JVI Journals.ASM.org

GEM

Decoding Viral Infection by Ribosome Profiling

Noam Stern-Ginossar

Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

Ribosome profiling is an emerging technique that uses deep sequencing to monitor translation in live cells. Studies using ribosome profiling have already provided novel insights into the identities and amounts of the proteins being produced in cells, as well as novel insights into the mechanism of protein synthesis and translation regulation. Application of ribosome profiling to cells infected with human cytomegalovirus and Kaposi's sarcoma-associated herpesvirus revealed unanticipated complexity in the coding capacity of herpesviruses. Here, I discuss these results and how the application of ribosome profiling to cells infected with other viruses can reveal novel insights into the process of infection.

Genome-wide analyses of gene expression have so far focused on the identification of mRNA species as measured either by microarray or, more recently, by RNA sequencing (RNA-seq). Ribosome profiling is an innovative technique that uses deep sequencing to monitor translation and provide a novel opportunity to monitor gene expression at the level of translation (1, 2). Nuclease footprinting was recognized as a way to determine the positions of ribosomes on mRNAs more than 45 years ago (3). Cycloheximide-treated ribosomes were shown to physically protect fragments of 30 nucleotides of the translated transcript from nuclease digestion, and these protected RNA fragments, named ribosome footprints, indicate the exact location of the ribosome (3, 4). At the time these observations were made, this approach could be applied only to footprints from *in vitro* translation of a single known transcript. These days, ribosome profiling uses high-throughput sequencing to characterize the complex pool of footprints that originate from all of the translating ribosomes in cells (Fig. 1). Since each protected fragment reflects the position of one ribosome on the mRNA, ribosome profiling provides a unique opportunity to map translation events in an unbiased way.

Viruses need the machineries of their host cells for their genome replication, particle assembly, and release of new viral progeny. Many viruses encode the enzymatic machinery required for replication or transcription of their genome, but—with the exception of the recently discovered giant viruses (5)—none of them is known to encode any part of the translational apparatus. This condition makes viruses totally dependent on the cellular machinery for their mRNA translation. Therefore, viruses have evolved a wealth of mechanisms to customize the translation apparatus to meet their specific needs. This is particularly true of RNA viruses, as they have to cope with the dependence of eukaryotic translation machinery on the recognition of the mRNA 5′ cap. This recognition, in its canonical form, permits the production of only a single protein from a given mRNA. Since RNA viruses need to express multiple proteins, they have evolved a variety of strategies to bypass this limitation. These strategies include the use of internal ribosome entry sites, leaky scanning, non-AUG initiation, reinitiation, ribosomal frameshifting, and readthrough (6). However, because of the high density of viral genomes and the ill-defined sequence features of these translational elements, the discovery of these elements *in silico* poses a great challenge.

Ribosome profiling is a robust experimentally based way to identify translation events. Therefore, applying it to cells infected with different viruses has a great potential to reveal viral elements that are being translated during infection and to provide further mechanistic insights into viral gene expression and regulation.

So far, ribosome profiling has been used to annotate two large DNA viruses, human cytomegalovirus (HCMV) (7) and Kaposi's sarcoma-associated herpesvirus (KSHV) (8). To identify the range of HCMV-translated open reading frames (ORFs) and to monitor their temporal expression, we infected human foreskin fibroblasts with HCMV and generated ribosome profiling and parallel RNA-seq libraries. Examination of the full range of HCMV translation products, as reflected by ribosome footprints, revealed many putative previously unidentified ORFs and a fraction of those was confirmed by means of mass spectrometry. These new ORFs originate from a variety of sources, including novel loci, antisense ORFs, short upstream ORFs, messages thought to be noncoding, and a range of alternative protein products that overlap annotated ORFs. In all of these categories, we also observed ORFs starting at near-cognate codons (codons differing from AUG by one nucleotide), especially CUG. Our RNA-seq experiments also revealed that the regulated use of alternative transcription start sites plays a broad role in enabling tight temporal control of HCMV protein expression and allowing multiple distinct polypeptides to be generated from a single genomic locus (7).

