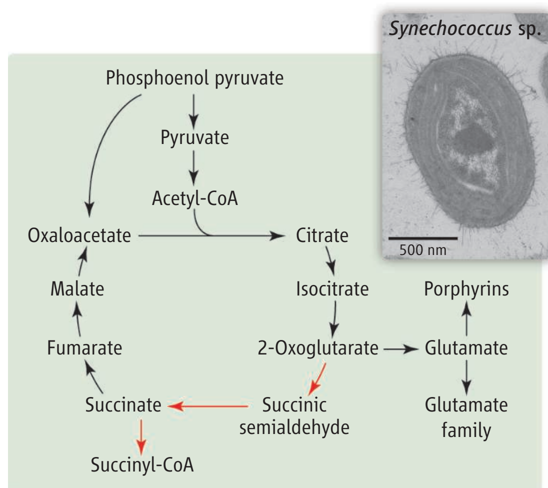


Cyanobacterial metabolism. Zhang and Bryant's experiments yield a closed TCA cycle in cyanobacteria. The red lines highlight reactions not previously known to occur in cyanobacteria. The results are based on data from the unicellular cyanobacterium *Synechococcus* sp. strain PCC 7002 and on recent genome annotations.

In the case of slow-growing microbes, this step enhances progress toward more definitive experiments in determination of whether kinetic constants, substrate specificity, and metabolite flux are consistent with a proposed physiological role in the organismal cytoplasm in which the proteins evolved.

The results will help to gain a more detailed understanding of cyanobacterial physiology. The lack of 2-oxoglutarate dehydrogenase, incomplete operation of a TCA cycle, and limited production of reductant in the dark were the first cornerstones for explaining why the cyanobacteria studied in the laboratory could grow only in the light with CO_2 as their carbon source (photoautotrophic metabolism) (*2*). It was subsequently shown that some cyanobacteria could grow in the dark (dark heterotrophic metabolism) (*6*), with sugars as their carbon source, although at rates that are slower than photoautotrophic rates by a factor of 2.5 to 5.0 (*6*–*8*). Given that some cyanobacteria do grow in the dark, an incomplete TCA cycle could not be solely responsible for the lack of dark growth by others. There must be limitations in metabolic control or other processes—not associated with a complete or incomplete TCA cycle, nor with transport (*9*)—that restrict robust dark heterotrophic growth of cyanobacteria.

The TCA cycle described by Zhang and Bryant differs from the conventional cycle in the lack of formation of succinyl-CoA as an intermediate. However, the requirement for succinate and/or succinyl-CoA in cyanobacteria is minimal. Synthesis of δ-aminolevulinic acid (the precursor of the pyrrole ring in porphyrin derivatives, heme, chlorophyll, and bilin pigments) is not from succinate and glycine, as it is in other bacteria; rather, the substrate is glutamate in glutamyl-tRNA (*10*). Nonetheless, the presence of genes encoding the α and β subunits of succinyl-CoA synthetase, previously also thought to be absent, have recently been annotated in the genomes of cyanobacteria (*11*), again excepting the marine picoplankton, although activity has yet to be demonstrated. Succinyl-CoA appears to be the end product of a side reaction of the TCA cycle in cyanobacteria (see the figure).

2-Oxoglutarate is a major precursor metabolite in cyanobacteria. It is the starting point for the biosynthesis of a wide variety of compounds in the cell, including phycobiliproteins, the main light-harvesting photosynthetic pigments in most cyanobacteria, which can constitute up to 30% of the cellular proteins (*12*). This biosynthetic demand is consistent with the 2-oxoglutarate concentration serving as a sensitive signal of carbon and nitrogen status (*13*, *14*). As Zhang and Bryant have now shown, 2-oxoglutarate can also be metabolized to succinate, adding another factor to modeling of carbon metabolic pathways (*15*) and metabolite flux. This information will be essential for precisely engineering cyanobacteria for biotechnological applications (*16*).

References and Notes

1. C. K. Pearce, N. G. Carr, *Biochem. J.* **105**, 45P (1967).
2. A. J. Smith *et al.*, *J. Bacteriol.* **94**, 972 (1967).
3. S. Zhang, D. A. Bryant, *Science* **334**, 1551 (2011).
4. H. Bothe *et al.*, *Arch. Microbiol.* **96**, 291 (1974).
5. This organism was first isolated in pure culture by C. Van Baalen and named ***Agmenellum quadruplicatum*** strain PR-6; this name appeared in publications from 1962 up to at least 1993. It was renamed *Synechococcus* sp. strain PCC 7002 in 1979 (*17*).
6. T. Khoja, B. A. Whitton, *Arch. Mikrobiol.* **79**, 280 (1971).
7. P. J. Bottomley, C. Van Baalen, *J. Gen. Microbiol.* **107**, 309 (1978).
8. M. L. Summers *et al.*, *J. Bacteriol.* **177**, 6184 (1995).
9. J. L. Ungerer *et al.*, *J. Bacteriol.* **190**, 8115 (2008).
10. G. P. O'Neill *et al.*, *J. Bacteriol.* **170**, 3810 (1988).
11. See genome.kazusa.or.jp.
12. N. Tandeau de Marsac, *J. Bacteriol.* **130**, 82 (1977).
13. O. Fokina, V. R. Chellamuthu, K. Forchhammer, K. Zeth, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19760 (2010).
14. M. I. Muro-Pastor *et al.*, *J. Biol. Chem.* **276**, 38320 (2001).
15. K.-H. Tang *et al.*, *Frontiers Microbiol*. **2**, 165 (2011).
16. D. C. Duca *et al.*, *Trends Biotechnol.* **29**, 95 (2011).
17. R. Rippka *et al.*, *J. Gen. Microbiol.* **111**, 1 (1979).

