



# Real-Time Temporal Dynamics of Bicistronic Expression Mediated by Internal Ribosome Entry Site and 2A Cleaving Sequence

Soomin Lee<sup>1</sup>, Jeong-Ah Kim<sup>1,2</sup>, Hee-Dae Kim<sup>3</sup>, Sooyoung Chung<sup>4</sup>, Kyungjin Kim<sup>1</sup>, and Han Kyoung Choe<sup>1,5,\*</sup>

<sup>1</sup>Department of Brain and Cognitive Sciences, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu 42988, Korea, <sup>2</sup>Department of Biological Sciences, Seoul National University, Seoul 08826, Korea, <sup>3</sup>Department of Basic Medical Sciences, University of Arizona College of Medicine-Phoenix, Phoenix, AZ 85004, USA, <sup>4</sup>Department of Brain and Cognitive Sciences, Scranton College, Ehwa Womans University, Seoul 03760, Korea, <sup>5</sup>Korea Brain Research Institute (KBRI), Daegu 41062, Korea

\*Correspondence: choehank@dgist.ac.kr

<https://doi.org/10.14348/molcells.2019.2427>

[www.molcells.org](http://www.molcells.org)

Multicistronic elements, such as the internal ribosome entry site (IRES) and 2A-like cleavage sequence, serve crucial roles in the eukaryotic ectopic expression of exogenous genes. For utilization of multicistronic elements, the cleavage efficiency and order of elements in multicistronic vectors have been investigated; however, the dynamics of multicistronic element-mediated expression remains unclear. Here, we investigated the dynamics of encephalomyocarditis virus (EMCV) IRES- and porcine teschovirus-1 2A (p2A)-mediated expression. By utilizing real-time fluorescent imaging at a minute-level resolution, we monitored the expression of fluorescent reporters bridged by either EMCV IRES or p2A in two independent cultured cell lines, HEK293 and Neuro2a. We observed significant correlations for the two fluorescent reporters in both multicistronic elements, with a higher correlation coefficient for p2A in HEK293 but similar coefficients for IRES-mediated expression and p2A-mediated expression in Neuro2a. We further analyzed the causal relationship of multicistronic elements by convergent cross mapping (CCM). CCM revealed that in all four conditions examined, the expression of the preceding gene causally affected the dynamics of the subsequent gene. As with the cross correlation, the predictive skill of p2A was higher than

that of IRES in HEK293, while the predictive skills of the two multicistronic elements were indistinguishable in Neuro2a. To summarize, we report a significant temporal correlation in both EMCV IRES- and p2A-mediated expression based on the simple bicistronic vector and real-time fluorescent monitoring. The current system also provides a valuable platform to examine the dynamic aspects of expression mediated by diverse multicistronic elements under various physiological conditions.

**Keywords:** 2A, expression dynamics, internal ribosome entry site, multicistronic elements, real-time fluorescent imaging

## INTRODUCTION

Multicistronic vectors are valuable tools for the co-expression of multiple genes employed in a variety of fields, including the biological sciences, bioengineering, and biomedical applications. In the case of plasmids or viral vectors for ectopic expression, the expression of multiple transgenes from the same cis-regulatory elements, including promoter, enhancer, and poly (A) signal, efficiently improves the packaging ca-

Received 19 November, 2018; revised 30 March, 2019; accepted 5 April, 2019; published online 26 April, 2019

eISSN: 0219-1032

©The Korean Society for Molecular and Cellular Biology. All rights reserved.

©This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>.

capacity for transgenes (Bouabe et al., 2008; Liu et al., 2017; Szymczak et al., 2004). In the design of transgenic and knock-in animals, multicistronic systems are used to express ectopic proteins, such as fluorescent reporters or DNA recombinases, while preserving the expression of endogenous genes of interest, offering platforms for fate mapping or cell type-specific genetic manipulation at a cellular resolution (Livet et al., 2007; Zhang et al., 2016). Multicistronic vectors also enable stoichiometric expression of transgenes for the proper formation of multicomponent protein complexes and the balanced progression of multifactor processes, facilitating gene therapy and cellular lineage manipulation (Szymczak et al., 2004; Takahashi and Yamanaka, 2006).

Among several strategies to achieve eukaryotic multicistronic expression, the internal ribosome entry site (IRES) and 2A cleaving sequence are most widely employed. IRES is an RNA sequence that was initially discovered in poliovirus RNA (Pelletier and Sonenberg, 1988) and encephalomyocarditis virus (EMCV) RNA (Jang et al., 1988) and later also identified in eukaryotic genes (Sarnow, 1989). To initiate cap-independent translation, the IRES functions as an independent platform for recruiting the ribosome, allowing multicistronic expression of open reading frames (ORFs) in a single mRNA (Hellen and Sarnow, 2001; Komar and Hatzoglou, 2011). In contrast to canonical cap-dependent translation, which is initiated by the recruitment of initiation factors, IRES-mediated translation involves IRES-transacting factors (ITAFs) (Komar and Hatzoglou, 2011), with each specific IRES requiring its own combination of ITAFs. Among a variety of IRES sequences from diverse sources, the EMCV IRES is one of the most popular options for the construction of multicistronic vectors.

