論文の内容の要旨

System Identification of Different Time-Scale Biological Phenomenon from Signal Transduction to Gene Expression

（シグナル伝達から遺伝子発現にかけて
異なる時間スケールで変動する生命現象のシステム同定）

氏名　土屋 貴穂

In intracellular signaling system, information of an extracellular stimulus is once encoded into combinations of distinct temporal patterns of phosphorylation of intracellular signaling at a scale of tens of minutes that are selectively decoded by downstream gene expression with a scale of hours to days, leading to cell fate decisions such as cell differentiation, proliferation and death (Behar and Hoffmann, 2010; Purvis and Lahav, 2013). In rat adrenal pheochromocytoma PC12 cells focused on in this study, nerve growth factor (NGF) induces cell differentiation mainly through sustained phosphorylation of ERK (Gotoh et al., 1990; Marshall, 1995; Qiu and Green, 1992; Traverse et al., 1992), whereas pituitary adenylate cyclase-activating polypeptide (PACAP) induces cell differentiation mainly through cAMP-dependent CREB phosphorylation (Akimoto et al., 2013; Gerdin and Eiden,11 2007; Saito et al., 2013; Vaudry et al., 2002; Watanabe et al., 2012), indicating that combinations of distinct temporal patterns of phosphorylation of

intracellular signaling induce cell differentiation in PC12 cells (Vaudry et al., 2002).

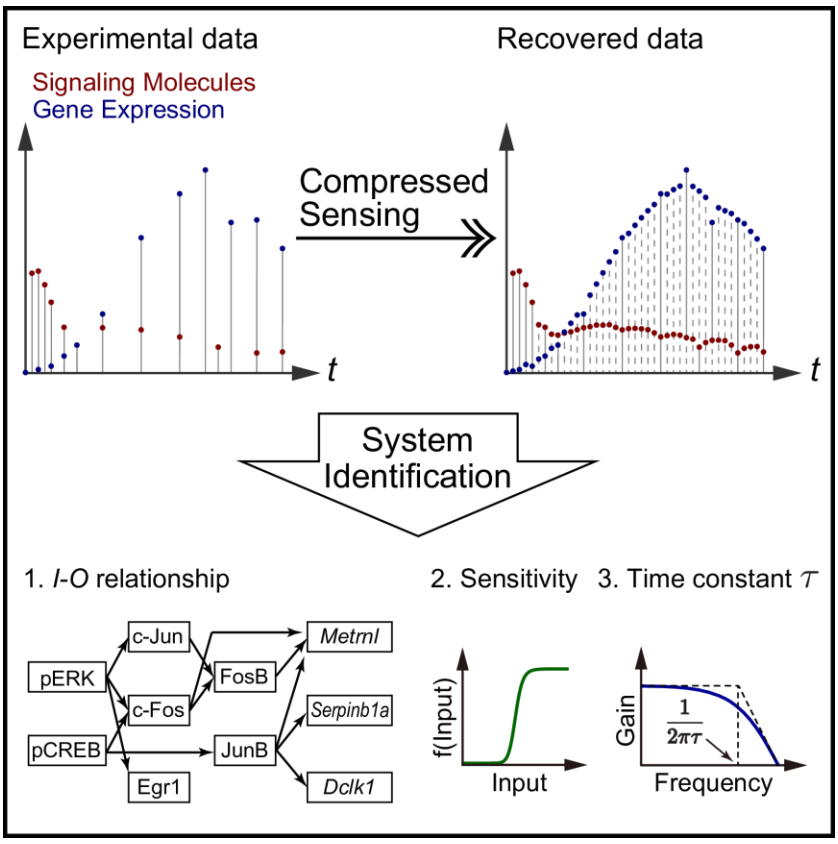
In PC12 cell differentiation, key genes are also identified as the downstream decoding genes essential for cell differentiation in PC12 cells, including *Metrnl*, *Dclk1*, and *Serpinb1a*, denoted as LP (latent process) genes, which are the decoders of neurite length information (Watanabe et al., 2012). Importantly, the expression levels of the LP genes, but not the phosphorylation level of ERK, correlate with neurite length regardless of types of extracellular stimuli. Thus, this unrevealed decoding mechanism of signaling (a shorter time scale) dependent LP gene expression (a longer time scale) is a key issue for understanding the mechanism of cell differentiation. Thus, I focused on this decoding mechanism in this study.

To identify decoding mechanisms by gene expression, the system identification method was employed for identifying input–output relationships from time series data without detailed prior knowledge of signaling pathways (Janes and Lauffenburger, 2006; Janes and Yaffe, 2006; Kholodenko et al., 2012; Ljung, 2010; Price and Shmulevich, 2007; Zechner et al., 2016). In the previous study, a system identification method based on time series data of signaling molecules and gene expression, denoted as the nonlinear autoregressive exogenous (NARX) model has been developed and applied it to the signaling-dependent immediate early genes (IEGs) expression during cell differentiation in PC12 cells (Saito et al., 2013). However, one of the difficulties of the NARX model is to require equally spaced dense time series data ideally. If the time scale of upstream and downstream molecules is different, such as signaling molecules (tens of minutes scale) and LP gene expression (a day scale) in this study (Doupé and Perrimon, 2014), it is technically difficult to obtain sufficient equally spaced dense time series data over the desirable time period due to experimental and budget limitations. Therefore, in reality, for a longer time scale experiment, unequally spaced sparse time series data rather than equally spaced dense time series data are available especially in biological experiments. However, no system identification method based on such sparse data due to different time scale exists.

Here I developed a system identification method by integrating the NARX model and a signal recovery technique in the field of compressed sensing originally developed for image analysis to biological sparse data of different time scales by recovering signals of missing time points (Summary Figure). I measured phosphorylation of ERK and CREB, IEGs expression products, and mRNAs of the decoder genes for neurite length in PC12 cell differentiation and performed the developed system identification, revealing the input–output relationships between signaling and gene expression with sensitivity such as graded or switch-like response and with time constant and gain, representing signal transfer efficiency (Summary Figure). Furthermore, I predicted and validated the identified system using pharmacological perturbation. The identifeid system was also princially consistent with the previous study (Saito et al., 2013).

I found that the LP genes depend only on the IEGs (c-Fos, FosB and/or JunB) but not other upstream molecules, and that the time constants of the LP genes are short except for *Serpinb1a*. This means that the timing of the final decoding step for neurite length information is not directly determined by the IEGs and LP genes, rather by the steps from extracellular stimuli to the IEGs. Furthermore, the identified system captured a selective NGF- and PACAP-signaling decoding system of neurite length information by LP gene expression by using different signaling pathways.

The developed system identification method in this study can solve different time-scale problem and can be broadly applied to different time scale biological phenomena, such as the cell cycle, development, regeneration, and metabolism involving ion flux, metabolites, and gene expression not limited to this study case. Thus, I provide a versatile method for system identification of various biological phenomena using data with different time scales.



Summary figure. Schematic overview of system identification of different time-scale biological phenomenon.