms could account for gene loss during evolution. One possibility is that in the absence of selection, genes are lost because of the neutral accumulation of mutations. Alternatively, gene loss events could be driven by selection. The observation that many of these gene-loss events are repeatedly observed supports this hypothesis. Repeated loss of all or part of the *Rbs* operon (whose products catabolize ribose) in *E. coli* provides a selective advantage in minimal glucose media (8). Quantitative analysis of alleles leading to eye reduction in *Astayanax* indicates that selection, possibly against the energetic cost of vision, is responsible for eye degeneration in cavefish populations (11). These studies suggest that haploid yeast that are propagated for long periods without mating partners should become sterile. Previous studies, however, showed that lineages that evolved higher growth rates and lower mating efficiencies appeared to segregate these traits independently (10). Here, we set out to directly test whether selection drives yeast to become sterile by determining whether mutations conferring sterility provide a selective advantage.

Results

Sterility Increases Growth Rate by Eliminating Unnecessary Gene Expression. We tested the hypothesis that sterile strains generally have a growth-rate advantage by isolating sterile mutants and testing their fitness. Haploid, a-mating type (*MAT*a) cells, arrest in G1 when exposed to the mating pheromone, alpha-factor (αF), and thus cannot form colonies on media containing αF. We initiated, from a single colony of haploid *MAT*a cells, a large number of parallel cultures that were plated onto either rich media or rich media containing αF. On rich media, the vast majority of cells form colonies, but on αF, only the small fraction of cells that have acquired mutations in pheromone-induced signaling can form colonies. From each culture, we randomly chose a single alpha-factor resistant (αF^R) or unselected colony and measured its relative growth rate by using a FACS-based competitive growth-rate assay that can detect growth-rate differences as small as 0.5%. The growth-rate coefficient is a measure of the growth-rate advantage over wild type. Fig. 1*A* shows the growth-rate coefficients (s_g) for 27 unselected clones and 45 αF^R clones. As a control we measured the relative growth rates of 24 similarly selected mutants that were resistant to canavanine, a toxic arginine analog. In each case, several clones have a low growth rate ($s_g < -1\%$), suggesting that these strains have become mitochondrial deficient or have acquired a deleterious mutation. Excluding clones with $s_g < -1\%$, the growth-rate coefficients of the unselected clones follow a tight distribution (Fig. 1*A*, $s_g = 0.08\% \pm 0.35\%$) indistinguishable from the distribution of the canavanine-resistant mutants (Fig. 1*A*, $s_g = 0.36\% \pm 0.48\%, P > 0.05$, Wilcoxon rank sum test); however, the growth-rate coefficients of the αF^R mutants show greater variation and a positive growth-rate advantage (Fig. 1*A*, $s_g = 1.48\% \pm 0.85\%, P < 10^{-7}$, Wilcoxon).

It appears from these data that at least some sterile mutants have a clear growth-rate advantage over wild type. To determine whether all sterile strains have a similar advantage, and to determine the basis for any growth-rate advantage in the sterile strains, we used a combination of 4 methods: Phenotypic characterization of the spontaneous αF^R mutants, growth-rate assays on targeted gene deletions within the mating pathway, mapping of the mutations in the most fit sterile strains, and expression analysis on αF^R strains both with and without a growth-rate advantage.

The yeast mating pathway is one of the best studied mitogen-activated protein (MAP) kinase cascades (12). At the beginning of the pathway is a pheromone receptor (Ste2 in *MAT*a or Ste3 in *MAT*α) that binds the cognate mating pheromone. Receptor stimulation activates a heterotrimeric G protein (consisting of Gpa1, Ste18, and Ste4), which in turn, activates a MAP kinase cascade (consisting of the MAP kinase kinase kinase, Ste11, the MAP kinase kinase, Ste7, the MAP kinases Fus3 and Kss1, and

GENETICS

A stochastic spectral analysis of transcriptional regulatory cascades

Aleksandra M. Walczak[a,1], Andrew Mugler[b], and Chris H. Wiggins[c]

[a]Princeton Center for Theoretical Science, Princeton University, Princeton, NJ 08544; and Departments of [b]Physics and [c]Applied Physics and Applied Mathematics, Center for Computational Biology and Bioinformatics, Columbia University, New York, NY 10027

The past decade has seen great advances in our understanding of the role of noise in gene regulation and the physical limits to signaling in biological networks. Here, we introduce the spectral method for computation of the joint probability distribution over all species in a biological network. The spectral method exploits the natural eigenfunctions of the master equation of birth–death processes to solve for the joint distribution of modules within the network, which then inform each other and facilitate calculation of the entire joint distribution. We illustrate the method on a ubiquitous case in nature: linear regulatory cascades. The efficiency of the method makes possible numerical optimization of the input and regulatory parameters, revealing design properties of, e.g., the most informative cascades. We find, for threshold regulation, that a cascade of strong regulations converts a unimodal input to a bimodal output, that multimodal inputs are no more informative than bimodal inputs, and that a chain of up-regulations outperforms a chain of down-regulations. We anticipate that this numerical approach may be useful for modeling noise in a variety of small network topologies in biology.

gene regulation | noise | signaling | biological network | information theory

Transcriptional regulatory networks are composed of genes and proteins, which are often present in small numbers in the cell (1, 2), rendering deterministic models poor descriptions of the counts of protein molecules observed experimentally (3–9). Probabilistic approaches have proven necessary to account fully for the variability of molecule numbers within a homogenous population of cells. A full stochastic description of even a small regulatory network proves quite challenging. Many efforts have been made to refine simulation approaches (10–14), which are mainly based on the varying step Monte Carlo or "Gillespie" method (15, 16). Yet expanding full molecular simulations to larger systems and scanning parameter space is computationally expensive. However, the interaction of many protein and gene types makes analytical methods hard to implement. A wide class of approximations to the master equation, which describes the evolution of the probability distribution, focuses on limits of large concentrations or small switches (17–19). Approximations based on timescale separation of the steps of small signaling cascades have been successfully used to calculate escape properties (20, 21). Variational techniques have also been used to calculate approximate distributions (22).

In this article we introduce a method for calculating the steady-state distributions of chemical reactants (code is available at http://specmark.sourceforge.net). The procedure, which we call the *spectral method*, relies on exploiting the natural basis of a simpler problem from the same class. The full problem is then solved numerically as an expansion in this basis, reducing the master equation to a set of linear algebraic equations. We break up the problem into two parts: a preprocessing step, which can be solved algorithmically; and the parameter-specific step of obtaining the actual probability distributions. The spectral method allows for huge computational gains with respect to simulations.

We illustrate the spectral method for the case of regulatory cascades: downstream genes responding to concentrations of transcription factors produced by upstream genes that are linked to external cues. Cascades play an important role in a diversity of cellular processes (23–25), from decision making in development (26) to quorum sensing among cells (27). We take a coarse-grained approach, modeling each step of a cascade with a general regulatory function that depends on the copy number of the reactant at the previous step (Fig. 1). Although the method as implemented describes arbitrary regulation functions, we optimize the information transmission in the case of the most biologically simple regulation function: a discontinuous threshold, in which a species is created at a high or low rate depending on the copy count of the species directly upstream. In the next sections, we outline the spectral method and present in detail our findings regarding signaling cascades.