10.1126/science.1215655

MOLECULAR BIOLOGY

Translation Goes Global

Robert B. Weiss[1] and John F. Atkins[1,2]

An assessment of all ribosomes engaged in protein synthesis expands the view of the potential size of the cellular proteome.

Global information about the extent of proteins being synthesized in a given cell has lagged behind our knowledge of genomic sequence information and RNAs. Now, a recent study highlights the extent to which this deficit is being rectified (*1*). Ingolia *et al*. expand the use of a technique they call ribosome profiling (*2*) to take "snapshots" of the complexity of mammalian proteomes—the constellation of proteins present in a cell or one of its compartments—focusing on messenger RNAs (mRNAs) being actively decoded to synthesize protein products and revealing several surprises.

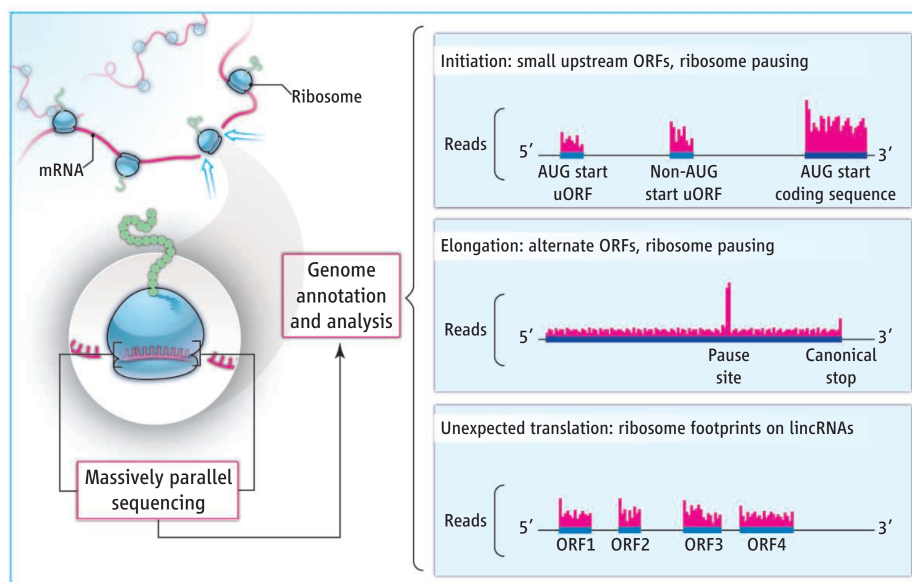
[1]Department of Human Genetics, University of Utah, Salt Lake City, UT 84112, USA. [2]BioSciences Institute, University College Cork, Cork, Ireland. E-mail: bob.weiss@genetics.utah.edu; j.atkins@ucc.ie

Although the translational output of individual mRNAs in a given eukaryotic cell can vary by a factor of 10 or more, it is the ratio of multiple specific mRNAs to proteins being synthesized that is important for understanding global gene expression (*3*). Earlier efforts to measure this ratio included microarray analysis of RNA isolated from polysomes (mRNAs with multiple translating ribosomes) (*4*, *5*) and translational profiling using affinity purification of epitope-tagged ribosomes (*6*, *7*). However, these techniques do not yield the positional and quantitative information revealed by ribosome profiling. Given the number of interactions relevant to most human traits, increased focus is on regulatory rather than structural gene mutations. In this regard, the study of Ingolia *et al*. provides important pointers to the unexpected

PHOTO CREDIT: D. A. BRYANT/PENNSYLVANIA STATE UNIVERSITY

Downloaded from www.sciencemag.org on December 19, 2011

Published by AAAS



Where the ribosomes lie. Ribosomes are "frozen" on actively translating mRNAs, after which RNase digestion (arrows) leaves small (30-nucleotide) ribosome footprints that are converted to DNA and sequenced. Aligning these footprints to the genome reveals ribosome positioning during translation. Compounds that block translation initiation cause footprints to accumulate at start codons (AUG) and uORFs within 5′ untranslated regions (UTRs). Genome-wide analysis reveals the locations of unannotated translation products, ribosomal pause sites, and an unexpected class of short, polycistronic ribosome–associated coding RNAs.

diversity of translational regulatory mechanisms, such as short upstream open reading frames (uORFs).