Similar extensive translation of novel, mainly short, proteins was also identified in KSHV (8). The functional relevance of these novel ORFs is still an open question, but the number and characterized functions of small peptides are continuously increasing and there is overwhelming evidence of the abundance and relevance of small peptides (9, 10). Of course, the act of ribosome engagement could be important rather than the translation product itself. Translation of ORFs upstream of expressed genes has a well-known regulatory function (11). In addition, the very act of ribosome binding can change the structure and availability of the mRNA for translation of other parts of the message or affect other aspects of RNA metabolism such as RNA stability (11). Thus,

Accepted manuscript posted online 25 March 2015

Citation Stern-Ginossar N. 2015. Decoding viral infection by ribosome profiling. J Virol 89:6164–6166. doi:10.1128/JVI.02528-14.

Editor: F. Goodrum

Address correspondence to noam.stern-ginossar@weizmann.ac.il.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.02528-14

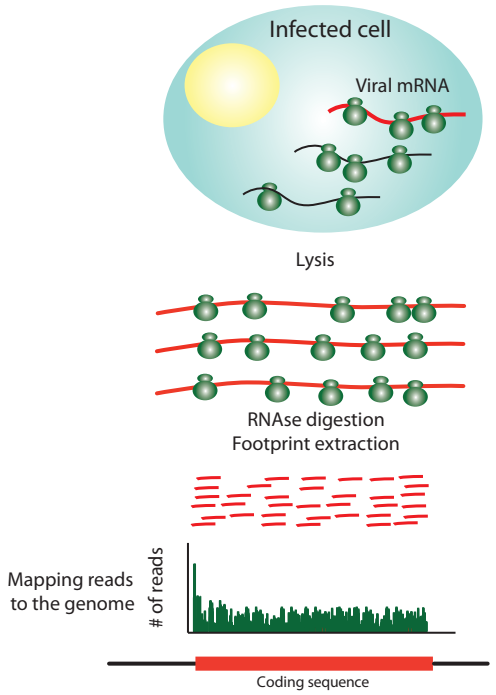

FIG 1 Ribosome profiling allows the identification of translated elements. Cells infected with a given virus are treated with cycloheximide and then lysed. After nuclease digestion, ribosome footprints are isolated and converted to deep-sequencing libraries. Reads are then mapped to the viral genome and allow mapping of translated ORFs in an unbiased quantitative manner.

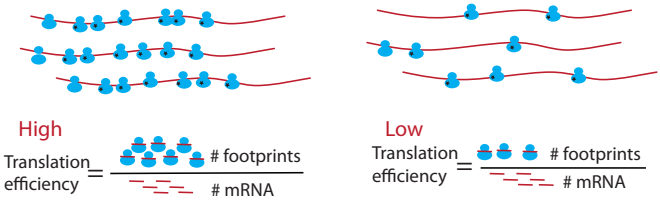

FIG 2 Parallel RNA-seq and ribosome-profiling measurements provide a platform to calculate translation efficiency of expressed genes. By dividing the number of footprints by the number of mRNA reads at every time point during infection, one can accurately calculate the translation efficiency of every expressed gene.

functional ORFs might be retained in regions of the genome where transcription occurs, even if no functional polypeptide is produced.

Although noncanonical translation is thought to be less significant in DNA viruses, ribosome profiling provided a wealth of examples of leaky scanning, non-AUG initiation, and reinitiation in the HCMV and KSHV genomes. Since ribosome profiling was shown to be a powerful tool to identify readthrough (12) and allows the identification of shifts in the translated frame (7), its application to cells infected with additional viruses should allow the global identification and characterization of translated elements. This identification and characterization should enable the generation of a much better catalogue of the coding capacities of different viruses and has the potential to reveal novel functional features.

Regardless of their cellular role, all translated polypeptides can serve as antigens, even if they are rapidly degraded and never accumulate within the cell. In fact, breakdown products from cotranslational degradation may be preferentially targeted for display as antigens (13). The adaptive immune system thus records signatures of foreign protein expression. We therefore mined this record by testing the antigenicity of several novel reading frames we identified in HCMV. We reasoned that if humans with a history of HCMV infection displayed a memory T cell response to novel peptides, as they do to canonical HCMV proteins (14), it would indicate that these peptides were produced in the course of the normal viral life cycle in a human host. Indeed, we were able to show robust cellular immune responses of human T cells from HCMV-seropositive donors against a series of peptides that originated from several short reading frames that we had identified by ribosome profiling (15). These responses were absent from HCMV-seronegative individuals, supporting the natural exposure of HCMV-infected individuals specifically to these newly annotated translation products (15).