Self-cleaving 2A-like peptide is another popular element in multicistronic expression systems. In contrast to the internal ribosomal recruitment that is mediated by the IRES sequence, the ribosome skips the 2A-like sequence during the translation of an mRNA to produce two separate polypeptides. The 2A peptide was first discovered in foot-and-mouth disease virus (Ryan et al., 1991), followed by the discovery of 2A-like sequences in equine rhinitis A virus, porcine teschovirus-1, and theosea asigna virus (Szymczak et al., 2004). Neither a eukaryotic counterpart of 2A-like sequences nor a mechanistic understanding of 2A-mediated ribosomal skipping has been extensively explored. Instead, the value of 2A-like sequences in constructing multicistronic vectors has been widely appreciated. Among a variety of 2A-like sequences, porcine teschovirus-1 2A (p2A) has been shown to be cleaved with the highest efficiency in several mammalian cells and has been widely utilized (Kim et al., 2011).

Several studies have characterized key properties of multicistronic elements, including their codon usage, expression rate, cleavage efficiency, and order of genes, to establish the optimal usage of these elements (Kim et al., 2011; Liu et al., 2017; Martinez-Salas, 1999; Mizuguchi et al., 2000; Park et al., 2014). Although these previous studies demonstrated the principle of using multicistronic element-based multiple gene expression, the dynamics of multicistronic element-mediated expression remain largely unknown. The temporal dynamics of gene expression is increasingly considered important because not only the level of gene expression but also the tem-

poral pattern of gene expression, ranging from minutes to days, is known to play a critical role in gene function (Hafner et al., 2017; Storch et al., 2002; Zhang et al., 2014). Therefore, understanding the temporal dynamics of multicistronic element-mediated expression will provide an important basis for controlling dynamic gene expression patterns and their functions.

Here, we quantitatively monitored the expression profiles of two distinct fluorescent proteins, which were engineered to have short half-lives for the monitoring of fine temporal changes, bridged by one of two bicistronic elements, either EMCV IRES or p2A, using real-time fluorescent imaging of HEK-293T cells. We also examined the temporal correlations of IRES and p2A in the neuronal blastoma Neuro2a cell line. The expression profiles of the fluorescent proteins were analyzed in terms of correlation and causality based on cross correlation analysis and cross convergent mapping.

## MATERIALS AND METHODS

### Plasmids

The EMCV IRES sequence was derived from pLVX-IRES\_Puro (Clontech, USA). The self-cleaving 2A sequence of porcine teschovirus-1 (Kim et al., 2011) was synthesized by Vector-builder (USA). We selected fluorescent proteins with short maturation times (Evdokimov et al., 2006; Shaner et al., 2004) and fast refolding kinetics (Fisher and DeLisa, 2008). Fluorescent reporters were cloned by polymerase chain reaction (PCR) from plasmids containing turboGFP (Evrogen, Russia) and mCherry (Clontech) and were conjugated to a nuclear localization signal to facilitate the quantification of fluorescent intensity and a PEST motif to reduce the half-life of the reporter. All required components were incorporated into the pCMV-tag (Stratagene, USA) plasmid by step-wise overlapping PCR and were validated by sequencing.

### Cell culture

Materials for cell culture were obtained from Thermo Fisher Scientific (USA). HEK and Neuro2a cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS), 100 U/ml penicillin/streptomycin, 4 mM glutamine, and 1 mM sodium pyruvate in a humidified atmosphere containing 5% CO<sub>2</sub> at 37°C. For imaging, cells were seeded onto and maintained in glass-bottomed 35-mm dish (SPL Life Sciences, Korea). For transient transfection, plasmids were introduced into either HEK or Neuro2a cells with TransIT-X2 (Mirus Bio, USA), according to the manufacturer's instruction. During imaging, 5% FBS-containing medium was used to reduce cellular motility and cell division.

### Imaging system

Time lapse images were acquired using an LSM LIVE confocal microscope (Zeiss, Germany) equipped with a chamber suitable for maintenance of cultured cells (humidified atmosphere containing 5% CO<sub>2</sub> at 37°C). Using a 10× objective lens, images were acquired every 10 minutes at a single fixed point for at least 12 hours. Laser intensity and gain were selected based on the maximal levels that did not induce no-

ticeable photo bleaching after the imaging session.

### Quantification and analysis

We quantified the fluorescent signals from cells that were discernible throughout each imaging session. The center and region of interest (ROI) of the cell were tracked using the Circadian gene expression toolbox (Sage et al., 2010) implemented in Fiji (Schindelin et al., 2012). The average ROI intensity was normalized using the background intensity and then using the maximal value of the given cell from the entire imaging session. Convergent cross mapping (CCM) was performed using the rEDM package (Sugihara et al., 2012) implemented in R software (R Development Core Team, 2010). Data were plotted using the ggplot2 package in R (Wickham, 2016).

