RT-PCR Enters the Realm of Stochastic Gene Expression

Sanjay Tyagi, Ph.D.

With the introduction of fluorogenic probes and fluorometric thermal cyclers that enable the monitoring of amplification reactions in real time, the sensitivity and precision of gene expression analyses has improved dramatically. Utilizing these assays, often called quantitative or real-time RT-PCR, populations of mRNAs can now be analyzed from samples as small as a single cell. The reliability of these assays was improved significantly by the development of techniques such as laser-capture dissection for the isolation of individual cells from tissue slices.

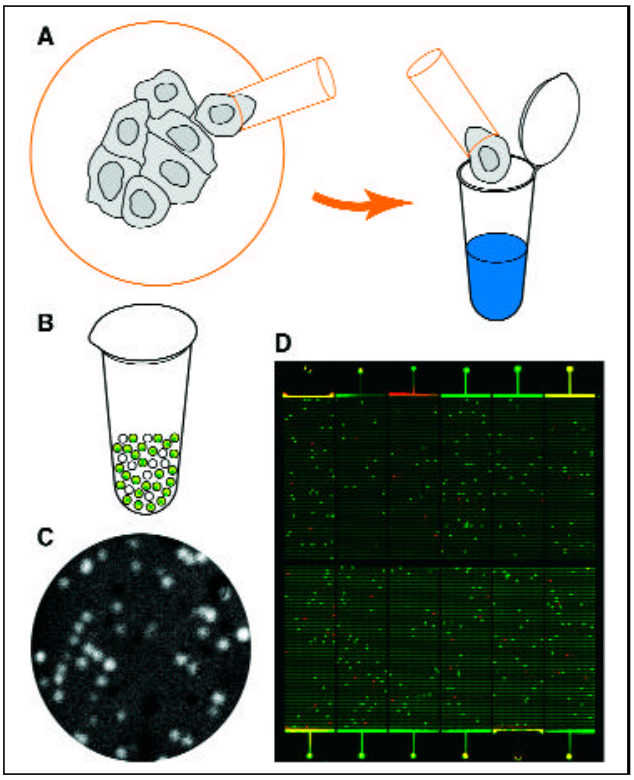

Figure 1. Digital RT-PCR and new formats that extend its dynamic range. (A) Individual cells are isolated from tissue slices or are picked from clusters of cells grown in petri dishes. A single cell is deposited into a solution that lyses the cell and exposes its RNA. (B) A schematic representation of BEAMing, a procedure in which the RNA content of a cell is distributed into PCR chambers formed by microdroplets in a water-oil emulsion. Each droplet also contains a magnetic bead that can bind to and gets labeled by the amplified DNA. The number of beads that become fluorescent indicates the number of microdroplets that received an mRNA molecule and thus indicate the number of copies of mRNA that were originally present in the cell. (C) Molecular colonies in a gel. The mRNA content of a cell is distributed within a flat gel in which RT-PCR is carried out. Since the diffusion of amplicons is limited in the gel, each mRNA molecule leads to the generation of a molecular colony that can be probed in situ with molecular beacons. The number of colonies produced at the end of the reaction reflect the number of mRNA molecules in the source cell. (D) RT-PCR on a microfluidic chip. The RNA content of the cell is distributed among the channels of a microfluidic chip through ports that are visible at the top and bottom of each group of channels. The reaction chambers are created by closing a series of valves that intersect the channels. The image shows the reaction chambers that exhibited amplification of one of two different mRNAs. The amplification products were detected using a pair of green- and red-labeled TaqMan probes.

Images courtesy of Alexander Chetverin, Luigi Warren, and Stephen Quake.

Improved sensitivity, however, exposes problems inherent in amplification-based gene detection. In real-time PCR, the threshold cycle (the number of cycles of amplification required before fluorescence signals first become clearly visible) is linearly proportional to the logarithm of the number of template molecules initially present in the sample. However, when the sample possesses only a few target molecules, that relationship is subject to statistical uncertainties during the early stages of PCR. A sample containing only eight template molecules may behave like a sample possessing 25 molecules or one containing only two molecules. Sometimes researchers can overcome this statistical limitation by analyzing the sample many times and then calculating the mean threshold cycle as a measure of target-copy number.

