Extraction of Biomolecular Signals Controlling Complex Behavior of Biological Cells

Yuichi Sakumura and Katsuyuki Kunida
Nara Institute of Science and Technology, Nara, Japan
E-mail: saku@bs.naist.jp Tel: +81-743-72-5571
E-mail: kkunida@bs.naist.jp Tel: +81-743-72-5545

***Abstract*— Cell deformation and migration are one of the important biological functions related to various biological phenomena. In this study, using the time-series data of cell morphological changes (protrusion and retraction of cell leading edge) and activities of regulatory molecules, we propose a novel method of data analysis for extracting molecular activity patterns corresponding to specific morphological changes based on the reverse correlation analysis. This method was applied to the time-series data of the three representative regulatory molecules measured from cultured cells. We extracted the dynamical molecular activities for four types of the morphological change; constant protrusion and retraction, and transient protrusion and retraction. Although it has been difficult to link cell movement and biochemical findings, we confirmed that our results are consistent with previous findings in the field of biochemistry.**

I. INTRODUCTION

Cells receive extracellular stimuli and convert them into cellular functions such as cell proliferation, differentiation, development, and deformation followed by migration. In particular, cell deformation and migration are related to various biological functions, so their principle elucidation is an important research subject in cell biological field [1]. Most cellular functions are controlled by biomolecules such as proteins and genes.

Cell has structural assembly of proteins called cytoskeleton [2]. Remodeling the cytoskeleton dynamically at the cell edge, cell can protrude and retract for its deformation and migration. Cytoskeleton deformation is controlled by a biochemical reaction network composed of various biomolecules. Among them, the activity of the molecules called Rho family GTPases (G proteins) plays a central role in the regulation of cytoskeleton. The main molecules of G proteins are RhoA, Rac1, and Cdc42 [3], [4]. Cdc42 and Rac1 promote skeletal formation [5]–[8], and RhoA leads to myosin II-mediated skeletal contraction [9]–[11]. These biochemical studies show the static relationship between the constant activity of the G proteins and the static phenotype of cell morphology.

However, dynamic morphological control during cell movement is still unclear. There are studies in which there is no logical consistency between G proteins activity and cell deformation; RhoA contributes to protrusion rather than retraction, and Cdc42 results in morphology stabilization rather than protrusion [12], [13]. These results imply almost the opposite effect of the G proteins to the result in the biochemical observation under the static conditions. Although Dynamic conditions are different from static conditions because they involve temporal changes, the differences are too large to maintain a logical consistency.

There are three reasons why it is difficult to elucidate dynamic movement control by G protein. The first difficulty lies in the limitation of observation in biological experiments. At present, it is difficult to simultaneously observe multiple G protein activities in the same cell. Therefore, cell movement control by coordination of the three G proteins cannot be evaluated. The second difficulty is the presence of other factors that control cell movement. As cell motility is always subject to internal and external mechanical constraints [14], [15], the G protein signal is not always reflected in cell motility. The third difficulty is the problem of mathematical analysis methods. Since multiple Rho G proteins cannot be observed simultaneously and there are complex relationships with other regulatory factors, there is no method to mathematically evaluate the dynamic relationship between molecules and cell movement.

In this study we propose the novel methodology called Motion-Triggered Average (MTA). The MTA is an application of the Spike-Triggered Average (or Reverse correlation) method often used in neuroscience [16]–[20]. The present work is the first application to images of cell movement. Since edge movement of cells is perturbed by various molecular activities, thus the molecular activity patterns that results in a specific edge movement include noise that is considered white noise. If the molecular activity time series of the specific edge movement is collected and averaged, the noise is offset and the original activity series are approximated.

The MTA has the following advantages. The first is that different molecular activity sequences observed independently can be associated with specific edge movement patterns. This solves analytically the experimental constraint, that is, that molecules can only be observed independently. The second is that the dynamic relationship between molecules and cell edge movement can be freely extracted. There is no restriction on the setting of the edge movement pattern, and a molecular activity sequence for the movement can be obtained. The third point is that the causality between molecular activity and edge