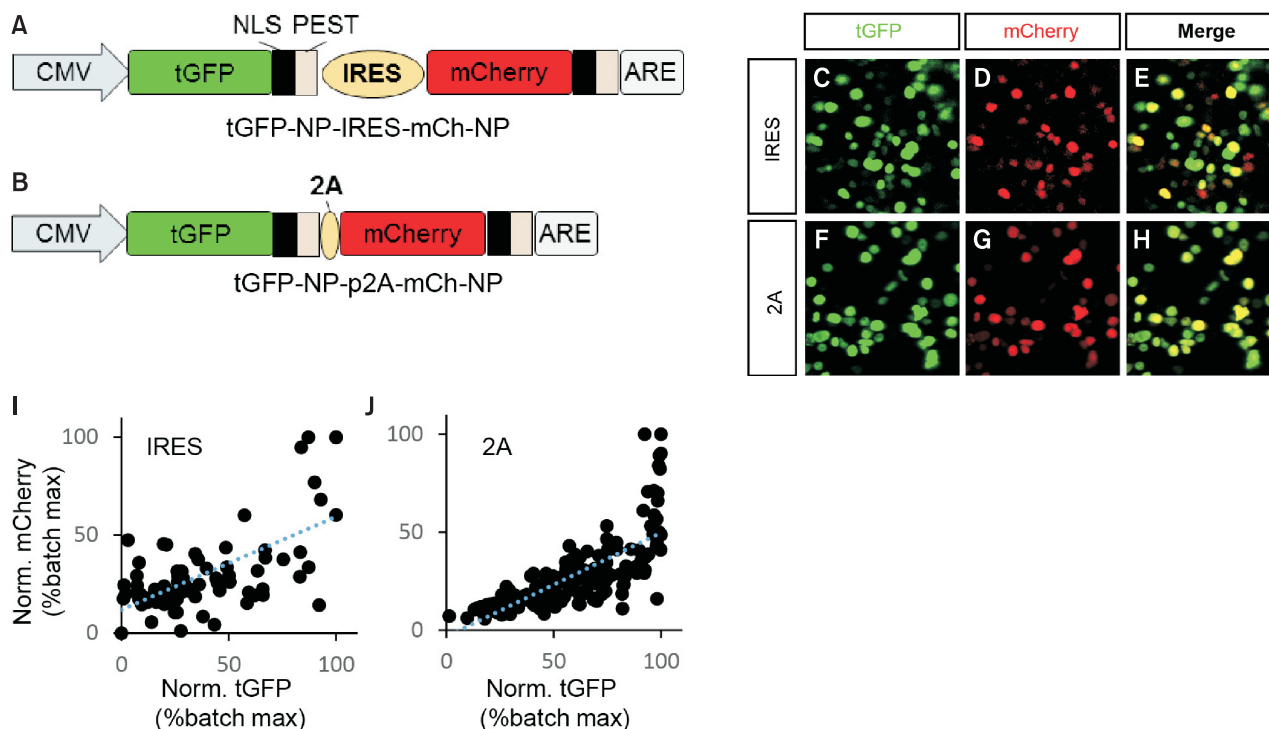
## RESULTS

To compare the temporal dynamics of bicistronic reporter expression using EMCV IRES and p2A sequences, we generated constructs that each expressed two independent fluorescent reporters, destabilized nuclear turboGFP (tGFP-NP) and destabilized nuclear mCherry (mCh-NP), linked by either EMCV IRES (Fig. 1A; tGFP-NP-IRES-mCh-NP) or the p2A sequence (Fig. 1B; tGFP-NP-p2A-mCh-NP). The transcription of tGFP-NP and mCh-NP, connected by either IRES or p2A, was driven by the CMV promoter, while the translation of the fluores-

cent proteins was initiated by cap-dependent translation or a cap-independent linker-dependent mechanism, respectively. Both constructs harbored an AU-rich element (ARE) in the 3' untranslated region (UTR) to destabilize the mRNA. Destabilization at both the mRNA and protein levels enables sensitive monitoring of dynamic and subtle changes in fluorescent levels.

To test the expression dynamics of the construct, we ectopically expressed either tGFP-NP-IRES-mCh-NP or tGFP-NP-2A-mCh-NP in HEK293 cells (Figs. 1C-1H). Epifluorescent microscopy revealed robust expression of both tGFP and mCh in each group. Most of the cells expressing tGFP co-expressed mCherry, with varying levels of expression. In order to assess the correlation between the tGFP signal and mCh signal, we quantified the intensities of the green and red fluorescence from the nuclei. Green and red fluorescent signals observed in both IRES- and 2A-connected constructs exhibited significant positive linear correlations (Figs. 1I and 1J) (IRES:  $P = 4.896 \times 10^{-10}$ , 81 cells from 3 batches; 2A:  $P = 2.2 \times 10^{-16}$ , 177 cells from 2 batches). Notably, the coefficient of determination ( $R^2$ ) was greater for the 2A-connected fluorescent reporters ( $R^2 = 0.4044$  for IRES and  $0.5606$  for 2A), suggesting that the 2A cleaving sequence may result in a more reliable stoichiometric relationship between connected ORFs than IRES.