The number of mRNA molecules in a cell varies from zero to a few thousand, with most species of mRNA being present in less than a hundred copies. Hence, when RNA species are amplified from a sample containing all the RNA from a single cell, there is substantial uncertainty in the measured copy number. To overcome this problem, researchers developed more reliable ways to perform single-cell RT-PCR. One of these methods is based on digital PCR. In digital RT-PCR, a sample containing RNA from a single cell is divided into a large number of reaction wells, so that each well is likely to receive only a single template molecule. Fluorescent probes, such as molecular beacons or TaqMan probes, are utilized to light up the wells that contain amplified template molecules. The number of illuminated wells provides a direct measure of the number of target molecules in the sample.

Originally, digital PCR was performed in 96- or 384-well format and therefore had a dynamic range of 0ñ100 target molecules. However, the dynamic range of these assays has been improved by the introduction of assay platforms possessing thousands of individual wells (Figure 1). In one format, there are 3,000 wells etched into the surface of a glass slide, each with a volume of a few nanoliters. The sample is distributed among the wells, and real-time PCR is performed. In a related format, 1,200 chambers are created in a microfluidic device by the intersection of rows and columns of channels and valves.[1]

In a third format, termed BEAMing (Beads, Emulsions, Amplification, and Magnetics), the sample is distributed into hundreds of thousands of microdroplets in a thermostable water-oil emulsion.[2] Each microdroplet also contains a paramagnetic bead that binds to the amplified DNA and becomes

labeled by fluorescent reporters associated with the amplified DNA. At the end of the amplification, the beads are separated from the emulsion and counted by a fluorescence-activated cell sorter. In the fourth format, the sample is diluted in a gel, and PCR is performed within the gel. Since the gel limits the diffusion of the amplicons during PCR, molecular colonies form at the positions of the original target molecules.[3] The number colonies indicate the number of targets that were in the original sample. Even more interesting formats are likely to follow.

Believing that the results obtained from exponential amplification of a few molecules will always suffer from statistical vagaries despite these innovations, other investigators have tried to do away with PCR altogether. A particularly attractive alternative is to perform in situ hybridization with several oligonucleotide probes against the mRNA target. Each of these probes is labeled with multiple fluorophores so that when they all bind to the same mRNA molecule at the same time, the target molecule appears as a fine fluorescent spot under a fluorescence microscope. All of the spots present in the cell can simply be counted, providing an accurate and integral value for the number of target mRNA molecules expressed in the cell.[4]

Stochastic Forces

Armed with these precise methods, as researchers examine the expression of mRNAs in individual cells, they are finding that gene expression in cells itself is subject to stochastic forces. Warren et al., utilizing digital RT-PCR, analyzed phenotypically identical hematopoietic stem cells for the expression of transcription factor gene PU.1 and expression of the housekeeping gene GAPDH. They found that each cell expressed very different numbers of mRNAs for each of these genes. Using real-time RT-PCR, Bengtsson et al. found that even though insulin producing cells in mouse pancreatic islets are identical in other respects, the level of their insulin mRNA expression varies as much as 10-fold(6). Similarly, in situ hybridization analyses of cultured cells shows that the number of mRNA molecules varies extensively from cell to cell (Figure 2).[5]

Recognizing that large cell-to-cell variations in gene expression are the norm rather than the exception, many investigators have begun to explore the origins and consequences of these variations. Utilizing their in situ hybridization approach, Raj et al. found that large-scale variations in gene expression occur between isogenic population of cells, because mRNAs are not synthesized at a steady rate but are synthesized in bursts beginning and ending in a random manner. Cells that exhibit a large number of mRNAs are those that at the moment of observation are in the middle of a burst of RNA synthesis. Cells that are observed to possess just a few mRNAs have either not experienced a burst yet or had produced a burst of synthesis so long before observation that most of the mRNA molecules had been degraded. They also found that bursts of mRNA synthesis in different genes occur independently of each other.

