Meeting report

Why genomics is more than genomes

Jeffrey G Lawrence

Address: Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA. E-mail: jlawrenc@pitt.edu

Published: 16 November 2004

Genome ***Biology*** 2004, **5:**357

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2004/5/12/357

A report on the 2004 meeting on Molecular Genetics of Bacteria and Bacteriophages, Cold Spring Harbor, USA, 25-29 August 2004.

The meetings on the molecular genetics of bacteria and bacteriophages are the oldest and smallest of the conferences that originated at the Cold Spring Harbor Laboratory, which was the venue for this year's meeting. The toolbox of molecular genetics has expanded immensely since the first of the 'phage meetings' was held at Cold Spring Harbor in the 1950s. Then, bacteriophages were used as model systems to elucidate fundamental aspects of biology at the molecular level, including the nature of mRNA, colinearity of genes and proteins, the action of restriction enzymes and DNA ligase, roles of molecular chaperones and anti-termination. Much was understood in the absence of genomic data. Ultimately, the release of the complete genome sequence of bacteriophage λ in 1982 perhaps raised more questions about λ biology than it answered. Complete genome sequences of prokaryotes number nearly 200 at the time of writing and they offer a powerful route to understanding aspects of a microorganism's biology. Yet, as shown at the 2004 meeting, there are always new aspects of biology to be discovered that are not at all evident from a genome sequence but are crucial for its interpretation.

Successful inference of gene regulation from genome sequences has progressed steadily in the past few decades, as the binding sites for regulatory proteins have become better characterized and algorithms for finding cognate sites have improved. Yet the ability to make inferences about regulation on the basis of sequence alone is limited by incomplete knowledge of the mechanisms used by the cell. In the late 1970s, translational attenuation was discovered in both *Salmonella enterica* and *Escherichia coli*. This mechanism for regulating gene expression depends on interactions between the ribosome, RNA polymerase and structures in the mRNA that either permit or inhibit transcription. The *S. enterica his* operon and the *E. coli trp* operon provided models for the detection of leader peptides and alternative terminator/anti-terminator mRNA structures in other systems. Yet Tina Henkin (Ohio State University, Columbus, USA) has recently proposed that many Gram-positive bacteria regulate genes for amino-acid biosynthesis and tRNA synthetases using mRNA secondary structures (termed T-box RNAs) that bind directly to uncharged tRNAs, eliminating the need for the ribosome to serve as a sensing molecule. At the meeting Henkin described biochemical experiments that clearly demonstrated binding of uncharged $tRNA^{Gly}$ to the *glyQS* leader mRNA and showed that anti-termination *in vitro* is sensitive to the ratio of charged to uncharged tRNA.

Uncharged tRNA is not the only relatively small molecule that binds directly to an mRNA to control gene expression. Ali Nahvi (Yale University, New Haven, USA), who was awarded the 2004 Nat L. Sternberg prize for outstanding dissertation research at the meeting, described the mechanisms of action of some of the nine classes of riboswitches, which control more than 70 genes in *Bacillus subtilis*, one of many organisms in which they have been found. Riboswitches are 5′ untranslated mRNA leaders that bind directly to small effector molecules such as cobalamin, *S*-adenosylmethionine, purines, flavin mononucleotide, glycine, lysine, glucosamine 6-phosphate and thiamine pyrophosphate. Conformational changes in the mRNA after small-molecule binding effect both transcriptional and translational control. The central role of many of these small-molecule ligands in metabolism, as well as the protein-free nature of riboswitch mechanism of action, leads Nahvi to speculate that riboswitches may represent an evolutionarily ancient mechanism of gene regulation.