Method

We calculate the steady-state joint distribution for L chemical species in a cascade (Fig. 1). The approach we take involves two key observations: the master equation, being linear,* benefits from solution in terms of its eigenfunctions; and the behavior of a given species should depend only weakly on distant nodes given the proximal nodes.

The second of these observations can be illustrated succinctly by considering a three-gene cascade in which the first gene may be eliminated by marginalization. For three species obeying $s \xrightarrow{q_s} n \xrightarrow{q_n} m$ as in Fig. 1, we have the linear master equation

$$
\begin{aligned}
\dot{p}_{snm} = &\ \tilde{\rho}[g p_{(s-1)nm} - g p_{snm} + (s+1)p_{(s+1)nm} - s p_{snm}] \\
&+ q_s p_{s(n-1)m} - q_s p_{snm} + (n+1)p_{s(n+1)m} - n p_{snm} \\
&+ \rho[q_n p_{sn(m-1)} - q_n p_{snm} + (m+1)p_{sn(m+1)} - m p_{snm}]. \quad [1]
\end{aligned}
$$

Here, time is rescaled by the second gene's degradation rate, so that each gene's creation rate (g, q_s, or q_n) is normalized by its respective degradation rate; $\tilde{\rho}$ and ρ are the ratios of the first and third gene's degradation rate to the second's, respectively.

To integrate out the first species, we sum over s. We then introduce g_n, the effective regulation of n, by

$$
\sum_s q_s p_{nms} = p_{nm} \sum_s q_s p_{s|nm} \approx p_{nm} \sum_s q_s p_{s|n} \equiv g_n p_{nm}. \quad [2]
$$

Here, we have made the Markovian approximation that s is conditionally independent of m given n. Generally speaking, the probability distribution depends on all steps of the cascade. However, since there are no loops in the cascades we consider here, we assume in Eq. 2 that at steady state each species is not affected by species two or more steps away in the cascade. The validity

Evolutionary dynamics in set structured populations

Corina E. Tarnita[a], Tibor Antal[a], Hisashi Ohtsuki[b], and Martin A. Nowak[a,1]

[a]Program for Evolutionary Dynamics, Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; and [b]Department of Value and Decision Science, Tokyo Institute of Technology, Tokyo 152-8552, Japan

Evolutionary dynamics are strongly affected by population structure. The outcome of an evolutionary process in a well-mixed population can be very different from that in a structured population. We introduce a powerful method to study dynamical population structure: evolutionary set theory. The individuals of a population are distributed over sets. Individuals interact with others who are in the same set. Any 2 individuals can have several sets in common. Some sets can be empty, whereas others have many members. Interactions occur in terms of an evolutionary game. The payoff of the game is interpreted as fitness. Both the strategy and the set memberships change under evolutionary updating. Therefore, the population structure itself is a consequence of evolutionary dynamics. We construct a general mathematical approach for studying any evolutionary game in set structured populations. As a particular example, we study the evolution of cooperation and derive precise conditions for cooperators to be selected over defectors.

cooperation | game | social behavior | stochastic dynamics

Human society is organized into sets. We participate in activities or belong to institutions where we meet and interact with other people. Each person belongs to several sets. Such sets can be defined, for example, by working for a particular company, living in a specific location, going to certain restaurants, or holding memberships at clubs. There can be sets within sets. For example, the students of the same university have different majors, take different classes, and compete in different sports. These set memberships determine the structure of human society: they specify who meets whom, and they define the frequency and context of meetings between individuals.

We take a keen interest in the activities of other people and contemplate whether their success is correlated with belonging to particular sets. It is therefore natural to assume that we do not only imitate the behavior of successful individuals, but also try to adopt their set memberships. Therefore, the cultural evolutionary dynamics of human society, which are based on imitation and learning, should include updating of strategic behavior and of set memberships. In the same way as successful strategies spawn imitators, successful sets attract more members. If we allow set associations to change, then the structure of the population itself is not static, but a consequence of evolutionary dynamics.

There have been many attempts to study the effect of population structure on evolutionary and ecological dynamics. These approaches include spatial models in ecology (1–8), viscous populations (9), spatial games (10–15), and games on graphs (16–19).

We see "evolutionary set theory" as a powerful method to study evolutionary dynamics in structured populations in the context where the population structure itself is a consequence of the evolutionary process. Our primary objective is to provide a model for the cultural evolutionary dynamics of human society, but our framework is applicable to genetic evolution of animal populations. For animals, sets can denote living at certain locations or foraging at particular places. Any one individual can belong to several sets. Offspring might inherit the set memberships of their parents. Our model could also be useful for studying dispersal behavior of animals (20, 21).

Let us consider a population of N individuals distributed over M sets (Fig. 1). Individuals interact with others who belong to the same set. If 2 individuals have several sets in common, they interact several times. Interactions lead to payoff from an evolutionary game.

The payoff of the game is interpreted as fitness (22–26). We can consider any evolutionary game, but at first we study the evolution of cooperation. There are 2 strategies: cooperators, C, and defectors, D. Cooperators pay a cost, c, for the other person to receive a benefit, b. Defectors pay no cost and provide no benefit. The resulting payoff matrix represents a simplified Prisoner's Dilemma. The crucial parameter is the benefit-to-cost ratio, b/c. In a well-mixed population, where any 2 individuals interact with equal likelihood, cooperators would be outcompeted by defectors. The key question is whether dynamics on sets can induce a population structure that allows evolution of cooperation.

Individuals update stochastically in discrete time steps. Payoff determines fitness. Successful individuals are more likely to be imitated by others. An imitator picks another individual at random, but proportional to payoff, and adopts his strategy and set associations. Thus, both the strategy and the set memberships are subject to evolutionary updating. Evolutionary set theory is a dynamical graph theory: who interacts with whom changes during the evolutionary process (Fig. 2). For mathematical convenience we consider evolutionary game dynamics in a Wright–Fisher process with constant population size (27). A frequency-dependent Moran process (28) or a pairwise comparison process (29), which is more realistic for imitation dynamics among humans, give very similar results, but some aspects of the calculations become more complicated.

The inheritance of the set memberships occurs with mutation rate v: with probability $1 - v$, the imitator adopts the parental set memberships, but with probability v a random sample of new sets is chosen. Strategies are inherited subject to a mutation rate, u. Therefore, we have 2 types of mutation rates: a set mutation rate, v, and a strategy mutation rate, u. In the context of cultural evolution, our mutation rates can also be seen as "exploration rates": occasionally, we explore new strategies and new sets.

