L. Sperzel* and S. Ji, Rutgers University

Decoding pathway-specific RNA waves of budding yeast undergoing glucose-galactose shift - the sounds of cell language

Abstract

DNA microarrays allow cell biologists to measure time-dependent changes in the intracellular levels (also called 'gene expression profiles', 'RNA trajectories', 'RNA concentration waves', or just 'RNA waves') of thousands of RNA molecules simultaneously. J. Garcia-Martinez et al. (Mol Cell 15, 303-313 (2004)) measured the time-course of the intracellular levels of about 6000 RNA molecules in budding yeast at 0, 5, 120, 360, 450, and 850 minutes after glucose-galactose shift. Visual inspections indicated that the shapes of the RNA trajectories (or concentration waves) belonging to the glycolysis (22 RNAs) and oxidative phosphorylation (13 RNAs) pathways (the two pathways directly affected by the glucose-galactose shift) are generally opposite to each other and differ from those of the RNA waves belonging to the other pathways of which there are about 200. To make this visual impression more quantitative, we transformed about 1000 of the original (or primary) RNA trajectories into secondary RNA trajectories or waves by using the software ViDaExpert (Visualization of high-dimensional Data Expert) developed by Zinovyev and Gorban in 2000 (http://bioinfo-out.curie.fr/projects/vidaexpert/). Our original data can be represented as a cluster of 1000 points in the 6-dimensional RNA trajectory space, each point representing the shape of one RNA trajectory or wave. ViDaExpert finds an optimal 2-dimensional 'principal grid' that has the property of being as close to all the points in the 6-dimensional RNA trajectory space as possible and associates each RNA trajectory with one of the $n2$ nodes of the principal grid that is closest to it, where n is the linear dimension of the principal grid. By plotting the number of RNA trajectories associated with the nth node against n (the node number or the node address), it is possible to construct a secondary RNA wave or ribonic spectrum, "ribon" indicating an RNA trajectory. Since the topology of the principal grid can be adjusted by changing the stretching coefficient λ, and the bending coefficient μ, it is possible to generate an indefinitely large number of secondary RNA waves or spectra from a given set of N points in the 6-dimensional RNA trajectory space. In the present case, $N = 1000$, which divides into about 50 distinct subclusters, each subcluster associated with one metabolic pathway. We have analyzed a total of

13 metabolic pathways using ViDaExpert with 8 sets of parameters, $n, \lambda, \text{and} \mu$, thereby generating 8 different secondary RNA waves, from which the "average RNA wave", or the "average ribonic spectrum", can be calculated for each metabolic pathway by averaging the number of RNA trajectories mapping onto a given node number. The average RNA waves or ribonic spectra so obtained exhibited three distinct regions, denoted as A (comprising node numbers from 1 to 20), B (from 100 to 120), and C (from 210-230), with the following properties:

(i) The average RNA waves or ribonic spectra of the glycolysis and oxidative phosphorylation pathways have a low peak in Region B, the ratio of this peak to the next largest peak in each spectrum being 0.20 to 0.45, which is small compared to the corresponding ratio of 0.93 (with the standard deviation of 0.163 and the coefficient of variation of 17.5
(ii) The average ribonic spectrum of the glycolytic pathway has a major peak in Region C, that of the oxidative phosphorylation pathway has a major peak in Region A, and those of all the other pathways (except that of transport) has major peaks in Region B. In order to account for these observations, we postulate (i) that the shape of an RNA trajectory is determined by two opposing processes transcription catalyzed by a transcritosome and transcript degradation catalyzed by a 'degradasome' (a term borrowed from bacteriology), (ii) that there exists at least one pair of transcriptosome and degradasome for each metabolic pathway that are functionally coupled to act as a unit which is identified with the 'metabolon' [1], the 'hyperstructure' [2], or the SOWAWN (Self-Organizing-Whenever-And -Wherever-Needed) machines [3], and
(iii) that a pathway-specific metabolon can exist in at least three distinct states the (+)-, (-)-, or (0)-states, effectuating, respectively, an increase, a decrease, or a steady state in the intracellular levels of RNA molecules belonging to a metabolic pathway. Since all of the metabolic pathways examined so far obey the blackbody (or Planck) radiation equation [3, 4, 5] and since the single-molecule enzymic activity data of cholesterol oxidase also obey the same radiation equation (which is interpreted as an evidence that a thermal activation step is implicated in single-molecule enzymic catalysis) [3, 6], it is here suggested that the pathway-specific metabolons are thermally activated to occupy one of the three states, i.e., (+)-, (-)-, and (0)-states, for varying time periods of time, thereby producing pathway-specific average RNA waves with

major peaks in Regions A, B, or C in the ribonic spectra. Finally, the possibility suggests itself that these RNA waves can act as the sounds of cell language whose design features were found to be isomorphic with those of human language [7, 8].

References:

1. Srere, P.A. (1987). Complexes of Sequential Metabolic Enzymes. Ann. Rev. Biochem. 56: 89-124.
2. Norris, V. et al. (1999). Hypothesis: Hyperstructures regulate bacterial structure and the cell cycle. Biochimie 81: 915-920.
3. Ji, S. (2011). Molecular Theory of the Living Cell: Concepts, Molecular Mechanisms, and Biomedical Applications. Springer, New York (to appear).
4. Ji, S. and So, K. (2009). The universal law of thermal transitions applicable to blackbody radiation, single-molecule enzymology and whole-cell metabolism. Abstract B1, The 102nd Statistical Mechanics Conference, Rutgers University, Piscataway, N.J., December 13-15.
5. Ji, S. and So, K. (2010). This conference.
6. Ji, S. (2008). Modeling the single-molecule enzyme kinetics of cholesterol oxidase based on Planck's radiation formula and the principle of enthalpy-entropy compensation. In: Short Talk Abstracts, The 100th Statistical Mechanics Conference, December 13-16, Rutgers University, Piscataway, N.J.
7. Ji, S. (1997). Isomorphism between cell and human languages: molecular biological, bioinformatics and linguistic implications. BioSystems 44: 17-39.
8. Ji, S. (1999). The Linguistics of DNA: Words, Sentences, Grammar, Phonetics, and Semantics. Ann. N. Y. Acad. Sci. 870: 411-417.