While T-boxes and riboswitches can be detected by virtue of their mRNA secondary structure in much the same way as can translational attenuators, Qi Meng (University of Illinois, Urbana-Champaign, USA) described a transcriptional attenuator that could not be predicted from first principles,

primarily because a critical sequence involved in anti-terminator formation is not present in the genome. Meng described attenuation at the *B. subtilus pyrG* gene, which encodes CTP synthetase; this gene is derepressed during cytidine starvation. The predicted transcript begins with a GGGC tetranucleotide. Depleted cytosine pools result in reiterative transcription and the synthesis of an extended poly(G) tail on the 5′ end of the mRNA; this poly(G) tract forms half of an anti-terminator which partners with a downstream mRNA sequence and permits transcription. When intracellular CTP pools are high, no reiterative transcription occurs, a terminator structure is formed in the mRNA and RNA polymerase fails to transcribe the *pyrG* gene. Alteration of the fourth nucleotide in the DNA encoding the transcript to another base allows derepression of the *pyrG* gene in response to depletion of the cognate nucleotide. These elegant experiments show that mRNA secondary structures may depend on sequences not encoded by the genome, but which are synthesized during specific cellular conditions. Only detailed knowledge of the molecular biology of *Bacillus* could allow this regulatory mechanism to be elucidated.

Large numbers of genes are controlled by the 'alarmone' ppGpp, which signals protein starvation. The ppGpp molecule has long been known to be produced by the RelA protein, but its mode of action has been elusive. Irina Artsimovich (Ohio State University) and Richard Gourse (University of Wisconsin, Madison, USA) independently reported structural biology studies, and *in vitro* and *in vivo* assays, suggesting that ppGpp is stabilized in its binding to RNA polymerase by the action of the DksA protein. Like the transcription factor GreA, DksA may extend a coiled-coil domain through the secondary pore of RNA polymerase to the active site. There, DksA aspartate residues may coordinate a magnesium ion bound to ppGpp phosphates, stabilizing the ppGpp-RNA-polymerase complex. DksA could also destabilize the binding of RNA polymerase to DNA via its interactions in the secondary pore. The concerted effects of the DksA interaction appear to increase transcription in some weakly transcribed genes (such as those encoding amino-acid biosynthetic enzymes), and decrease the transcription of others (such as those encoding structural RNAs). This mode of action explains how a single molecule can differentially affect transcription after binding to RNA polymerase.

Gene regulation can also occur post-translationally, and a novel example from bacteriophage P1 was presented by Ryland Young (Texas A&M University, College Station, USA). Here, the phage lysozyme, Lyz, is made in an inactive form not as a pro-protein but as the active-length protein, which is exported to the host bacterium's periplasm. The lysozyme amino-terminal domain becomes embedded in the cell membrane and activation of the enzyme only occurs when the P1-encoded holin protein triggers membrane depolarization, releasing the lysozyme from the lipid bilayer. After release, disulfide-bond isomerization occurs, utilizing the now-available Cys13 and freeing the catalytic Cys51 residue. So, in this case subcellular localization prevents the lysozyme from adopting an active conformation until the signal, a change in membrane potential, is received.

The genome can be considered as a static source of information that is interpreted differentially by the cell depending upon environmental conditions. But, Steven Finkel (University of Southern California, Los Angeles, USA) reported reproducible genomic rearrangements in response to long-term starvation of *E. coli*. Here, microarrays were used to quantify the amount of DNA in cells that had remained in stationary phase for up to three years. Like work reported earlier this year by Daniel Dykhuizen and Antony Dean on the adaptation of *E. coli* to different sugar sources, Finkel described how the endpoints of some genomic rearrangements that occurred during adaptation to starvation could be mapped to insertion sequences, which provided local DNA identity to promote either duplication or deletion of genome segments. This study shows once again that bacterial cells can respond to external conditions and adapt by means other than point mutation or horizontal gene transfer.

As a whole, the more than 100 short seminars and a similar number of posters presented a dizzying array of mechanisms by which the information found in microbial genomes can be dynamically employed by bacterial cells. While the understanding of microbiology imparted by genome analysis is tremendous, that knowledge is implemented through understanding of such mechanisms, and it would appear that we have only begun to unpack the toolboxes assembled by bacteria and their phages over the past 3,500 million years.