We study the mutation-selection balance of cooperators versus defectors in a population of size N distributed over M sets. In the supporting information (SI) *Appendix*, we show that cooperators are more abundant than defectors (for weak selection and large population size) if $b/c > (z - h)/(g - h)$. The term z is the average number of sets 2 randomly chosen individuals have in common. For g we pick 2 random, distinct, individuals in each state; whether they have the same strategy, we add their number of sets to the average, otherwise we add 0; g is the average of this average over the stationary distribution. For understanding h we must pick 3 individuals at random: then h is the average number of sets the first 2 individuals have in common

ized by host cells packed full of fluorescent parasites unable to egress efficiently. These grossly swollen cells often detached from the monolayer and could be found floating in the culture medium giving the appearance of a hot air balloon convention.

Fibroblast cell lines derived from CAPNS1 KO mouse embryos lack both calpain-1 and -2 activity (30), and infection of these cells with T. gondii produced the same swollen cell phenotype observed in CAPNS1 knock-down experiments (Fig. 3C). Transgenic expression of CAPNS1 in the KO mutants restores calpain-1 and -2 activity (30) and also complemented the T. gondii egress defect. Parasite tachyzoites were readily able to invade (Fig. 3D) and replicate (Fig. 3E) in WT, CAPNS1 KOs, and CAPNS1-complemented fibroblasts, demonstrating that the impact of host cell calpains on T. gondii infection is specific to egress. Plaque assays showed a ~13-fold reduction in plaque size for T. gondii in CAPNS1 mutants versus parental MEF cells or CAPNS1-complemented KOs (Fig. 3, F and G). In contrast to P. falciparum, which rarely emerge from calpain-depleted erythrocytes (Fig. 2), some T. gondii parasites did eventually manage to escape from calpain-deficient fibroblasts, yielding a small plaque phenotype.

In summary, in addition to the many roles that parasite-encoded cysteine proteases play in the biology of infection and pathogenesis (25), the apicomplexans Plasmodium falciparum and Toxoplasma gondii both exploit host cell calpains to facilitate escape from the intracellular parasitophorous vacuole and/or host plasma membrane. The precise mechanism of calpain-mediated parasite egress is unknown, but calpains play a role in remodeling of the cytoskeleton and plasma membrane during the migration of mammalian cells (31), and activated calpain-1 can degrade erythrocyte cytoskeletal proteins in vitro and during P. falciparum infection in vivo (fig. S4). The calcium responsible for calpain activation during parasite infection may be supplied through the action of a parasite-encoded perforin recently implicated in T. gondii egress (32). The parasitophorous vacuole was labeled by the calcium-specific dye Fluo-4-AM during late schizogony, and depletion of internal calcium with the membrane-permeant chelator EGTA-AM blocked parasite egress, whereas removal of calcium from the culture medium did not (fig. S5). We suggest a model in which a calcium signal triggered late during parasite infection activates host cell calpain, which relocalizes to the host plasma membrane, cleaving cytoskeletal proteins to facilitate parasite egress (fig. S6). Because parasites that fail to escape from their host cells are unable to proliferate, this suggests an intriguing strategy for anti-parasitic therapeutics.

References and Notes
1. M. Nishi, K. Hu, J. M. Murray, D. S. Roos, J. Cell Sci. 121, 1559 (2008).
2. K. Hu et al., Mol. Biol. Cell 13, 593 (2002).
3. S. Glushakova, D. Yin, T. Li, J. Zimmerberg, Curr. Biol. 15, 1645 (2005).
4. B. L. Salmon, A. Oksman, D. E. Goldberg, Proc. Natl. Acad. Sci. U.S.A. 98, 271 (2001).
5. M. W. Black, G. Arrizabalaga, J. C. Boothroyd, Mol. Cell. Biol. 20, 9399 (2000).
6. K. Nagamune et al., Nature 451, 207 (2008).
7. T. Hadley, M. Aikawa, L. H. Miller, Exp. Parasitol. 55, 306 (1983).
8. M. E. Wickham, J. G. Culvenor, A. F. Cowman, J. Biol. Chem. 278, 37658 (2003).
9. S. Arastu-Kapur et al., Nat. Chem. Biol. 4, 203 (2008).
10. S. Yeoh et al., Cell 131, 1072 (2007).
11. D. C. Greenbaum et al., Science 298, 2002 (2002).
12. S. Glushakova, J. Mazar, M. F. Hohmann-Marriott, E. Hama, J. Zimmerberg, Cell. Microbiol. 11, 95 (2009).
13. Materials and methods are available as supporting material on Science Online.
14. I. Russo, A. Oksman, B. Vaupel, D. E. Goldberg, Proc. Natl. Acad. Sci. U.S.A. 106, 1554 (2009).
15. E. M. Pasini et al., Blood 108, 791 (2006).
16. D. E. Croall, K. Ersfeld, Genome Biol. 8, 218 (2007).
17. D. E. Goll, V. F. Thompson, H. Li, W. Wei, J. Cong, Physiol. Rev. 83, 731 (2003).
18. S. C. Murphy et al., PLoS Med. 3, e528 (2006).
19. R. A. Hanna, R. L. Campbell, P. L. Davies, Nature 456, 409 (2008).
20. T. Moldoveanu, K. Gehring, D. R. Green, Nature 456, 404 (2008).
21. S. Gil-Parrado et al., Biol. Chem. 384, 395 (2003).
22. A. E. Bianco, F. L. Battye, G. V. Brown, Exp. Parasitol. 62, 275 (1986).
23. M. Hanspal, V. K. Goel, S. S. Oh, A. H. Chishti, Mol. Biochem. Parasitol. 122, 227 (2002).
24. L. Weiss, J. Johnson, W. Weidanz, Am. J. Trop. Med. Hyg. 41, 135 (1989).
25. P. J. Rosenthal, Int. J. Parasitol. 34, 1489 (2004).
26. P. Dutt et al., BMC Dev. Biol. 6, 3 (2006).
27. J. S. Arthur, J. S. Elce, C. Hegadorn, K. Williams, P. A. Greer, Mol. Cell. Biol. 20, 4474 (2000).
28. D. S. Roos, R. G. Donald, N. S. Morrissette, A. L. Moulton, Methods Cell Biol. 45, 27 (1994).
29. K. A. Joiner, D. S. Roos, J. Cell Biol. 157, 557 (2002).
30. N. Dourdin et al., J. Biol. Chem. 276, 48382 (2001).
31. A. Huttenlocher et al., J. Biol. Chem. 272, 32719 (1997).
32. B. F. C. Kafsack et al., Science 323, 530 (2009); published online 18 December 2008 (10.1126/science.1165740).
33. We thank R. W. Doms, M. Marti, and M. Klemba for critical discussions; the Penn Proteomics Core for mass spectrometry; and M. A. Lampson for help with imaging. P. falciparum expressing GFP were provided by O. S. Harb. P.H.D. and D.P.B. are funded by National Research Service Awards, and D.S.R. is an Ellison Medical Foundation Senior Scholar in Global Infectious Disease, supported by grants from NIH. D.C.G. was supported by the Ritter Foundation, the Penn Genome Frontiers Institute, and the Penn Institute for Translational Medicine and Therapeutics.