Additional information that is obtained from ribosome-profiling experiments is the accurate measurement of viral gene expression. Although microarray-based approaches have been used for the quantification of herpesvirus gene expression during infection (16, 17), those studies were limited by incomplete annotation of transcripts and ORFs, especially in cases in which viral elements are encoded by overlapping or bicistronic transcripts. The data obtained from ribosome profiling provide a unique opportunity to quantify the rate of viral protein production throughout the course of infection. This is particularly useful in the case of large DNA viruses in which viral proteins are expressed in defined kinetics, and the time of expression could provide insights into biological functions. However, these measurements could also be helpful to unravel if there are differences in the production rates of different viral proteins encoded by RNA viruses, as one might expect from the different need for enzymatic and structural proteins.

Since ribosome-profiling experiments are performed on cells infected with viruses, an additional important aspect of probing infection by ribosome profiling is the ability to globally measure the translational changes that occur in cellular genes during viral infection. Control of protein production reflects both the regulation of mRNA levels and the efficiency with which these messages are translated into proteins. A prime means by which viruses influence host physiology is interaction with the cellular translation machinery. Not only do viruses depend on this machinery to translate their own mRNAs, they also must block host defenses that inactivate the cellular translation apparatus. Still, genome-wide analyses of gene expression have so far focused on measurements of changes in transcript abundance. Performing parallel RNA-seq and ribosome-profiling measurements allows direct assessment of differential translation of any cellular and viral gene during infection (Fig. 2). This global examination of translation efficiency provides opportunities to address fundamental questions; such as what is the effect that infection with different viruses has on the spectrum of host mRNAs that are being translated and whether, and to what extent, a given virus possesses mechanisms to commandeer the translation machinery to more effectively translate its own mRNAs at the expense of its host counterparts. The ability to elucidate translational reprogramming of both the virus and the host can reveal novel modules different viruses rely on.

Lastly, recent advances in mass spectrometry techniques allow accurate measurements of changes in the human proteome during infection (18). Since these measurements provide quantitative evaluations of steady-state protein levels during infection and ribosome profiling provides quantitative measurements that reflect the rate of protein synthesis, integration of these measurements could facilitate global identification of the degradation landscape of proteins during infection (as anticorrelation between synthesis rates and protein levels suggests degradation). Therefore, this approach could serve as an excellent platform on which to identify without bias cellular proteins that are actively degraded during infection. Since active degradation of host proteins during infection is likely to indicate biological importance, this integration will provide a powerful way to identify pathways that are central for infection and are directly targeted by viruses.

The genomic era began with the sequencing of the bacterial DNA virus φX174 in 1977 (19) and the mammalian DNA virus simian virus 40 (20) the following year. Since then, extraordinary advances in sequencing technology have enabled the determination of a vast array of viral genomes. However, because of the high density of these genomes, deciphering their coding potential remains a great challenge. Ribosome profiling offers an unprecedented opportunity to better annotate viruses, and in addition, application of the method to infected cells can help reveal novel aspects of the complex interaction between viruses and their hosts.