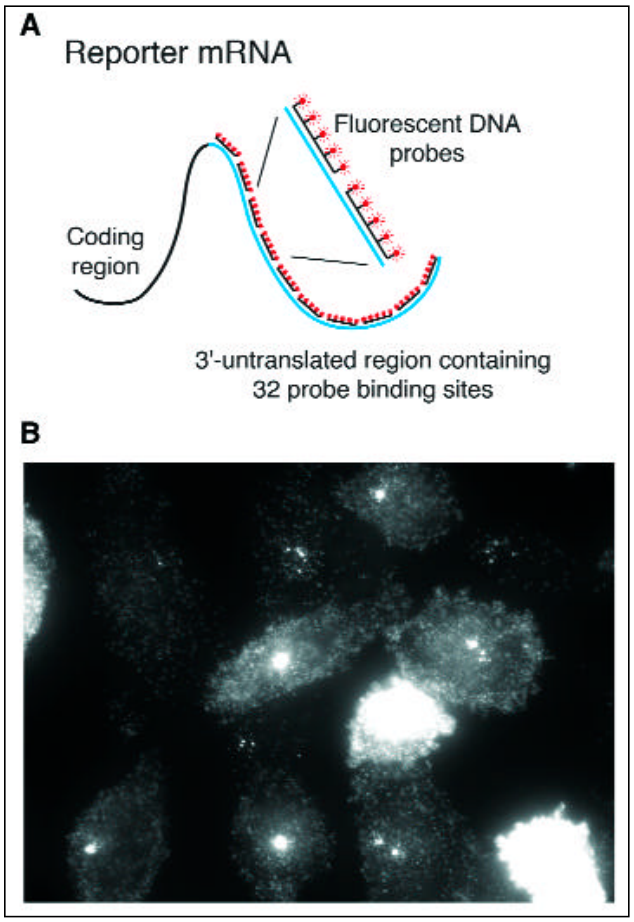

Figure 2. Cell-to-cell variation in gene expression explored with the help of a reporter gene. (A) A schematic representation of the mRNA encoded by a reporter gene. The mRNA can be labeled so intensely with multiple fluorescent probes that each individual mRNA molecule can be seen as a well-defined fluorescent spot within the cell. (B) Cell-to-cell variation in the number of mRNA molecules present in a group of otherwise identical cells. A copy of the reporter gene containing an inserted segment of tandemly repeated target sequences was stably integrated into the genome of a cell line. Fluorescence in situ hybridization was performed using a probe that binds to the repeated sequence. Each spot corresponds to an individual mRNA molecule. Clusters of spots are visible in the nucleus of some cells, where recently synthesized mRNAs have not yet diffused away from the gene locus.

The origin of stochastic mRNA synthesis may lie in unique mechanisms that open up the chromatin in which the gene is embedded, rendering it conducive for mRNA synthesis, and then close down the chromatin, shutting off synthesis.

How do cells achieve their characteristic homogenous phenotypes, given that the number of molecules of a given type of mRNA in each cell is so variable? Part of the answer is that proteins generally stay around in cells longer then mRNAs do. Preexisting pools of proteins receive periodic supplements as a consequence of transient bursts of mRNA synthesis. Since the size of the protein pools is relatively large, it is buffered against variations in mRNA level. Thus,

levels of proteins in cells vary less then levels of mRNAs in cells. However, the lifetimes of different proteins are different from one another. Thus, variations in the level of short-lived proteins are more subject to variations in the level of their respective mRNAs. To cope with such variations, organisms may have developed other, yet unidentified, mechanisms. In some situations, these variations will even be beneficial, serving as an extragenetic substrate for adaptation to transient variations in the environment.

The ultimate aim of gene expression analyses is to determine the levels of different proteins that are present in a cell; some investigators like to extend the profiling to the levels of cellular metabolites. Assuming that levels of proteins are proportional to the levels of mRNAs, biologists have used mRNA analyses as surrogates for more difficult to obtain protein profiles. However, the divergence of mRNA levels and protein levels from cell-to-cell highlights the limitations of this paradigm.

Conventional RT-PCR and microarray-based analyses are typically performed by homogenizing the RNA contents of a large number of cells. The results are then presented as being representative of the behavior of the entire cell population. Single-cell RT-PCR and in situ hybridization procedures are revealing significant and intriguing alterations to that simplistic assumption.

Acknowledgements

The author is supported by National Institutes of Health Grant GM-070357.

References

1. Warren L, Bryder D, Weissman IL and Quake SR, Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. Proceedings of the National Academy of Sciences USA 103:17807-17812, 2006.

2. Diehl F, Li M, He Y, Kinzler KW, Vogelstein B and Dressman D, BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions. Nature Methods 3:551-559, 2006.

3. Samatov TR, Chetverina HV and Chetverin AB, Real-time monitoring of DNA colonies growing in a polyacrylamide gel. Anal Biochem 356:300-302, 2006.

4. Femino AM, Fay FS, Fogarty K and Singer RH, Visualization of single RNA transcripts in situ. Science 280:585-590, 1998.

5. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S, Stochastic mRNA synthesis in mammalian cells. PloS Biol 4:e309, 2006.

6. Bengtsson M, Stahlberg A, Rorsman P, Kubista M, Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. Genome Research, 15:1388-1392, 2005.

Sanjay Tyagi, Ph.D., is at the Public Health Research Institute, New Jersey Medical School, University of Medicine and Dentistry of New Jersey. E-mail: sanjay@phri.org.