Human Induced Pluripotent Stem Cells Free of Vector and Transgene Sequences

Junying Yu,[1,2,3]* Kejin Hu,[3] Kim Smuga-Otto,[1,2,3] Shulan Tian,[1,2] Ron Stewart,[1,2] Igor I. Slukvin,[3,4] James A. Thomson[1,2,3,5]*

Reprogramming differentiated human cells to induced pluripotent stem (iPS) cells has applications in basic biology, drug development, and transplantation. Human iPS cell derivation previously required vectors that integrate into the genome, which can create mutations and limit the utility of the cells in both research and clinical applications. We describe the derivation of human iPS cells with the use of nonintegrating episomal vectors. After removal of the episome, iPS cells completely free of vector and transgene sequences are derived that are similar to human embryonic stem (ES) cells in proliferative and developmental potential. These results demonstrate that reprogramming human somatic cells does not require genomic integration or the continued presence of exogenous reprogramming factors and removes one obstacle to the clinical application of human iPS cells.

The proliferative and developmental potential of both human embryonic stem (ES) cells and human induced pluripotent stem (iPS) cells offers unprecedented access to the differentiated cells that make up the human body (1–3). In addition, iPS cells can be derived with a specific desired genetic background, including patient-specific iPS cells for disease models and for transplantation therapies, without the problems associated with immune rejection. Reprogramming of both mouse and human somatic cells into iPS cells has been achieved

by expressing combinations of factors such as OCT4, SOX2, c-Myc, KLF4, NANOG, and LIN28 (2–4). Initial methods used to derive human iPS cells used viral vectors, in which both the vector backbone and transgenes are permanently integrated into the genome (2, 3). Such vectors can produce insertional mutations that interfere with the normal function of iPS cell derivatives, and residual transgene expression can influence differentiation into specific lineages (2), or even result in tumorigenesis (5). Vector integration–free mouse iPS cells have been derived from

difference observed between these values may imply a breakdown of electron-hole symmetry because of many-body effects (8, 30). However, some contribution to the electron-hole asymmetry observed in Fig. 2C could also be caused by the screening of the tip electric field (band bending), requiring a small correction to the energy scale in Fig. 2C.

Spatial variation of the LL energies can be used to map fluctuations of the local potential (31), as we show for the $n = 0$ LL (Fig. 3). This level, composed of electron and hole carriers, occurs exactly at the Dirac point, as a direct consequence of the chiral solutions of the Dirac equation that describes graphene's low-energy electronic structure. Figure 3, A and B, shows a topographic image and corresponding spatial map of the lowest LL$_n$ energies (vertical) and dI/dV intensities (color scale) for $n = 0$, -1, -2, and -3 along the line marked in Fig. 3A. The average position of E_D was 55.2 meV above E_F, with a SD of ± 1.9 meV (Fig. 3C) (20). By far, the largest variation (in this image and generally in our measurements) corresponded to a subsurface rotational domain boundary that occurs in the center of the image [the top graphene layer is atomically continuous over the boundary (fig. S3)]. Away from such boundaries, the spatial fluctuations were much smaller: ≈ 0.5 meV, as seen for the region from 250 to 400 nm in Fig. 3. The Dirac point energy map showed an extremely smooth potential (hence, small carrier density fluctuations) for this low-doped graphene sheet, in contrast to the electron and hole puddles observed for exfoliated graphene on SiO$_2$ substrates (10). In surveying the sample, a variation of $\sim\pm 25$ meV in E_D was observed over distances of many tens of micrometers (see E_D differences between Figs. 2 and 3). The larger density fluctuations on SiO$_2$ substrates apparently result from charged impurities in the SiO$_2$ substrate. The smooth charge/potential contour in epitaxial graphene could be the result of screening of the interface potential fluctuations by the graphene multilayer, and the crystalline SiC substrate may be more homogenous than the amorphous SiO$_2$ substrate with respect to trapped charges.

The dI/dV spectra in Fig. 2 show a direct measurement of graphene magnetic quantization expected for massless Dirac fermions. This result implies that the topmost layer of epitaxial graphene closely approximates an isolated sheet of graphene. We attribute this isolation to the presence of rotational stacking faults between the graphene layers in epitaxial graphene grown on the carbon face of SiC, which effectively decouples the electronic structure of the layers (9). A variety of rotational stacking angles were found in STM topographic images of the surface (fig. S2) (13). Slight rotations of one layer with respect to the next create moiré super periods superimposed on the atomic lattice (fig. S2) (9, 13). The ≈ 0.02-nm peak-to-peak height modulation originates from periodically varying the alignment of top-layer atoms with those below, but the exact source of image

contrast is still a subject of debate (32). Our survey of the carbon-face grown sample showed moiré patterns of various periods in almost every location examined, with spectra similar to those seen in Fig. 2. We expect this structure of multilayer epitaxial graphene to be important for future studies of Dirac point physics in graphene.

Note added in proof: A new publication reports the observation of graphene LLs in STS measurements over a graphene flake on graphite (33).