K. Kunida receives funding from a Grant-in-Aid for Young Scientists (# 19K20400).

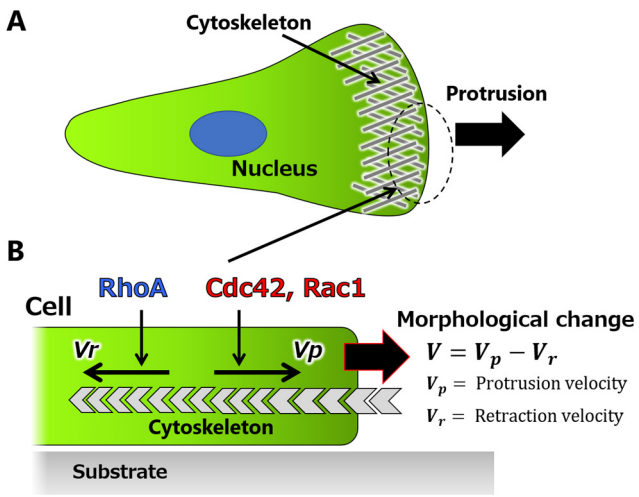

Fig. 1: Cell migration and biochemical reaction network that controls cell movement. (A) Dynamical intracellular cytoskeletal structure formed by actin molecules. (B) Cytoskeleton formation is regulated by the upstream biochemical signals, Cdc42, Rac1, and RhoA.

movement can be addressed because both of the time series can be obtained instead of the correlations.

We applied the MTA to the quantitative data of G protein activity and movement of the cultured cells. RhoA, Rac1, or Cdc42 activity was individually measured in multiple cells. As reference cell movements, we adopted four kinds of movements; constant protrusion and retraction, and transient protrusion and retraction. As a result, similar time series of G protein activity were obtained in different cells for the same reference movement. The resultant relationships between the activity and the movement were consistent with the biochemical causality described above. These results indicate that the intracellular molecular activity extraction using MTA is appropriate methodology.

II. Research Background

A. Biochemical Reaction Network that Controls Cell Migration

Cells do not slide and move in the same shape, but move by protruding and retracting the cell leading edge (Fig. 1A). Protrusion and retraction are largely dependent on the shape of the cytoskeleton, and the three important Rho-family G proteins (Cdc42, Rac1, and RhoA) have a significant influence on this cytoskeleton remodeling. Each G protein activity plays a role in protrusion and retraction. As the biochemical reaction network, Cdc42 and Rac1 control protrusion [5]–[8], and RhoA regulates retraction [9]–[11] (Fig. 1B).

B. Dynamic Measurement of Regulatory Molecules

In order to measure the activity of G protein during cell migration, it is common to use live-cell imaging measurement. A fluorescent biosensor based on Green Fluorescence Protein (GFP) enables protein activity and localization to be monitored in living cells. Molecular activity can be quantified by image brightness values [21].

C. Previous Research

The current mainstream analysis is the time-shifted cross-correlation between time series of G protein activity and cell movement [12], [13], [22]–[25]. The cross-correlation coefficient is obtained by shifting one time series to the future or the past time, and the causality is evaluated by the sign of the shift time at which it becomes maximum.

There are several problems with this time-shift cross-correlation. The first is that the dynamic relationship between sequences is lost in cross correlation. Since the similarity between time series is evaluated by time averaging, the relationship with time change cannot be evaluated. The second is that the obtained cross-correlation coefficient is very small (it hardly exceeds 0.5). This implies that the time series of molecules and movements have little synchrony in the timing of increase and decrease. The third is that the positive or negative of the shift time at which the correlation coefficient is maximum does not necessarily indicate a causal relationship. The condition for increasing the correlation coefficient is that the timings of increase and decrease of the two series coincide with each other. Consider the case where molecular activity is slowly increasing and decreasing, while the cell edge rapidly protrudes and retracts. In this case, when the series of edge movement is shifted to the future, the correlation between the two series is maximized. However, it cannot be argued that edge movement caused molecular activity change.

From the above, this study proposes a new data analysis method for extracting dynamic relationship from data of morphological change and G protein activity.

III. Proposed Method

A. Motion-Triggered Average Analysis

In this study, we propose Motion-Triggered Average (MTA) inspired by Spike-Triggered Average (STA) (Fig. 2). STA collects and averages input stimuli before and after the time at which the neuron generates an action potential. Thereby, an input stimulus that triggers an action potential can be extracted without considering a complex signal processing from input to output. Although the electrical response of the neuron is expressed as the discrete timing of the action potential generation, the response of the cell movement is not discrete timing but a various pattern of cell edge displacement. Therefore, it is difficult to estimate which molecular activity pattern causes cell movement by assembling elementary processes of molecular reactions in a bottom-up manner. Using MTA, however, activity patterns can be extracted for any cell edge movement that we specify.

