

## SYSTEMS BIOLOGY

# Excavating the Functional Landscape of Bacterial Cells

Howard Ochman and Rahul Raghavan

There is a poster that charts the metabolic pathways of the model bacterium *Escherichia coli*, and our laboratory uses it as a reference to convert each new bacterial genome sequence into an atlas of encoded functions. This is a particularly satisfying endeavor and has been best applied to those host-dependent bacteria whose highly reduced genomes contain a subset of the genes in *E. coli*: If a particular gene or pathway is eroded or absent, it is assumed that the bacterium is deficient in that activity and no longer has a need for the trait in its current circumstances. However, this linear mapping of genes to function rarely considers how a cell actually accomplishes the processes it has retained. Genes are viewed simply as performing a specified function despite the many internal and external factors that might affect their implementation. Three papers in this issue—by Güell *et al.* on page 1268 (1), Yus *et al.* on page 1263 (2), and Kühner *et al.* on page 1235 (3)—report features of transcriptional control and protein organization that are much more subtle and intricate than were previously considered possible in bacteria, and in many ways,

appear similar to mechanisms in eukaryotes. Using *Mycoplasma pneumoniae*, one of the causative agents of what is sometimes called “walking pneumonia,” the authors of all three studies performed comprehensive and detailed investigations on the organism’s small genome that have greatly increased our understanding of bacterial cellular systems (see the figure). Although numerous technological developments assisted these analyses, much of their beauty comes from the organism itself. Unlike the >4000-kb genomes of the most popular bacterial model systems *E. coli* (4) and *Bacillus subtilis* (5), *M. pneumo-*

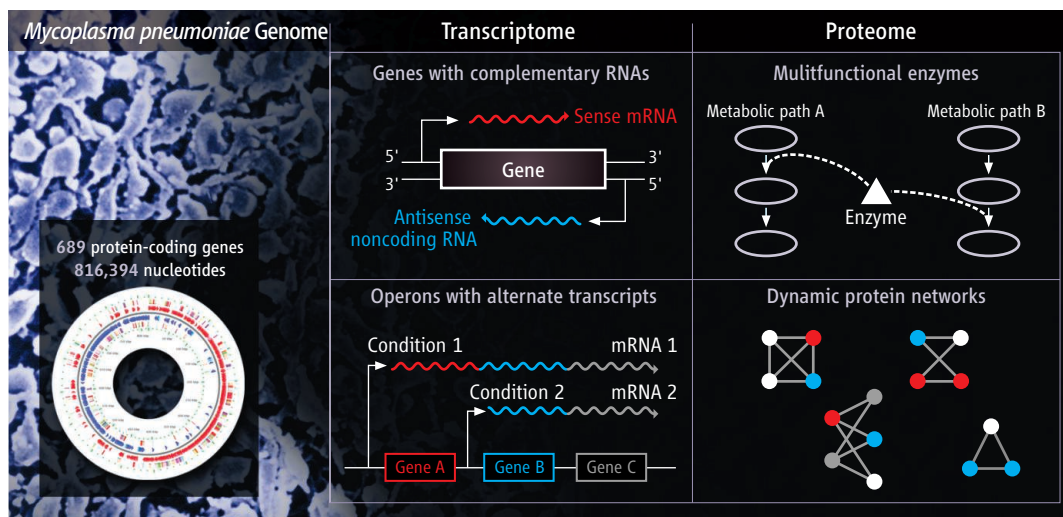
*niae* has a more tractable genome size of only 816 kb, encoding proportionately fewer genes and regulatory elements. But unlike other bacteria with diminutive genomes, the smallest of which are host-dependent symbionts of insects (6), *M. pneumoniae* can be cultivated in vitro, allowing manipulation and detailed scrutiny of external parameters on its physiology.

Güell *et al.* analyzed the expression patterns of the complete set of RNAs encoded in the *M. pneumoniae* genome, revealing a com-

plex transcriptional landscape. A large fraction of protein-coding genes have RNAs encoded on the complementary strand of genomic DNA, suggesting a double-stranded RNA-based regulation of gene expression akin to a process observed in eukaryotes (7). Genome-wide profiling of transcripts produced under various growth conditions revealed increased or decreased expression of contiguous genes (depending on the context), progressively descending levels of transcripts within operons (an operon includes several genes that are controlled by common regulatory elements), and the production of alternative transcripts from a single locus, indicating that bacterial operons are much more dynamic and versatile than previously thought. The authors have begun to decode these complicated regula-

tory signatures by assembling groups of coexpressed genes and searching for sequence motifs common to each group. Although this approach identified some novel regulatory elements, it seems likely that many remain unrecognized. Together, these findings suggest the presence of a highly structured, multifaceted regulatory machinery, which is unexpected because bacteria with small genomes contain relatively few transcription factors (8, 9).

Yus *et al.* integrated biochemical, structural, and computational information to



**Compact but complex.** Analyses of the RNAs and proteins of the bacterium *M. pneumoniae* raise questions about the complexity of regulatory mechanisms in organisms with small genomes.

reconstruct the metabolic network of *M. pneumoniae*, and then used this metabolic map to develop a culture medium containing the minimal requirements necessary for cultivating the bacterium. The metabolic pathways in *M. pneumoniae*, compared to those of bacteria with larger genomes, contain few branches or redundancies. However, its metabolic potential is augmented by the presence of numerous enzymes that perform multiple functions. And as observed in the RNA analyses of Güell *et al.*, the reduction in the metabolic network of *M. pneumoniae* is accompanied by intricate changes in the pattern of regulation, further suggesting a complex system that involves multifunctional enzymes, post-translational modifications of proteins, and chemical messengers.