References and Notes

1. C. Berger et al., *J. Phys. Chem. B* **108**, 19912 (2004).
2. C. Berger et al., *Science* **312**, 1191 (2006); published online 12 April 2006 (10.1126/science.1125925).
3. K. S. Novoselov et al., *Nature* **438**, 197 (2005).
4. Y. Zhang, Y. W. Tan, H. L. Stormer, P. Kim, *Nature* **438**, 201 (2005).
5. T. Matsui et al., *Phys. Rev. Lett.* **94**, 226403 (2005).
6. G. Li, E. Y. Andrei, *Nat. Phys.* **3**, 623 (2007).
7. M. L. Sadowski, G. Martinez, M. Potemski, C. Berger, W. A. de Heer, *Phys. Rev. Lett.* **97**, 266405 (2006).
8. R. S. Deacon, K. C. Chuang, R. J. Nicholas, K. S. Novoselov, A. K. Geim, *Phys. Rev. B* **76**, 081406R (2007).
9. J. Hass et al., *Phys. Rev. Lett.* **100**, 125504 (2008).
10. J. Martin et al., *Nat. Phys.* **4**, 144 (2008).
11. E. Rossi, S. Das Sarma, *Phys. Rev. Lett.* **101**, 166803 (2008).
12. W. A. de Heer et al., *Solid State Commun.* **143**, 92 (2007).
13. Additional text and data are available on *Science* Online.
14. C. Toke, P. E. Lammert, V. H. Crespi, J. K. Jain, *Phys. Rev. B* **74**, 235417 (2006).
15. M. Arikawa, Y. Hatsugai, H. Aoki, *Phys. Rev. B* **78**, 205401 (2008).
16. N. Ashcroft, N. Mermin, *Solid State Physics* (Brooks Cole, London, 1976).
17. V. P. Gusynin, S. G. Sharapov, *Phys. Rev. B* **71**, 125124 (2005).
18. S. G. Sharapov, V. P. Gusynin, H. Beck, *Phys. Rev. B* **69**, 075104 (2004).
19. I. A. Luk'yanchuk, Y. Kopelevich, *Phys. Rev. Lett.* **93**, 166402 (2004).
20. All uncertainties reported represent 1 SD in the measured quantity.
21. E. Rollings et al., *J. Phys. Chem. Solids* **67**, 2172 (2006).
22. A. Bostwick, T. Ohta, T. Seyller, K. Horn, E. Rotenberg, *Nat. Phys.* **3**, 36 (2007).
23. M. Morgenstern, J. Klijn, C. Meyer, R. Wiesendanger, *Phys. Rev. Lett.* **90**, 056804 (2003).
24. Y. Zhang et al., *Phys. Rev. Lett.* **96**, 136806 (2006).
25. K. I. Bolotin et al., *Solid State Commun.* **146**, 351 (2008).
26. M. Orlita et al., *Phys. Rev. Lett.* **101**, 267601 (2008).
27. E. A. Henriksen et al., *Phys. Rev. Lett.* **100**, 087403 (2008).
28. G. M. Rutter, J. N. Crain, N. P. Guisinger, P. N. First, J. A. Stroscio, *J. Vac. Sci. Technol. A* **26**, 938 (2008).
29. J. Yan, E. A. Henriksen, P. Kim, A. Pinczuk, *Phys. Rev. Lett.* **101**, 136804 (2008).
30. P. E. Trevisanutto, C. Giorgetti, L. Reining, M. Ladisa, V. Olevano, *Phys. Rev. Lett.* **101**, 226405 (2008).
31. M. Morgenstern, C. Wittneven, R. Dombrowski, R. Wiesendanger, *Phys. Rev. Lett.* **84**, 5588 (2000).
32. W. T. Pong, C. Durkan, *J. Phys. D Appl. Phys.* **38**, R329 (2005).
33. G. Li, A. Luican, E. Y. Andrei, *Phys. Rev. Lett.* **102**, 176804 (2009).
34. We thank A. MacDonald, H. Min, M. Stiles, and the NIST graphene team for valuable comments and discussions and C. Berger, N. Sharma, M. Sprinkle, S. Blankenship, A. Band, and F. Hess for their technical contributions to this work. Portions of this work were supported by NSF (grant ECCS-0804908), the Semiconductor Research Corporation Nanoelectronics Research Initiative (INDEX program), and the W. M. Keck Foundation. Graphene production facilities were developed under NSF grant ECCS-0521041.

Direct Detection of Abortive RNA Transcripts in Vivo

Seth R. Goldman,[1] Richard H. Ebright,[2] Bryce E. Nickels[1]*

During transcription initiation in vitro, prokaryotic and eukaryotic RNA polymerase (RNAP) can engage in abortive initiation—the synthesis and release of short (2 to 15 nucleotides) RNA transcripts—before productive initiation. It has not been known whether abortive initiation occurs in vivo. Using hybridization with locked nucleic acid probes, we directly detected abortive transcripts in bacteria. In addition, we show that in vivo abortive initiation shows characteristics of in vitro abortive initiation: Abortive initiation increases upon stabilizing interactions between RNAP and either promoter DNA or sigma factor, and also upon deleting elongation factor GreA. Abortive transcripts may have functional roles in regulating gene expression in vivo.

During transcription, RNA polymerase (RNAP) synthesizes the first ~ 8 to 15 nucleotides (nt) of RNA as an RNAP-promoter initial transcribing complex (1–3) [using a "scrunching" mechanism (4)]. Upon synthesis of an RNA transcript with a threshold length of ~ 8 to 15 nt, RNAP breaks its interactions with promoter

DNA, escapes the promoter, and enters into processive synthesis of RNA as an RNAP-DNA transcription elongation complex (1–3) [using a "stepping" mechanism (5)]. In transcription reactions in vitro, the RNAP-promoter initial transcribing complex can engage in tens to hundreds of cycles of synthesis and release of short RNA transcripts (abortive initiation) (1–3, 6–8). Abortive initiation competes with productive initiation in vitro and, as such, is a critical determinant of promoter strength and a target of transcription regulation in vitro (1–3, 7–13). It has been proposed that abortive initiation likewise occurs in vivo

[1]Department of Genetics and Waksman Institute, Rutgers University, Piscataway, NJ 08854, USA. [2]Department of Chemistry, Waksman Institute, and Howard Hughes Medical Institute, Rutgers University, Piscataway, NJ 08854, USA.

*To whom correspondence should be addressed. E-mail: bnickels@waksman.rutgers.edu

editing levels. The enlarged set of nonrepetitive RNA editing targets may help unravel rules of RNA editing in human diseases and behavior.

References and Notes

1. B. L. Bass, *Annu. Rev. Biochem.* **71**, 817 (2002).
2. K. Nishikura, *Nat. Rev. Mol. Cell Biol.* **7**, 919 (2006).
3. M. Higuchi *et al.*, *Nature* **406**, 78 (2000).
4. R. Brusa *et al.*, *Science* **270**, 1677 (1995).
5. M. Singh *et al.*, *J. Biol. Chem.* **282**, 22448 (2007).
6. M. J. Palladino, L. P. Keegan, M. A. O'Connell, R. A. Reenan, *Cell* **102**, 437 (2000).
7. H. Lomeli *et al.*, *Science* **266**, 1709 (1994).
8. S. Maas, Y. Kawahara, K. M. Tamburro, K. Nishikura, *RNA Biol.* **3**, 1 (2006).
9. J. Shendure, H. Ji, *Nat. Biotechnol.* **26**, 1135 (2008).
10. Materials and methods are available as supporting material on *Science* Online.
11. E. Y. Levanon *et al.*, *Nucleic Acids Res.* **33**, 1162 (2005).
12. B. Hoopengardner, T. Bhalla, C. Staber, R. Reenan, *Science* **301**, 832 (2003).
13. W. M. Gommans *et al.*, *RNA* **14**, 2074 (2008).
14. D. R. Clutterbuck, A. Leroy, M. A. O'Connell, C. A. Semple, *Bioinformatics* **21**, 2590 (2005).
15. J. Ohlson, J. S. Pedersen, D. Haussler, M. Ohman, *RNA* **13**, 698 (2007).
16. G. J. Porreca *et al.*, *Nat. Methods* **4**, 931 (2007).
17. A. G. Polson, B. L. Bass, *EMBO J.* **13**, 5701 (1994).
18. K. Nishikura *et al.*, *EMBO J.* **10**, 3523 (1991).
19. K. A. Lehmann, B. L. Bass, *Biochemistry* **39**, 12875 (2000).
20. A. Athanasiadis, A. Rich, S. Maas, *PLoS Biol.* **2**, e391 (2004).
21. M. Blow, P. A. Futreal, R. Wooster, M. R. Stratton, *Genome Res.* **14**, 2379 (2004).
22. E. Eisenberg *et al.*, *Trends Genet.* **21**, 77 (2005).
23. D. D. Kim *et al.*, *Genome Res.* **14**, 1719 (2004).
24. E. Y. Levanon *et al.*, *Nat. Biotechnol.* **22**, 1001 (2004).
25. N. Tian, X. Wu, Y. Zhang, Y. Jin, *RNA* **14**, 211 (2008).
26. C. M. Burns *et al.*, *Nature* **387**, 303 (1997).
27. M. Kohler, N. Burnashev, B. Sakmann, P. H. Seeburg, *Neuron* **10**, 491 (1993).
28. B. Sommer, M. Kohler, R. Sprengel, P. H. Seeburg, *Cell* **67**, 11 (1991).
29. T. R. Mercer *et al.*, *Neuroscientist* **14**, 434 (2008).
30. We thank R. Emeson, M. P. Ball, and F. Isaacs for critical reading of the manuscript; P. Wang and Z. Liu (BioChain Institute) for helping collect human samples; M. Higuchi and P. Seeburg for providing *ADAR2*$^{-/-}$ mouse brain cDNA; Harvard Biopolymers Facility for help with Illumina sequencing; and A. Ahlford, H. Ebling, and J. Santosuosso for assistance with Sanger sequencing. E.Y.L. was supported by the Machiah foundation. Funding came from National Human Genome Research Institute Centers of Excellence in Genomic Science grant to G.M.C. The Illumina sequencing data are deposited at the National Center for Biotechnology Information Short Read Archive under accession number SRA008181.