Ribosomal profiling has been enabled by the recent ability to sequence vast numbers of DNA fragments at once. It builds on an old discovery that the segment of mRNA contained within a ribosome can be isolated on the basis of its resistance to nucleases that destroy unprotected regions of the mRNA. Ribosomal profiling takes a global snapshot for all ribosomes engaged with mRNA in a cell at one instant. It does so by quantitatively converting the nuclease-protected mRNA to DNA and sequencing all the different DNAs at once. By using appropriate inhibitors of either initiation or ribosome progression, or first the initiation inhibitor and later the progression inhibitor, it was possible to identify the translation start sites, their distribution, and the speed of translating ribosomes. The technique was initially applied in yeast, revealing an unexpected number of uORFs (*2*), and subsequently in mammalian cells, to study the effects of microRNAs on translation (*8*).

In the new study, Ingolia *et al*. analyzed a mouse embryonic stem cell line and assessed the disparity between the amounts of individual mRNAs and the amount of their protein products (*1*) (see the figure). The translation rate was generally constant at close to 6 amino acids per second, without numerous small pauses at rare codons; 1500 major pauses, defined as a dwell time up to 25 times that of a standard translation event, were found in 1100 genes. These pause sites have a characteristic 3-codon signature (Pro-Pro-Glu), and analysis of their importance may reveal critical features for the subcellular localization and functions of the encoded proteins.

A startling feature was the extent of the new translational start sites identified. Of the ~5000 genes examined, 13,454 likely start sites [adenine-uracil-guanine (AUG)] were identified, with 65% of the mRNAs containing more than one start site and 16% with four or more. This greatly expands the list of mRNAs containing uORFs. uORFs have major regulatory functions and have been implicated in the synthesis of peptides that stimulate the immune system (*9*). There was a strong enrichment for start codons that, while not the canonical AUG, were related to it by single-nucleotide base substitutions. Only a small number had been identified previously, and the new number increases the potential importance of the hypothesis that start codon selection regulates the initiation of translation (*10*). Even among known coding sequences, 570 likely upstream starts were identified that would yield proteins with extra sequence at their amino-terminal ends. Also detected were 870 likely downstream starts that would yield amino-terminal truncations. These newly identified sites could potentially play partially competing roles. The number of upstream starts is 10 times greater than the number identified in a recent analysis based on coding sequence conservation (*11*), perhaps indicative of a wider regulatory role for upstream starts. The annotation implications for the human genome and regulatory studies are large. Most studies of the proteome identify products predicted to be present and rely on searches for a peptide of a specific mass. The work of Ingolia *et al*. expands our vision of the proteome, even though it does not address protein stability or modifications.

Another major finding concerns the relatively recently discovered class of more than a thousand lincRNAs, large mammalian RNAs that don't contain the characteristics of known protein-coding sequences (*12*). Ingolia *et al*. show that many putative lincRNAs have successive short segments that are translated at a rate similar to comparable classical protein coding sequences. This class of lincRNAs is renamed as short polycistronic ribosome–associated coding RNAs (sprcRNAs). Still, a substantial number of lincRNAs were found not to engage ribosomes. The new information will be crucial for investigations of the linc/sprcRNAs to clarify cytoplasmic versus nuclear functions.

Ingolia *et al*. also address protein synthesis changes that occur when pluripotent embryonic stem cells undergo differentiation; these findings should inform studies of numerous cell types of diverse species and issues. Ribosome profiling seems poised to inaugurate a new era of studies aimed at genome-wide information on protein synthesis (GWIPS), adding a new star in the firmament of genome-wide analysis. It is a fitting way to mark the 50th anniversary of the determination of the general nature of readout of the genetic code (*13*) and of identification of the first code word (*14*).

Downloaded from www.sciencemag.org on December 19, 2011

References

1. N. T. Ingolia *et al*., *Cell* **147**, 789 (2011).
2. N. T. Ingolia *et al*., *Science* **324**, 218 (2009); 10.1126/science.1168978.
3. B. Schwanhäusser *et al*., *Nature* **473**, 337 (2011).
4. Q. Zong, M. Schummer, L. Hood, D. R. Morris, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10632 (1999).
5. G. Johannes, M. S. Carter, M. B. Eisen, P. O. Brown, P. Sarnow, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13118 (1999).
6. M. Heiman *et al*., *Cell* **135**, 738 (2008).
7. E. Sanz *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13939 (2009).
8. H. Guo *et al*., *Nature* **466**, 835 (2010).
9. S. Malarkannan *et al*., *Immunity* **10**, 681 (1999).
10. I. P. Ivanov, G. Loughran, M. S. Sachs, J. F. Atkins, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18056 (2010).
11. I. P. Ivanov, A. E. Firth, A. M. Michel, J. F. Atkins, P. V. Baranov, *Nucleic Acids Res.* **39**, 4220 (2011).
12. M. Guttman *et al*., *Nature* **477**, 295 (2011).
13. F. H. C. Crick, L. Barnett, S. Brenner, R. J. Watts-Tobin, *Nature* **192**, 1227 (1961).
14. M. W. Nirenberg, J. H. Matthaei, *Proc. Natl. Acad. Sci. U.S.A.* **47**, 1588 (1961).

10.1126/science.1216974

CREDIT: Y. HAMMOND/SCIENCE

Published by AAAS