The MTA uses quantitative data on cell edge displacement and G protein activity near the edge (Fig. 2). The displacement is measured at an observation axis crossing the cell edge, and the molecular activity is obtained as the light intensity of the fluorescent protein. Let $x_i(t)$ be a time-series of cell edge

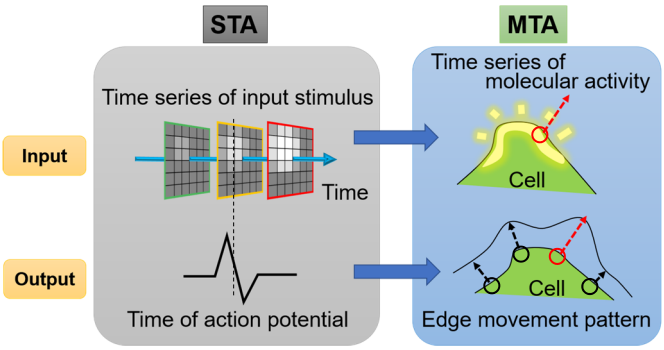

Fig. 2: Schematic comparison between spike-triggered average and motion triggered average. Both algorithms extract the average input stimuli that triggered a specific pattern in output response. Arrows in the right panel represent axes for measuring displacement and activity in the vicinity (circles), and the red one indicates the measurement axis selected based on correlation with a reference displacement.

displacement, where t is a relative time in $[-\tau, \cdot\cdot 0, \cdot\cdot \tau]$, and i is an data index in the measurement data set, $\{x_i(t)\}$. The time-series of the corresponding molecular activity, $a_i(t)$, is defined for each edge displacement $x_i(t)$. Here we introduce the reference edge displacement, $X_j(t)$, which is used for selecting the measured displacement from the data set based on the cross correlation between $X_j(t)$ and $x_i(t)$. The average time-series of molecular activity for the j-th reference displacement, $A_j(t)$, is defined as,

$$A_j(t) = \frac{1}{N_j}\sum_{i\in I_j} a_i(t) \quad \mathbf{s.t.}\ I_j = \{i \mid corr(X_j, x_i) \geq 0.8\}, \quad (1)$$

where N_j is the number of elements of the subset of the measurements, I_j.

We consider four reference edge displacements, $X_j(t)$ $(j = 1, 2, 3, 4)$, constant protrusion and retraction, and transient protrusion and retraction, while MTA can be applied to any type of reference in principle. Any complex cell edge movements can be constructed with a combination of these four movements. The difference in molecular activities between continuous and changing movements can likely give important information, as biology wants to know the signal processing mechanism of molecules that cause cell movement.

IV. Results

The experimental data was given from that in reference [25], and about 20,000 time series were obtained from five individual cells per molecule. The time series of the molecular activity by MTA from the cells were all similar (not shown). The mean time series over all cells with reference to constant movement is shown in Fig. 3. The reference displacement is monotonically increasing or decreasing (Fig. 3A). The extracted time series of three molecular activities show relative

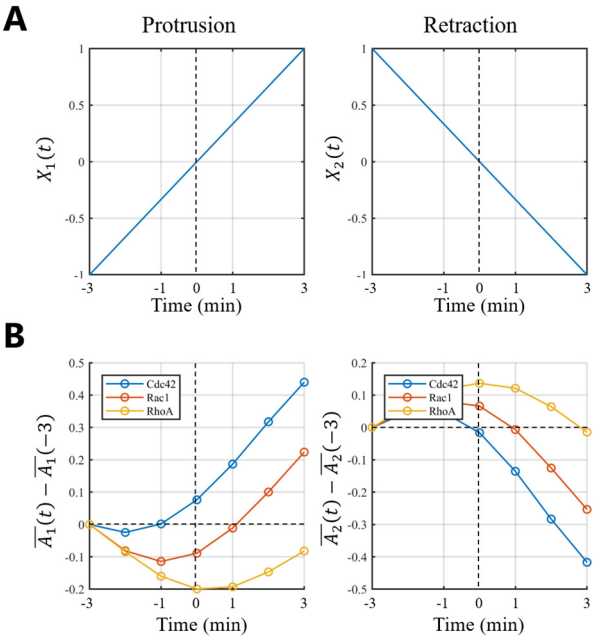

Fig. 3: MTA for constant movement as reference response. (A) Reference time series of constant edge displacement of protrusion (left) and retraction (right). Vertical axis represents relative coordinate. (B) MTA of G protein activities for the displacements given in (A). $\tau = 3$.