Unstable Tandem Repeats in Promoters Confer Transcriptional Evolvability

Marcelo D. Vinces,[1,2,3]* Matthieu Legendre,[1,4]* Marina Caldara,[1] Masaki Hagihara,[5] Kevin J. Verstrepen[1,2,3]†

Relative to most regions of the genome, tandemly repeated DNA sequences display a greater propensity to mutate. A search for tandem repeats in the *Saccharomyces cerevisiae* genome revealed that the nucleosome-free region directly upstream of genes (the promoter region) is enriched in repeats. As many as 25% of all gene promoters contain tandem repeat sequences. Genes driven by these repeat-containing promoters show significantly higher rates of transcriptional divergence. Variations in repeat length result in changes in expression and local nucleosome positioning. Tandem repeats are variable elements in promoters that may facilitate evolutionary tuning of gene expression by affecting local chromatin structure.

The genomes of most organisms are not uniformly prone to change because they contain hotspots for mutating events. An abundant class of sequences that mutate at higher frequencies than the surrounding genome is composed of tandem repeats (TRs, also known as satellite DNA), DNA sequences repeated adjacent to one another in a head-to-tail manner (*1*). Errors during replication make TRs unstable, generating changes in the number of repeat units that are 100 to 10,000 times more frequent than

point mutations (*2*). Variable TRs are often dismissed as nonfunctional "junk" DNA. However, some TRs located within coding regions (exons) have demonstrable functional roles. For example, TR copy numbers in genes such as *FLO1* in *Saccharomyces cerevisiae* generate plasticity in adherence to substrates (*3*). In canines, variable repeats located in *Alx-4* and *Runx-2* confer variability to skeletal morphology, which may have facilitated the diversification of domestic dogs bred by humans (*4*). Thus, repeats located in coding regions may increase the evolvability of proteins.

There is also evidence that repeats influence expression of certain genes (*5–7*). To investigate the involvement of TRs in gene expression variation, we first mapped and classified all repeats in the S288C yeast genome (*8*) (data set S1). TRs are enriched in yeast promoters (table S1). Of the ~5700 promoters in the genome, 25% (1455) contain at least one TR. Many TRs in promoters consist of short, A/T-rich sequences (table S2, fig. S1, and data set S2). Comparison of orthologous regions in genomes of different *S. cerevisiae* strains showed that many of the TRs are variable (data set S1). For example, 24.1% of

orthologous TR loci in promoters differ in the number of repeat units between the two fully sequenced strains, S288C and RM11 (*8*). To confirm this, we sequenced 33 randomly chosen promoter repeats in seven *S. cerevisiae* genomes (Fig. 1A, figs. S2 and S3, and data set S3). Twenty-five of the 33 TRs differed in repeat units in at least one of the seven strains. The repeat variation frequency is 40-fold higher than the frequency of insertions and deletions (indels) and of point mutations in the surrounding nonrepetitive sequence ($P < 10^{-15}$) (figs. S2 and S3).

To determine whether promoter TR variation affects gene expression, we compared repeat variablity to expression divergence (ED), which represents how fast the transcriptional activity of each gene evolves (*9–11*). Promoters containing TRs showed significantly ($P < 1.75 \times 10^{-4}$) higher amounts of ED than did promoters lacking TRs when comparing yeast species (*S. cerevisiae*, *S. paradoxus*, *S. mikatae*, and *S. kudriavzevii*) (Fig. 1, B to D, and fig. S4A) and *S. cerevisiae* strains (S288C and RM11) (Fig. 1, E to G, and fig. S4, B and C). This difference was independent of factors known to affect transcriptional divergence, for example, the presence of TATA boxes (fig. S5). Only promoters containing variable numbers of repeat units between strains or species showed the elevated ED (Fig. 1, D and G). Furthermore, when variable TRs were binned into variable and highly variable (10% most variable) groups, highly variable repeats displayed even higher ED. Hence, ED correlates not merely with TRs in promoters but more specifically with repeat number variation.

To directly test whether changes in promoter TRs affect transcriptional activity, we varied the TR repeat number in the promoters of yeast genes *YHB1*, *MET3*, and *SDT1* (Fig. 2 and fig. S6A). For each construct, expression increased as the length of the TR increased from zero, until a certain size was reached, after which expression dropped off. To determine whether natural variation between strains corresponded to similar changes in gene expression, we cloned promoters of several strains

[1]FAS Center for Systems Biology, Harvard University, 52 Oxford Street, Cambridge, MA 02138, USA. [2]Laboratory for Systems Biology, Flanders Institute for Biotechnology (VIB), Katholieke Universiteit Leuven (K.U. Leuven), B-3001 Heverlee, Belgium. [3]Genetics and Genomics Group, Centre of Microbial and Plant Genetics (CMPG), K.U. Leuven, Gaston Geenslaan 1, B-3001 Leuven (Heverlee), Belgium. [4]Structural and Genomic Information Laboratory, CNRS-UPR 2589, IFR-88, Université de la Méditerranée Parc Scientifique de Luminy, Avenue de Luminy, FR-13288 Marseille, France. [5]The Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, 567-0047, Japan.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: Kevin.Verstrepen@biw.vib-kuleuven.be

Synthetic Gene Networks That Count

Ari E. Friedland,[1]* Timothy K. Lu,[1,2]* Xiao Wang,[1] David Shi,[1]
George Church,[2,3] James J. Collins[1]†

Synthetic gene networks can be constructed to emulate digital circuits and devices, giving one the ability to program and design cells with some of the principles of modern computing, such as counting. A cellular counter would enable complex synthetic programming and a variety of biotechnology applications. Here, we report two complementary synthetic genetic counters in *Escherichia coli* that can count up to three induction events: the first, a riboregulated transcriptional cascade, and the second, a recombinase-based cascade of memory units. These modular devices permit counting of varied user-defined inputs over a range of frequencies and can be expanded to count higher numbers.