We then utilized real-time fluorescent imaging to quantify the temporal profile of multicistronic element-mediated ex-



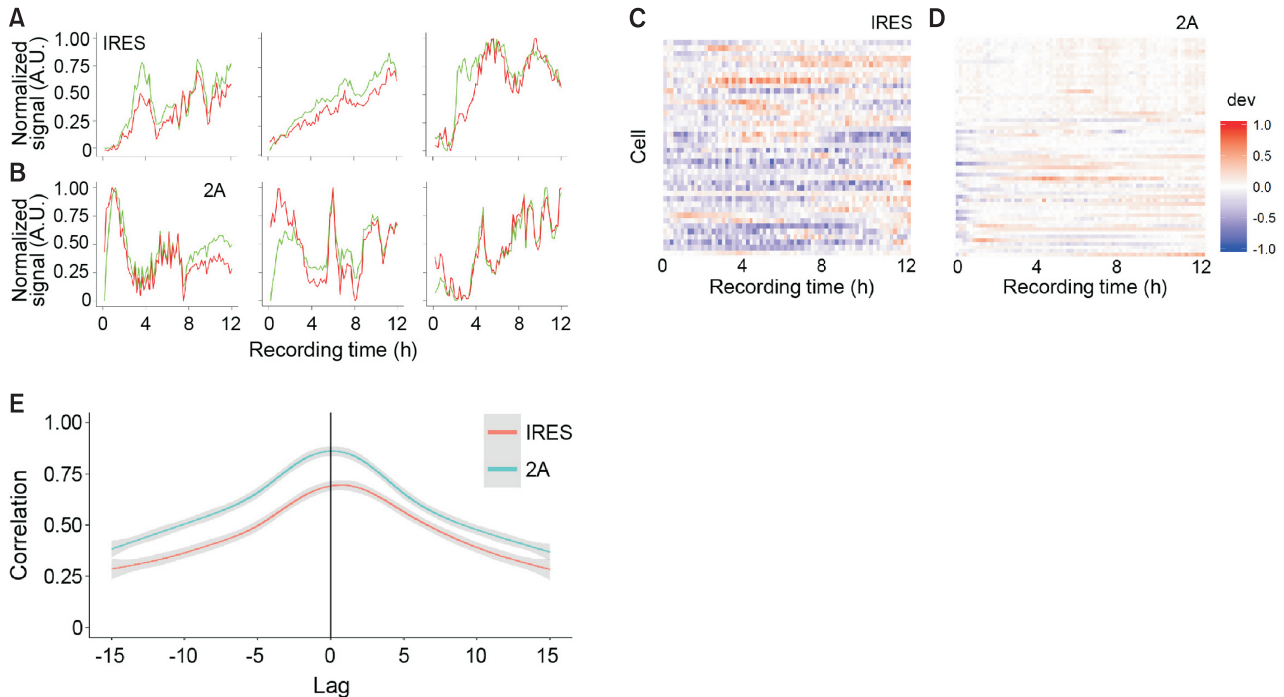
**Fig. 1. Design of constructs reporting the temporal dynamics of IRES- and p2A-mediated gene expression.** (A) Construct for IRES dynamics (tGFP-NP-IRES-mCh-NP). (B) Construct for 2A dynamics (tGFP-NP-2A-mCh-NP). (C-H) Representative fluorescent images of tGFP-NP-IRES-mCh-NP (C-E) and tGFP-NP-2A-mCh-NP (F-H) expressed in HEK293 cells. (I and J) Scatter plot of normalized (Norm.) tGFP and mCherry intensity connected by IRES (I) or 2A (J). Dashed blue line: linear regression line. NLS, nuclear localization sequence; PEST, destabilizing motif; CMV, CMV promoter; tGFP, turboGFP; ARE, AU-rich element.

pression of dual fluorescent proteins. The fluorescent levels of HEK293 cells expressing tGFP-NP-IRES-mCh-NP were monitored in real-time for 12 hours (Fig. 2A, Supplementary Video S1). Under the control of the CMV promoter, the tGFP signal in each cell exhibited a distinct dynamic profile without obvious regularity, presumably driven by stochastic and burst-like transcription (Arnaud et al., 2015; Yunger et al., 2010). The pattern of the mCh signal showed a dynamic profile similar to that of the tGFP signal but did not completely match the dynamics of tGFP. We then examined the fluorescent profiles of HEK293 cells expressing tGFP-NP-2A-mCh-NP (Fig. 2B, Supplementary Video S2). The profile of tGFP also exhibited a similar dynamicity to that of the tGFP signal from the IRES-connected bicistronic vector. In addition, the patterns of mCh and tGFP in each cell more closely matched each other when connected by 2A than when connected through IRES. To visualize the deviation of the mCh signal from the tGFP signal, we plotted the difference between the normalized tGFP signal and the normalized mCh signal in a heatmap (IRES: n = 40 cells from 2 batches; 2A: n = 60 cells from 3 batches) (Fig. 2C). As expected based on the representative profiles of tGFP and mCh connected by either IRES or 2A (Figs. 2A and 2B), the heatmap clearly demonstrated a stronger deviation between tGFP and mCh following IRES-mediated expression than following 2A-mediated expression. The direction of a deviation between the two fluorescent report-

ers is bidirectional rather than unidirectional. Moreover, the absolute difference between the two normalized fluorescent signals did not monotonically change in one direction. These suggest that the results are due to characteristics of EMCV IRES and p2A, but not due to either the differential decay rate between fluorescent reporter proteins or the intrinsic difference in synthesis rate between cap-dependent and multicistronic element-mediated expression.

