The number of genes (~23,000 genes) in the human genome is not much larger than those in lower eukaryotes, however the biological functions of human genome are more intricate and diverse due largely to the complex network of regulation of gene expression. The interaction between each step of gene expression, especially at transcriptional and translational levels, contributes to the large human proteome with estimated 1 million proteins for the vast physiological functions of the human body. As important as it is, gene expression profiling has become a routine experimental approach to study various physiological and pathological questions. Yet, most of current techniques rely on the quantitation of mRNA rather than simultaneous measurement of protein production, leading to a gap between mRNA and protein abundance and an inaccurate explanation of biological functions.

The completion of human genome sequencing project in 2003 has shifted the research paradigm from a gene-based approach to a systems biology trend. Consequently, numerous high throughput techniques have been dramatically improved, exemplified by DNA microarray and next-generation sequencing (NGS), which allow us to analyze gene expression at the genome level. Genome-wide gene expression profiling has been carried out for many diseases. The first global map of human gene expression was accomplished [1] with as many as 369 different types of cells, tissues and disease states. Without a doubt, these achievements fundamentally enhanced our understanding of disease mechanisms; however a major piece of information linking a gene to its function is missing- protein expression. Proteins are functional molecules to carry out the biological tasks of genes. Without knowing protein levels, the current studies are still being conducted at gene- or pathway-based levels in most research laboratories. When gene expression profiling is performed, one should be mindful to have their target genes verified by the lack of robust technique like RT-PCR, microarray and RNA-seq with NGS. The high dynamics and post-translational modifications of proteins pose additional chemical obstacles for protein analysis. Fortunately, mass spectrometry (MS)-based proteomics has advanced to a level sophisticated enough to identify and quantitate proteins on a genome scale, unveiling proteomes in many diseases and cellular processes [2].

The combination of NGS and MS drives us to revisit the importance of protein synthesis in gene expression. It is noteworthy to mention two major technical advances: one is the “4sU-seq” by NGS that can accurately measure the dynamics of newly synthesized mRNA through pulse labeling with 4-thiouridine at physiological conditions [3]; the other is the “quantitative MS” through stable isotope labeling by the lack of robust technique like RT-PCR, microarray and RNA-seq with NGS. The high dynamics and post-translational modifications of proteins pose additional chemical obstacles for protein analysis. Fortunately, mass spectrometry (MS)-based proteomics has advanced to a level sophisticated enough to identify and quantitate proteins on a genome scale, unveiling proteomes in many diseases and cellular processes [2].
is another efficient technique to confirm the translational activity of specific mRNA. Collectively, these strategies will foster a better understanding of the mechanisms that regulate gene expression and the consequence on biological function.

References

1. Lukk M, Kapushesky M, Nikkilä J, Parkinson H, Goncalves A, et al. (2010) A global map of human gene expression. Nat Biotechnol 28: 322-324.
2. Pontén F, Gry M, Fagerberg L, Lundberg E, Asplund A, et al. (2009) A global view of protein expression in human cells, tissues, and organs. Mol Syst Biol 5: 337.
3. Rabani M, Levin JZ, Fan L, Adiconis X, Raychowdhury R, et al. (2011) Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat Biotechnol 29: 436-442.
4. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A 100: 6940-6945.
5. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, et al. (2011) Global quantification of mammalian gene expression control. Nature 473: 337-342.
6. Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, et al. (2011) Comparative analysis of proteome and transcriptome variation in mouse. PLoS Genet 7: e1001393.
7. Foss EJ, Radulovic D, Shaffer SA, Ruderfer DM, Bedalov A, et al. (2007) Genetic basis of proteome variation in yeast. Nat Genet 39: 1369-1375.
8. Baerenfater K, Grossmann J, Grobel MA, Hull R, Hirsch-Hoffmann M, et al. (2008) Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics. Science 320: 938-941.
9. Ling J, Wells DR, Tanguay RL, Dickey LF, Thompson WF, et al. (2000) Heat shock protein HSP101 binds to the Fed-1 internal light regulatory element and mediates its high translational activity. The Plant Cell 12: 1213-1228.
10. Sonenberg N, Hinnebusch AG (2009) Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell 136: 731-745.
11. Zaher HS, Green R (2009) Fidelity at the molecular level: lessons from protein synthesis. Cell 136: 746-762.
12. Ling J, Morley SJ, Traugh JA (2005) Inhibition of cap-dependent translation via phosphorylation of eIF4G by protein kinase Pak2. EMBO J 24: 4094-4105.