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85718, USA. E-mail: hochman@email.arizona.edu

The genome-scale analysis of protein complexes in *M. pneumoniae* by Kühner *et al.* greatly expands our knowledge of protein-protein interactions within bacterial cells. Proteins often interact with one another to form functional complexes, and similar to the situation with eukaryotes (10), more than 90% of soluble proteins in *M. pneumoniae* serve as components of protein complexes. Surprisingly, the protein-interaction networks correlate poorly with genome organization and gene expression patterns—gene adjacency and coexpression were not good predictors of physically interacting proteins—again suggesting the presence of regulatory mechanisms not apparent from the compact genome structure.

How did these remarkable layers of gene regulation and the highly promiscuous behavior of proteins in *M. pneumoniae* arise? At first glance, these features may seem to be fine-tuned adaptations to the organism's current life-style, but this is not compatible with evidence for the reduced efficacy of selection that operates on the genomes of host-dependent bacteria. Bacteria that chronically associate with eukaryotic hosts undergo

bottlenecks at the time of transmission, and such reductions in long-term effective population size result in a relaxation of selection genome-wide. This instigates the accumulation and fixation of deleterious mutations in seemingly beneficial genes due to genetic drift, as well as in those genes rendered superfluous in the nutrient-rich host environment (11). In both cases, disrupted genes are eliminated by the pervasive mutational bias favoring deletions that is present in all bacteria, thereby reducing genome size (12).

As genes are lost, their roles are fulfilled by the remaining genes, much like the members of a downsized office-staff who perform tasks that previously were carried out by former co-workers. As evidence of this process, the smallest cellular genome, currently 144 kb for the insect symbiont *Hodgkinia cicadicola* (6), encodes only 15 transfer RNAs (tRNAs) to specify the 20 amino acids required to synthesize proteins. It is difficult to see how this could be adaptive, or even possible, but presumably several tRNAs must be assuming multiple roles.

The reduced genome of *M. pneumoniae* belies an underlying eukaryote-like cellular

organization replete with intricate regulatory networks and innovative pathways, revealing that there is no such a thing as a “simple” bacterium. The compound roles of individual genes and the need for additional regulatory mechanisms both may be hallmarks of reduced bacterial genomes, and the extraordinary information now available for *M. pneumoniae* sets a new standard for understanding systems-level questions about bacterial physiology and evolution.

#### References

1. M. Güell *et al.*, *Science* **326**, 1268 (2009).
2. E. Yus *et al.*, *Science* **326**, 1263 (2009).
3. S. Kühner *et al.*, *Science* **326**, 1235 (2009).
4. F. R. Blattner *et al.*, *Science* **277**, 1453 (1997).
5. F. Kunst, *Nature* **390**, 249 (1997).
6. J. P. McCutcheon, B. R. McDonald, N. A. Moran, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15394 (2009).
7. S. Katayama *et al.*, *Science* **309**, 1564 (2005).
8. R. Himmelreich *et al.*, *Nucleic Acids Res.* **24**, 4420 (1996).
9. S. Shigenobu *et al.*, *Nature* **407**, 81 (2000).
10. A. C. Gavin *et al.*, *Nature* **440**, 631 (2006).
11. C.-H. Kuo, N. A. Moran, H. Ochman, *Genome Res.* **19**, 1450 (2009).
12. C.-H. Kuo, H. Ochman, *Genome Biol. Evol.* **1**, 145 (2009).

10.1126/science.1183757

## COMPUTER SCIENCE

# What Can Virtual Worlds and Games Do for National Security?

V. S. Subrahmanian and John Dickerson

Military planners have long used war games to plan for future conflicts. Beginning in the 1950s, defense analysts began to develop computer-based models to predict the outcomes of military battles that incorporated elements of game theory. Such models were often restricted to two opposing forces, and often had a strict win-lose resolution. Today, defense analysts face situations that are more complex, not only in that conflicts may involve several opposing groups within a region, but also in that military actions are only part of an array of options available in trying to foster stable, peaceful conditions. For example, in the current conflict in Afghanistan, analysts must try to estimate how particular actions by their forces—building schools, burning

drug crops, or performing massive security sweeps—will affect interactions between the many diverse ethnic groups in the region. We discuss one approach to addressing this prediction problem in which possible outcomes are explored through computer-based virtual-world environments.

War games (1) are used to play out certain scenarios that an expert has designed. Partial information games (2) allow machine-derived models, such as stochastic opponent modeling agents (SOMA) (3), to guide the actions of players in the game based on stochastic decision rules and in the presence of partial information about the other players' situation. U.S. forces might use SOMA models to understand that a particular group might respond in one of a million ways, together with a probability distribution over those million responses.

Virtual worlds provide a software environment within which players can virtually “see”

Virtual environments based on behavioral models allow analysts to explore different outcomes of proposed actions in military conflicts.

these scenarios play out in front of them, understand the probabilities of the scenarios, understand what types of things another player might do, and explore “what ifs” with respect to opponents. Decision-makers gain simulated “experience” to guide real-world decisions by enumerating different ways that such complex interactions might play out over an extended period of time.

Underlying the output of a virtual world is a game tree, that is, a tree in which each node represents a “belief state.” A belief state is a probability distribution over states, where a state is a set of conditions that are true in a particular node. A state may include not only what a group might be thinking, but also what its offensive capabilities are, where it has forces, or whether opponents are actively engaging them.

Each outcome of a player's actions drops the game a level in a tree. For example, the U.S. might play, with a group playing at the

Department of Computer Science and Institute for Advanced Computer Studies (UMIACS), University of Maryland, College Park, MD 20742, USA. E-mail: vs@cs.umd.edu; johnd@umiacs.umd.edu