To further analyze the temporal correlation between the tGFP and mCh signal in IRES- and 2A-connected constructs, we calculated the cross-correlation between the normalized fluorescent signals obtained from live imaging (Fig. 2E). The correlation between the tGFP and mCh signal was consistently higher in the profiles of 2A-connected constructs than in those of IRES-connected constructs at all examined lags. Notably, the highest correlation coefficient between the fluorescent signals was found at lag 0 in both the IRES- and 2A-connected groups, suggesting that the translation of the second ORF, following either IRES or 2A, may not experience a detectable delay, at least when measured at a 10-minute interval. Together, these observations suggest that the 2A sequence provides a more robust reflection of the temporal pattern of the preceding ORF than the IRES sequence.

We next asked whether the temporal correlation of 2A-mediated bicistronic expression ubiquitously surpassed that of IRES-mediated expression in other cell lines. To ad-

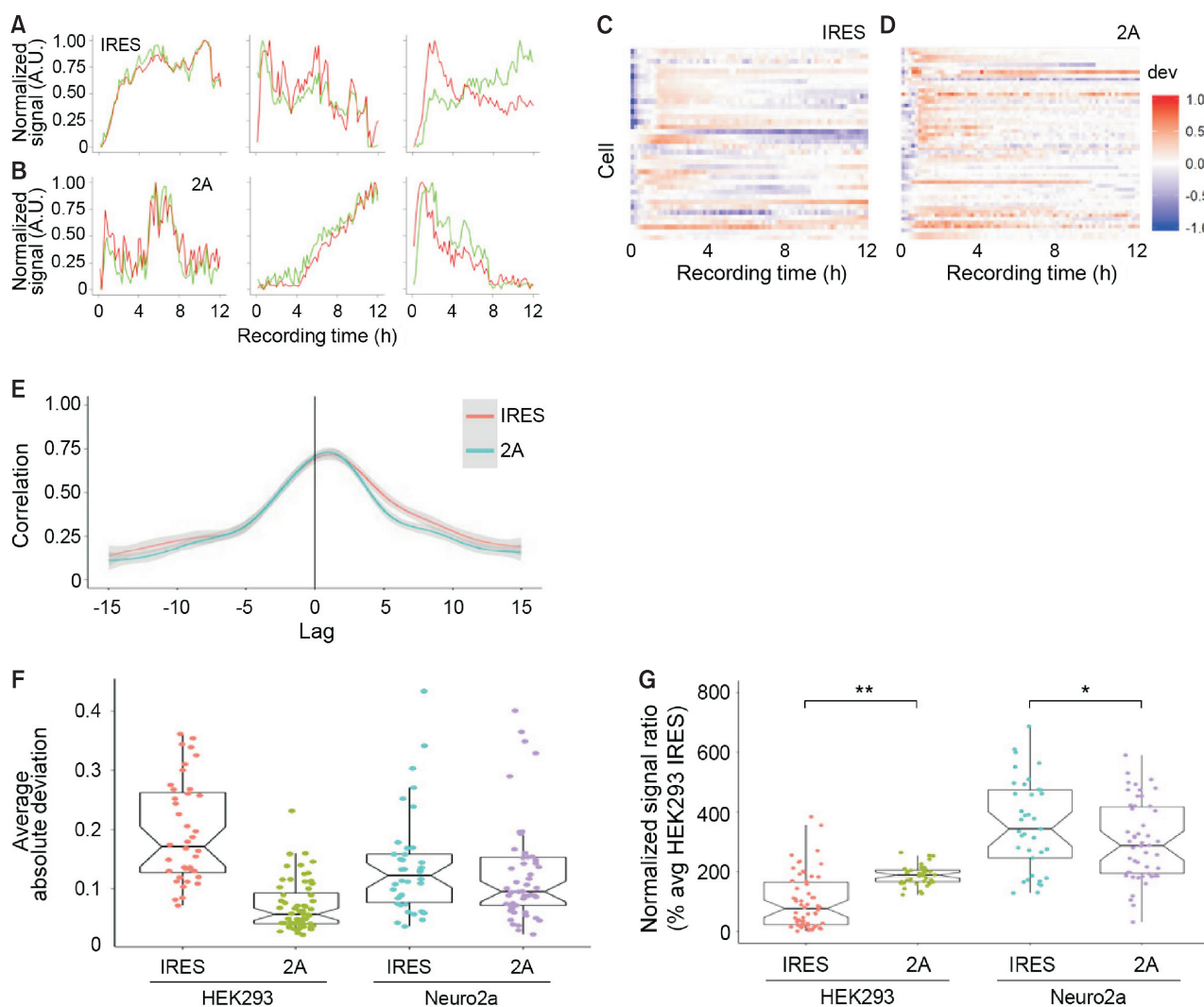


**Fig. 2. Quantitative analysis of temporal patterns of IRES- and 2A-mediated expression in HEK293 cells.** (A) Representative profiles of normalized tGFP (green) and mCh (red) expression from IRES bicistronic vector (tGFP-NP-IRES-mCh-NP). (B) Representative profiles of normalized tGFP (green) and mCh (red) expression from 2A bicistronic vector (tGFP-NP-2A-IRES-mCh-NP). (C and D) Heatmap showing the deviation of mCh signal from tGFP signal at the single-cell level following IRES- (C) and p2A-mediated (D) expression. Each row represents a cell. Deviation (dev) is color-coded according to the pseudocolor table on the right. (E) Cross-correlation diagram between tGFP and mCh signals. Red line: IRES-mediated expression. Blue line: 2A-mediated expression. Gray shadow: standard error. n = 40 cells from 2 batches for IRES, 60 cells from 3 batches for 2A.



dress this question, we monitored the reporter expression from tGFP-NP-IRES-mCh-NP and tGFP-NP-2A-mCh-NP in the Neuro2a neuroblastoma cell line. Similar to the reporter expression dynamics observed in HEK293 cells, tGFP expression in the Neuro2a exhibited fluctuations without any obvious regularity (Figs. 3A and 3B, Supplementary Video S3 and S4). In the case of IRES-mediated expression, the profile of mCh closely reflected that of tGFP. The 2A-mediated expression pattern of mCh also resembled that of tGFP, while the similarity between the mCh and tGFP profiles in Neuro2a cells