A counter is a key component in digital circuits and computing that retains memory of events or objects, representing each number of such as a distinct state. Counters would also be useful in cells, which often must have accurate accounting of tightly controlled processes or biomolecules to effectively maintain metabolism and growth. Counting mechanisms have been reportedly found in telomere length regulation (*1*, *2*) and cell aggregation (*3*). These system behaviors appear to be the result of a threshold effect in which some critical molecule number or density must be reached for the observed phenotypic change.

In this study, we first developed a counter, termed the riboregulated transcriptional cascade (RTC) counter, which is based on a transcriptional cascade with additional translational regulation. Two such cascades are illustrated in Fig. 1, A and C that can count up to two and three, respectively (hence, the designations RTC two-counter and RTC three-counter). For the RTC two-counter, the constitutive promoter $P_{LtetO-1}$ drives transcription of T7 RNA polymerase (RNAP), whose protein binds the T7 promoter and transcribes the downstream gene, in this case, green fluorescent protein (GFP). Both genes are additionally regulated by riboregulators (*4*), whose cis and trans

elements silence and activate posttranscriptional gene expression, respectively. The cis-repressor sequence [cr in Fig. 1] is placed between the transcription start site and the ribosome-binding site (RBS), and its complementarity with the RBS causes a stem-loop structure to form upon transcription. This secondary structure prevents binding of the 30*S* ribosomal subunit to the RBS, which inhibits translation. A short, transactivating, noncoding RNA (taRNA), driven by the arabinose promoter P_{BAD}, binds to the cis repressor in trans, which relieves RBS repression and allows translation. With this riboregulation, each node (i.e., gene) in the cascade requires both independent transcription and translation for protein expression. This cascade is able to count brief arabinose pulses [for pulse definition, see (*5*)] by expressing a different protein in response to each pulse (Fig. 1A). With cis-repressed T7 RNAP

Fig. 1. The RTC two-counter and RTC three-counter construct designs and results. (**A**) The RTC two-counter is a transcriptional cascade with two nodes. Shown at the bottom are expected expression profiles after zero, one, and two arabinose (Ara) pulses. (**B**) Mean fluorescence of three replicates of RTC two-counter cell populations over time, measured by a flow cytometer. Shaded areas represent arabinose pulse duration. (**C**) The RTC three-counter is a transcriptional cascade with three nodes. Shown at the bottom are expected expression profiles after zero, one, two, and three arabinose pulses. (**D**) Mean fluorescence of three replicates of RTC three-counter cell populations over time, measured by a flow cytometer. Shaded areas represent arabinose pulse duration.
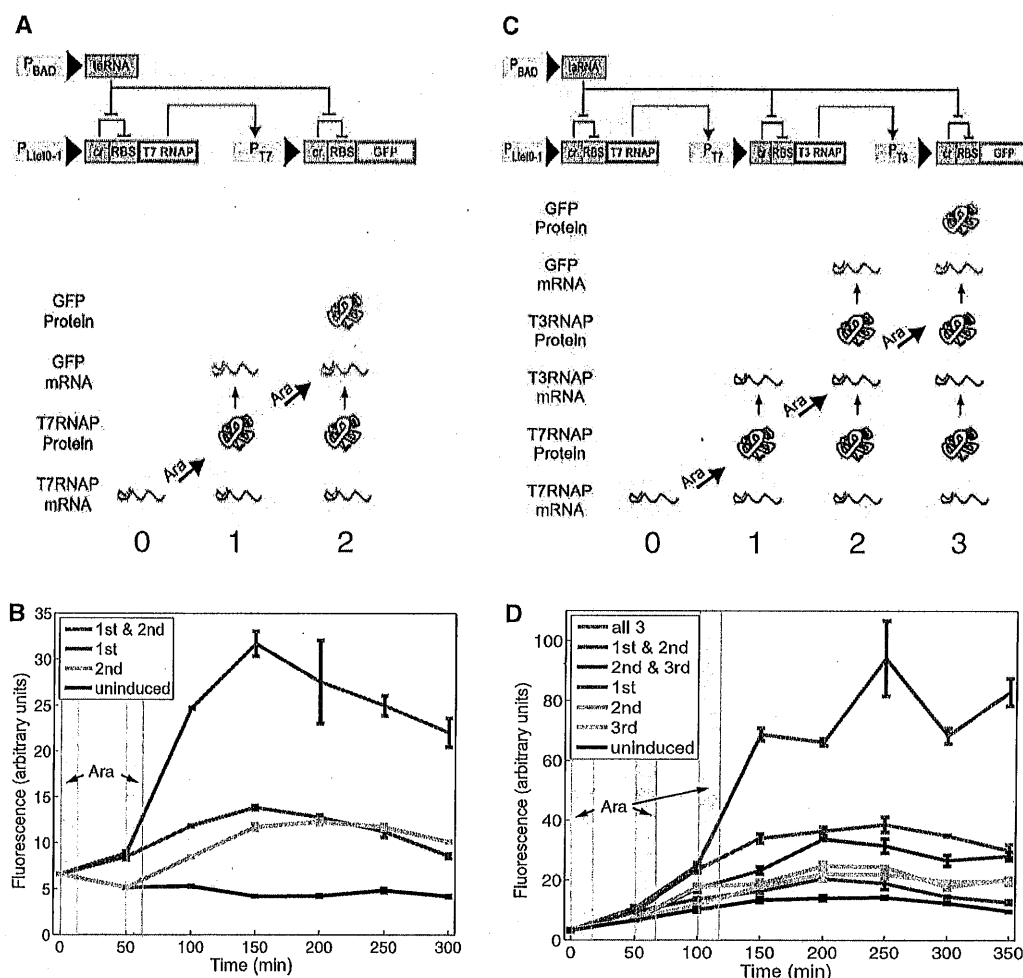
and the size and shape of the tRNA stems. Finally, a positively charged groove extends from the THUMP domain to the active site of the enzyme. The width of the groove formed between the two monomers could provide a snug fit for the acceptor stem; the length of the groove ensures that binding of the 3′-CCA at the THUMP domains places U8 adjacent to the active site of the CDD.

It is possible that C8 is the result of genetic drift. Alternatively, C8 may be beneficial for *M. kandleri* at the level of the tRNA gene, the primary tRNA transcripts, or the maturation steps that involve folding and modification of the tRNA. Because *M. kandleri* grows at extreme temperatures (up to 110°C), C8 may stabilize the genome at the tRNA gene or aid in tertiary folding and modification events of RNA molecules at such extreme conditions. Perhaps the 4-thiolation of U8 to S^4U8 mediated by ThiI might be regulated by CDAT8, as C8 would not be a ThiI substrate. A C8 may also be beneficial at the level of DNA, as a T8C mutation might prevent the interaction of tRNA genes with mobile genetic elements [e.g., (23, 24)] at this otherwise fixed and convenient target.