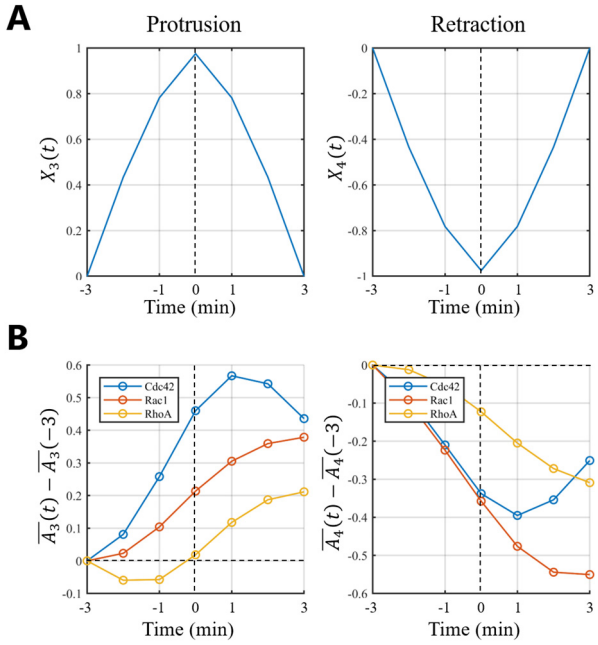

Fig. 4: MTA for transient movement as reference response. (A) Reference time series of transient edge displacement of protrusion (left) and retraction (right). Vertical axis represents relative coordinate. (B) MTA of G protein activities for the displacements given in (A). $\tau = 3$.

activity to the initial value at the relative time −3.

In constant protrusion, Cdc42 is activated after RhoA is inactivated (Fig. 3B left). Rac1 shows two other intermediate changes and clearly behaves differently from Cdc42 while it has the same function as Cdc42 biochemically. The molecular activities of the constant retraction became like their upside down of the protrusion. (Fig. 3B right). RhoA is activated first, and then Cdc42 is inactivated. These results are consistent with biochemical findings; Rac1 and Cdc42 promote cytoskeleton formation and RhoA shrinks the cytoskeleton.

Next, in order to examine changes in G protein activity with respect to velocity changes, MTA was performed using wave-like transient movement instead of constant one (Fig. 4). We used two fluctuating movements retraction after protrusion and protrusion after retraction as reference movement (Fig. 4A). The G protein activities change clearly different from those in the constant movement (Fig. 3). Especially, in the transient protrusion, RhoA activity is increased and Cdc42 activity is saturated and decreases in the latter half of the protrusion (fig. 4B). In the transient retraction, the active time series is not upside-down of the time series as shown in Fig. 3, but the trend is similar to the upside-down one. In the latter half, only Cdc42 changes from a decrease to an increase, and RhoA decreases even in the first half where the edge is retracted. These results are also consistent with the biochemical findings.

V. Summary

In this study, we proposed a novel data analysis method to extract time-series of molecular activity that control the specific pattern of morphological change based on reverse correlation analysis. We have extracted molecular activity data for morphological change patterns that protrude or retract continuously, and confirmed that analysis results consistent with previous biological findings can be obtained. This is the first study to quantitatively link the activity time series of three molecules with cell movement. In future, we will perform the system identification by integrating time-series data of molecular activity extracted by MTA analysis and mathematical models based on physical processes including biochemical reaction and aim to construct a quantitative cell migration model that secures robustly prediction performance.

Acknowledgment

We thank K. Aoki, T. Nakamura, N. Takagi, T. Tanaka, and K. Ikeda for helpful discussions.

References

[1] A. Lauffenburger and A. F. Horwitz, “Cell Migration: A Physically Integrated Molecular Process,” Cell, vol. 84, no. 3, pp. 359–369, 1996.

[2] T. D. Pollard and G. G. Borisy, “Cellular Motility Driven by Assembly and Disassembly of Actin Filaments,” *Cell*, vol. 112, no. 4, pp. 453–465, 2003.

[3] D. Nobes and A. Hall, “Rho, Rac, and Cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia,” Cell, vol. 81, no. 1, pp. 53–62, 1995.

[4] L. M. Machesky and A. Hall, “Role of Actin Polymerization and Adhesion to Extracellular Matrix in Rac- and Rho-induced Cytoskeletal Reorganization,” *The Journal of Cell Biology*, vol. 138, no. 4, pp. 913–926, 1997.

[5] H. Miki, T. Sasaki, Y. Takai, and T. Takenawa, “Induction of filopodium formation by a WASP-related actin-depolymerizing protein N-WASP,” *Nature*, vol. 391, no. 6662, p. 34208, 1998.