was not as high as that in HEK cells. A deviation plot clearly revealed that the disparity between 2A-mediated expression and IRES-mediated expression was minimal (Figs. 3C and 3D). Then, we analyzed the temporal correlation between the tGFP and mCh signal in IRES- and 2A-connected constructs (Fig. 3E). The correlation between the tGFP and mCh signal in Neuro2a was similar in the profiles of IRES-mediated constructs and those of 2A-mediated constructs. Consistent with HEK293, the highest correlation coefficient between the fluorescent signals was found at lag 0 in both the IRES- and

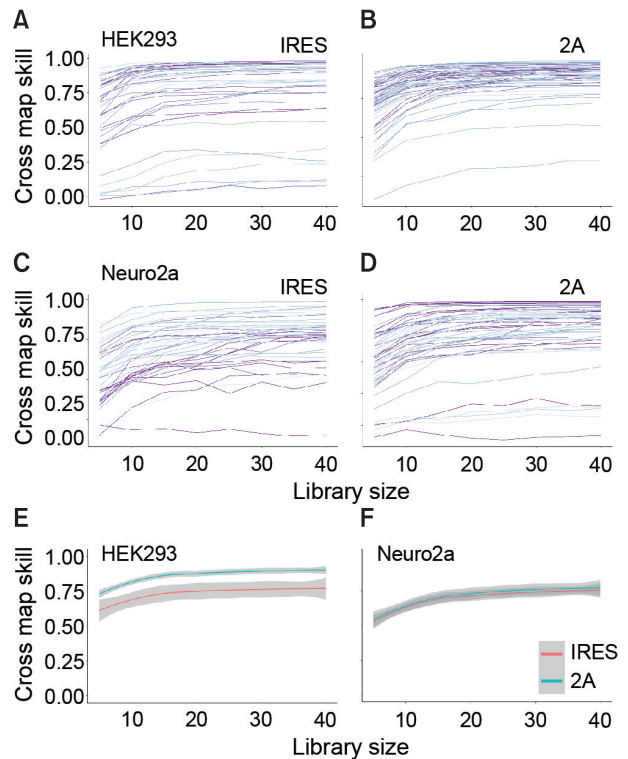


**Fig. 3. Quantitative analysis of temporal patterns of IRES- and 2A-mediated expression in Neuro2a cells.** (A) Representative profiles of normalized tGFP (green) and mCherry (red) expression from IRES bicistronic vector (tGFP-NP-IRES-mCh-NP). (B) Representative profiles of normalized tGFP (green) and mCherry (red) expression from 2A bicistronic vector (tGFP-NP-2A-IRES-mCh-NP). (C and D) Heatmap showing the deviation of mCherry signal from tGFP signal at a single-cell level following IRES- (C) and 2A-mediated (D) expression. Each row represents a cell. Deviation (dev) is color-coded according to the pseudocolor table on the right. (E) Cross-correlation diagram between tGFP and mCh signals. Red line: IRES-mediated expression. Blue line: 2A-mediated expression. Gray shadow: standard error. n = 38 cells from 2 batches for IRES, 52 cells from 3 batches for p2A. (F) Average absolute deviation of IRES- and 2A-mediated expression in HEK293 and Neuro2a cells. (G) Normalized signal ratio of mCh over tGFP of IRES- and 2A-mediated expression in HEK293 and Neuro2a cells. Lower and upper hinges: first and third quartiles. Lower and upper whiskers: smallest and largest value within 1.5 folds of the interquartile range from the hinge. Notch: a 95% confidence interval for median. \* $P < 0.05$ ; \*\* $P < 0.01$  by Tukey honestly significant difference test.