References and Notes

1. B. Teng, C. F. Burant, N. O. Davidson, *Science* 260, 1816 (1993).
2. A. P. Gerber, W. Keller, *Science* 286, 1146 (1999).
3. S. Maas, A. P. Gerber, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* 96, 8895 (1999).
4. J. Wolf, A. P. Gerber, W. Keller, *EMBO J.* 21, 3841 (2002).
5. H. Grosjean, F. Constantinesco, D. Foiret, N. Benachenhou, *Nucleic Acids Res.* 23, 4312 (1995).
6. S. Steinhauser, S. Beckert, I. Capesius, O. Malek, V. Knoop, *J. Mol. Evol.* 48, 303 (1999).
7. G. V. Börner, M. Mörl, A. Janke, S. Pääbo, *EMBO J.* 15, 5949 (1996).
8. M. A. Rubio *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 104, 7821 (2007).
9. J. R. Palmer, T. Baltrus, J. N. Reeve, C. J. Daniels, *Biochim. Biophys. Acta* 1132, 315 (1992).
10. A. I. Slesarev *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 99, 4644 (2002).
11. E. Westhof, P. Dumas, D. Moras, *J. Mol. Biol.* 184, 119 (1985).
12. C. N. Jones, C. I. Jones, W. D. Graham, P. F. Agris, L. L. Spremulli, *J. Biol. Chem.* 283, 34445 (2008).
13. T. Sterner, M. Jansen, Y. M. Hou, *RNA* 1, 841 (1995).
14. A. A. Smith, D. C. Carlow, R. Wolfenden, S. A. Short, *Biochemistry* 33, 6468 (1994).
15. L. Aravind, E. V. Koonin, *Trends Biochem. Sci.* 26, 215 (2001).
16. D. G. Waterman, M. Ortiz-Lombardia, M. J. Fogg, E. V. Koonin, A. A. Antson, *J. Mol. Biol.* 356, 97 (2006).
17. See supporting material on *Science* Online.
18. C. Prochnow, R. Bransteitter, M. G. Klein, M. F. Goodman, X. S. Chen, *Nature* 445, 447 (2007).
19. K. M. Chen *et al.*, *Nature* 452, 116 (2008).
20. L. G. Holden *et al.*, *Nature* 456, 121 (2008).
21. C. J. McCleverty, M. Hornsby, G. Spraggon, A. Kreusch, *J. Mol. Biol.* 373, 1243 (2007).
22. C. T. Lauhon, W. M. Erwin, G. N. Ton, *J. Biol. Chem.* 279, 23022 (2004).
23. R. Marschalek, T. Brechner, E. Amon-Bohm, T. Dingermann, *Science* 244, 1493 (1989).
24. Q. She, K. Brugger, L. Chen, *Res. Microbiol.* 153, 325 (2002).
25. We thank J. Yuan, P. O'Donoghue, and R. L. Sherrer for help and encouragement; K. O. Stetter for a gift of *M. kandleri* cells; and the staff at the Advanced Photon Source (beamline 24-ID), the National Synchrotron Light Source (beamlines X25 and X6A), and the Center for Structural Biology at Yale University. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (code 3G8Q). Supported by NIH grants GM22854 (D.S.) and AI078831 (Y.X.).

A Yeast Hybrid Provides Insight into the Evolution of Gene Expression Regulation

Itay Tirosh,[1] Sharon Reikhav,[1,2] Avraham A. Levy,[2]* Naama Barkai[1]*

During evolution, novel phenotypes emerge through changes in gene expression, but the genetic basis is poorly understood. We compared the allele-specific expression of two yeast species and their hybrid, which allowed us to distinguish changes in regulatory sequences of the gene itself (cis) from changes in upstream regulatory factors (trans). Expression divergence between species was generally due to changes in cis. Divergence in trans reflected a differential response to the environment and explained the tendency of certain genes to diverge rapidly. Hybrid-specific expression, deviating from the parental range, occurred through novel cis-trans interactions or, more often, through modified trans regulation associated with environmental sensing. These results provide insights on the regulatory changes in cis and trans during the divergence of species and upon hybridization.

Over the past years, extensive variations in gene expression were identified between closely related species (1, 2). The genetic basis of most differences, however, remains unknown. Expression divergence of a specific gene can result from mutations in its regulatory sequences, such as promoter elements (cis effects), or from mutations elsewhere in the genome that alter the abundance or activity of upstream regulators (trans effects). Distinguishing the relative contribution of cis and trans effects to the diver-

gence of gene expression is an essential step in elucidating its genetic basis.

When comparing different strains of the same species, cis and trans effects can be approximated by using linkage analysis (3–5). Alternatively, for comparison of different species that produce viable hybrids, cis versus trans effects can be distinguished by using the interspecific hybrid (6). Within the hybrid, both alleles of each gene are exposed to the same nuclear environment. Thus, differences in the expression of the two hybrid alleles reflect cis effects, whereas expression differences between the parental genes that disappear in the hybrid reflect trans effects. This approach was previously used to analyze several dozen *Drosophila* (6, 7), yeast (8), and maize (9) genes and recently also a single mammalian

chromosome (10), but has not yet been applied to a whole genome.

We designed a microarray that enables the measuring of allele-specific expression in a hybrid of *Saccharomyces cerevisiae* and *S. paradoxus*, two yeast species that diverged ~5 million years ago (fig. S1). Hybridization of genomic DNA from the two parental species verified the specificity of the microarray (Fig. 1A) and confirmed the lack of major variations in copy number. Biological repeats were highly correlated [Pearson correlation $(r) \sim 0.98$], significantly more so than the correlation in expression between species $(r \sim 0.85)$ or between the corresponding alleles within the hybrid $(r \sim 0.9)$ (Fig. 1B).

Using the array, we measured the expression profiles of the two parental species and their hybrid under four different growth conditions and quantified the relative contribution of cis and trans effects to the interspecies divergence (Fig. 1C). As expected, cis (but not trans) effects were correlated with sequence divergence at promoters and regulatory elements (fig. S3), whereas trans (but not cis) effects were enriched with genes whose expression was altered upon deletion of transcription or chromatin regulators (fig. S4 and S5). A small portion of the trans effects (<1%) were enriched within contiguous chromosomal regions that display correlated expression divergence, possibly indicating epigenetic effects (fig. S6).

The relative contribution of cis and trans effects to the divergence of gene expression varied under the different environmental conditions (Fig. 1D). In three of the conditions [rich media, heat shock, and the addition of an Rpd3p inhibitor, trichostatin A (TSA)], cis effects dominated, which is consistent with previous reports

[1]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. [2]Department of Plant Sciences, Weizmann Institute of Science, Rehovot, Israel.

*To whom correspondence should be addressed. E-mail: avi.levy@weizmann.ac.il (A.A.L.); naama.barkai@weizmann.ac.il (N.B.)