[6] R. Rohatgi *et al.*, “The Interaction between N-WASP and the Arp2/3 Complex Links Cdc42-Dependent Signals to Actin Assembly,” *Cell*, vol. 97, no. 2, pp. 221–231, 1999.

[7] A. J. Ridley, H. F. Paterson, C. L. Johnston, D. Diekmann, and A. Hall, “The small GTP-binding protein rac regulates growth factor-induced membrane ruffling,” *Cell*, vol. 70, no. 3, pp. 401–410, 1992.

[8] J. Ridley, “Rho GTPases and actin dynamics in membrane protrusions and vesicle trafficking,” *Trends Cell Biol*, vol. 16, no. 10, pp. 522–529, 2006.

[9] T. Leung, X. Chen, E. Manser, and L. Lim, “The p160 RhoA-binding kinase ROK alpha is a member of a kinase family and is involved in the reorganization of the cytoskeleton.,” *Mol Cell Biol*, vol. 16, no. 10, pp. 5313–5327, 1996.

[10] T. Matsui *et al.*, “Rho-associated kinase, a novel serine/threonine kinase, as a putative target for small GTP binding protein Rho.,” *Embo J*, vol. 15, no. 9, pp. 2208–2216, 1996.

[11] S. Etienne-Manneville and A. Hall, “Rho GTPases in cell biology,” *Nature*, vol. 420, no. 6916, p. 629, 2002.

[12] M. Machacek *et al.*, “Coordination of Rho GTPase activities during cell protrusion,” *Nature*, vol. 461, no. 7260, p. 99, 2009.

[13] Y. Tsukada, K. Aoki, T. Nakamura, Y. Sakumura, M. Matsuda, and S. Ishii, “Quantification of Local Morphodynamics and Local GTPase Activity by Edge Evolution Tracking,” *PLoS Computational Biology*, vol. 4, no. 11, p. e1000223, 2008.

[14] K. Baba *et al.*, “Gradient-reading and mechano-effector machinery for netrin-1-induced axon guidance,” *eLife*, vol. 7, 2018.

[15] A. Mogilner and J. Zhu, “Cell Polarity: Tension Quenches the Rear,” *Curr Biol*, vol. 22, no. 2, pp. R48–R51, 2012.

[16] Z. Mainen and S. T. Science, “Reliability of spike timing in neocortical neurons,” 1995.

[17] Boer and P. Kuyper, “Triggered Correlation,” Ieee T Bio-med Eng, vol. BME-15, no. 3, pp. 169–179, 1968.

[18] P. Z. Marmarelis and K.-I. Naka, “White-Noise Analysis of a Neuron Chain: An Application of the Wiener Theory,” *Science*, vol. 175, no. 4027, pp. 1276–1278, 1972.

[19] L. Paninski, “Convergence properties of three spike-triggered analysis techniques,” *Netw Comput Neural Syst*, vol. 14, no. 3, pp. 437–464, 2003.
[20] T. Sharpee, N. C. Rust, and W. Bialek, “Analyzing Neural Responses to Natural Signals: Maximally Informative Dimensions,” *Neural Comput*, vol. 16, no. 2, pp. 223–250, 2006.
[21] K. Aoki and M. Matsuda, “Visualization of small GTPase activity with fluorescence resonance energy transfer-based biosensors,” *Nat Protoc*, vol. 4, no. 11, p. 1623, 2009.
[22] J. Odde, E. M. Tanaka, S. S. Hawkins, and H. M. Buettner, “Stochastic dynamics of the nerve growth cone and its microtubules during neurite outgrowth,” Biotechnology and Bioengineering, vol. 50, no. 4, pp. 452–461, 1996.
[23] A. Ponti, M. Machacek, S. Gupton, C. Waterman-Storer, and G. Danuser, “Two Distinct Actin Networks Drive the Protrusion of Migrating Cells,” *Science,* vol. 305, no. 5691, pp. 1782–1786, 2004.
[24] L. Ji, J. Lim, and G. Danuser, “Fluctuations of intracellular forces during cell protrusion.,” *Nature cell biology*, vol. 10, no. 12, pp. 1393–400, 2008.
[25] K. Kunida, M. Matsuda, and K. Aoki, “FRET imaging and statistical signal processing reveal positive and negative feedback loops regulating the morphology of randomly migrating HT-1080 cells,” *J Cell Sci*, vol. 125, no. 10, pp. 2381–2392, 2012.