REFERENCES

1. **Ingolia NT, Lareau LF, Weissman JS.** 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell **147:**789–802. http://dx.doi.org/10.1016/j.cell.2011.10.002.
2. **Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS.** 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science **324:**218–223. http://dx.doi.org/10.1126/science.1168978.
3. **Steitz JA.** 1969. Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. Nature **224:**957–964. http://dx.doi.org/10.1038/224957a0.
4. **Wolin SL, Walter P.** 1988. Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J **7:**3559–3569.
5. **Jeudy S, Abergel C, Claverie J-M, Legendre M.** 2012. Translation in giant viruses: a unique mixture of bacterial and eukaryotic termination schemes. PLoS Genet **8:**e1003122. http://dx.doi.org/10.1371/journal.pgen.1003122.
6. **Firth AE, Brierley I.** 2012. Non-canonical translation in RNA viruses. J Gen Virol **93:**1385–1409. http://dx.doi.org/10.1099/vir.0.042499-0.
7. **Stern-Ginossar N, Weisburd B, Michalski A, Le VT, Hein MY, Huang SX, Ma M, Shen B, Qian SB, Hengel H, Mann M, Ingolia NT, Weissman JS.** 2012. Decoding human cytomegalovirus. Science **338:**1088–1093. http://dx.doi.org/10.1126/science.1227919.
8. **Arias C, Weisburd B, Stern-Ginossar N, Mercier A, Madrid AS, Bellare P, Holdorf M, Weissman JS, Ganem D.** 2014. KSHV 2.0: a comprehensive annotation of the Kaposi's sarcoma-associated herpesvirus genome using next-generation sequencing reveals novel genomic and functional features. PLoS Pathog **10:**e1003847. http://dx.doi.org/10.1371/journal.ppat.1003847.
9. **Andrews SJ, Rothnagel JA.** 2014. Emerging evidence for functional peptides encoded by short open reading frames. Nat Rev Genet **15:**193–204. http://dx.doi.org/10.1038/nrg3520.
10. **Ladoukakis E, Pereira V, Magny EG, Eyre-Walker A, Couso JP.** 2011. Hundreds of putatively functional small open reading frames in Drosophila. Genome Biol **12:**R118. http://dx.doi.org/10.1186/gb-2011-12-11-r118.
11. **Barbosa C, Peixeiro I, Romão L.** 2013. Gene expression regulation by upstream open reading frames and human disease. PLoS Genet **9:**e1003529. http://dx.doi.org/10.1371/journal.pgen.1003529.
12. **Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS.** 2013. Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. eLife **2:**e01179. http://dx.doi.org/10.7554/eLife.01179.
13. **Yewdell JW.** 2007. Plumbing the sources of endogenous MHC class I peptide ligands. Curr Opin Immunol **19:**79–86. http://dx.doi.org/10.1016/j.coi.2006.11.010.
14. **Sylwester AW, Mitchell BL, Edgar JB, Taormina C, Pelte C, Ruchti F, Sleath PR, Grabstein KH, Hosken NA, Kern F, Nelson JA, Picker LJ.** 2005. Broadly targeted human cytomegalovirus-specific $CD4^+$ and $CD8^+$ T cells dominate the memory compartments of exposed subjects. J Exp Med **202:**673–685. http://dx.doi.org/10.1084/jem.20050882.
15. **Ingolia NT, Brar GA, Stern-Ginossar N, Harris MS, Talhouarne GJ, Jackson SE, Wills MR, Weissman JS.** 2014. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. Cell Rep **8:**1365–1379. http://dx.doi.org/10.1016/j.celrep.2014.07.045.
16. **Chambers J, Angulo A, Amaratunga D, Guo H, Jiang Y, Wan JS, Bittner A, Frueh K, Jackson MR, Peterson PA, Erlander MG, Ghazal P.** 1999. DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. J Virol **73:**5757–5766.
17. **Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, DeRisi JL.** 2002. Microarray-based detection and genotyping of viral pathogens. Proc Natl Acad Sci U S A **99:**15687–15692. http://dx.doi.org/10.1073/pnas.242579699.
18. **Weekes MP, Tomasec P, Huttlin EL, Fielding CA, Nusinow D, Stanton RJ, Wang ECY, Aicheler R, Murrell I, Wilkinson GW, Lehner PJ, Gygi SP.** 2014. Quantitative temporal viromics: an approach to investigate host-pathogen interaction. Cell **157:**1460–1472. http://dx.doi.org/10.1016/j.cell.2014.04.028.
19. **Sanger F, Nicklen S, Coulson AR.** 1977. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A **74:**5463–5467. http://dx.doi.org/10.1073/pnas.74.12.5463.
20. **Fiers W, Contreras R, Haegemann G, Rogiers R, Van de Voorde A, Van Heuverswyn H, Van Herreweghe J, Volckaert G, Ysebaert M.** 1978. Complete nucleotide sequence of SV40 DNA. Nature **273:**113–120. http://dx.doi.org/10.1038/273113a0.