2A-connected groups.

To quantitatively compare the deviation between two reporters, we calculated average absolute deviation of each cell (Fig. 3F). Two-way ANOVA revealed significant effects in type of multicistronic element and in interaction between type of multicistronic element and cell line (effect of interaction:  $F_{1,186} = 24.923$ ,  $P < 0.001$ ; effect of multicistronic element:  $P < 0.001$ ; effect of cell line:  $P > 0.05$ ). As expected from deviation heatmaps, average absolute deviation of IRES-mediated expression was significantly higher than 2A-mediated expression only in HEK293 (adjusted  $P < 0.001$ ), but not in Neuro2a (adjusted  $P = 0.9126$ ). We also calculated the ratio of average mCh intensity over the average tGFP intensity (Fig. 3G). The two-way ANOVA revealed significant effects in the type of cell line and in the interaction between types of multicistronic elements and cell lines (effect of cell line:  $F_{1,186} = 123.6$ ,  $P < 0.001$ , effect of interaction:  $F_{1,186} = 20.97$ ,  $P < 0.001$ ; effect of multicistronic element:  $F_{1,186} = 0.3026$ ,  $P = 0.5829$ ). The pairwise *post hoc* comparison indicated that the ratio of 2A-mediated expression was significantly higher than that of IRES-mediated expression in HEK293 cells, and that the ratio of IRES-mediated expression was significantly higher than that of 2A-mediated expression in Neuro2a cells, but not as much as the difference observed in HEK293 cells. These analyses suggest that 2A-mediated bicistronic expression reflects the expression dynamics of the preceding gene at least as well as IRES-mediated expression.

To examine whether the expression dynamics of the preceding gene causally affect those of the subsequent fluorescent reporter, we performed CCM of tGFP and mCh profiles (Sugihara et al., 2012). When time-series data causally affect another time-series, the first time-series leaves a trace on the latter time-series that can be utilized to predict the first time-series from the latter time-series. As this relationship is causal, predictions are more precise (higher cross map skill), when more data points are used in the prediction (library size). We plotted the predictive skill of mCh profiles in cross mapping tGFP profiles versus the library size for all cells analyzed in Figs. 2 and 3. For the majority of HEK293 cells (39 out of 40 cells), the cross map skill of mCh expression profiles mediated by IRES monotonically increased as the library size increased (Fig. 4A). Similarly, the cross map skill of the 2A-mediated mCh expression profiles in all examined cells (60 cells) monotonically increased with library size (Fig. 4B). In the case of Neuro2a cells, both IRES- and 2A-mediated mCh expression profiles exhibited similar ranges of cross map skills (Figs. 4C and 4D), with the cross map skills of all but one cell in each group increasing monotonically as the library size increased. On average, however, although the predictive skills of both IRES- and 2A-bridged fluorescent signals increased monotonically as the library size increased, the level of cross map skill was one standard error higher in 2A-mediated expression than in IRES-mediated expression in HEK293 cells (Fig. 4E). In contrast, the average predictive skills of the expression profiles of both multicistronic elements were similar in Neuro2a cells (Fig. 4F). Thus, the CCM of IRES- and 2A-mediated expression suggests that, while there is a causal relationship in both IRES- and 2A-bridged constructs, the causal link is tighter in 2A-bridged constructs in certain cell



**Fig. 4. Causality analysis of IRES- and p2A-mediated translational dynamics.** (A-D) Convergent cross mapping skill of mCherry signal in mapping tGFP signal at a single-cell level. Each line represents a single cell observed among IRES-connected (A) and 2A-connected constructs (B) in HEK293 cells or IRES-connected (C) and 2A-connected constructs (D) in Neuro2a cells. (E and F) Average cross mapping skill of IRES- and 2A-mediated expression in HEK293 cells (E) and Neuro2a cells (F). Red line: IRES-mediated expression. Blue line: p2A-mediated expression. Gray shadow: standard error.

types.

## DISCUSSION

We compared the expression dynamics mediated by the multicistronic elements EMCV IRES and p2A using real-time imaging of two short-lived fluorescent reporters, tGFP-NP and mCh-NP. In HEK293 cells, the temporal profiles between tGFP expression and mCh expression were more highly correlated in 2A-bridged plasmids than in IRES-bridged plasmids. In contrast, IRES- and p2A-mediated expression in Neuro2a cells did not exhibit noticeable differences in the correlation between the two fluorescent reporters. CCM analysis, which examines the causal link between two time-series data sets, revealed a consistent pattern, showing better predictive efficiency of the 2A linker in HEK293 cells but similar efficiencies in Neuro2a cells. Based on the examination of two cell lines, we suggest that 2A-bridged multicistronic expression reflects the dynamics of the preceding ORF at least as well as EMCV IRES-bridged expression. Utilizing the experimental and analytic platform that we established here, further investigation

promises to reveal the temporal dynamics of multicistronic element-mediated expression in various cell lines under a variety of physiological conditions.

Although there are differences in expression dynamics mediated by multicistronic elements, it should be noted that EMCV IRES- and 2A-mediated expression profiles exhibited causal relationships between tGFP-NP expression and mCh-NP expression in both HEK293 and Neuro2a cells (Fig. 4). In parallel, correlation coefficients between the reporters indicated that the time-series of tGFP-NP and mCh-NP expression is quite similar without a detectable lag, at least at a 5-minute resolution (Figs. 2E and 3E). This suggests that reporter expression bridged by multicistronic elements may reflect the essence of the expression dynamics of the preceding endogenous gene at the level of tens-of-minutes to hours. This highlights the value of real-time monitoring of IRES- or 2A-sequence-mediated reporters as faithful surrogate markers in gene expression dynamics at a single-cell level. Indeed, it has been well demonstrated that 2A-mediated luciferase (F2A-dsLuc2) robustly reports circadian rhythms at a single-cell level (Suter et al., 2011). If properly supported by computational methods, this technology can be used to estimate the parameters of stochastic gene expression.

However, the matching of multicistronic element-mediated expression to the gene expression profiles of co-expressed genes may be limited by several factors. First, multicistronic element-mediated expression may not reflect abundance dynamics if the stability of the genes of interest differs. While these elements direct the stoichiometric matching of the rate of translation, the abundance of the protein is determined not only by translation but also by post-translational processing including degradation. Therefore, IRES- or 2A-like sequence-bridged designs will provide temporally matched expression of connected genes only when the half-lives of the genes are within comparable ranges. Second, the translation rate initiated by multicistronic elements may be affected by physiological conditions. IRES is not only found in viral genomes but is also widely found in eukaryotic genomes. IRES-dependent translation is directed by ITAFs, as well as eukaryotic translation initiation factors (eIFs) (Komar and Hatzoglou, 2011), with the set of ITAFs and eIFs required to initiate IRES-directed translation varying among IRES sequences. The ratio of cap-dependent translation to IRES-dependent translation is thus affected by various physiological and pathological conditions, including cellular stress, nutrient status, cell proliferation, circadian rhythm and differentiation (Balvay et al., 2009; Kim et al., 2010; Komar and Hatzoglou, 2011). Therefore, translation mediated by different IRES sequences will be differentially affected by cellular conditions. To our knowledge, the regulatory mechanism underlying 2A-directed cleavage through ribosomal skipping has not yet been identified, leaving room for cellular state-dependent variations in 2A-like sequence-mediated multicistronic expression as well. Further investigation of the dynamic properties of subtypes of IRES and 2A-like sequences is required to establish optimal multicistronic elements independent of physiological and pathological variations.

Thus far, we have discussed the values and limits of dynamic multicistronic expression mediated by EMCV IRES and

p2A. Although these elements cannot provide universally invariant synchronization of gene expression in a multicistronic system, EMCV IRES and p2A provide relatively well-correlated and causal relationships of temporal expression of connected genes under normal growth conditions. Progress in gene therapy, genome editing, and genetic approaches require reliable multicistronic elements to mediate finely tuned temporal and spatial expression of a set of genes (Daigle et al., 2018; Li et al., 2018; Liu et al., 2017; Szymczak et al., 2004). Further exploration of multicistronic elements based on the platform presented here will facilitate the identification of optimal multicistronic element-based system for the synchronous expression of genes of interests.

*Note: Supplementary information is available on the Molecules and Cells website ([www.molcells.org](http://www.molcells.org)).*

## ACKNOWLEDGMENTS

This work was supported by grants (NRF-2017R1C1B2008775, NRF-2017R1A4A1015534, and NRF-2018M3C7A1022310) from the National Research Foundation of Korea (NRF) of the Ministry of Science and ICT, and by KBRI basic research program through Korea Brain Research Institute funded by the Ministry of Science and ICT (17-BR-04).

## ORCID

Soomin Lee <https://orcid.org/0000-0002-5896-0771>  
Han Kyong Choe <https://orcid.org/0000-0001-7849-7094>

## REFERENCES

- Arnaud, O., Meyer, S., Vallin, E., Beslon, G., and Gandrillon, O. (2015). Temperature-induced variation in gene expression burst size in metazoan cells. *BMC Mol. Biol.* 16, 20.
- Balvay, L., Soto Rifo, R., Ricci, E.P., Decimo, D., and Ohlmann, T. (2009). Structural and functional diversity of viral IRESes. *Biochim. Biophys. Acta* 1789, 542-557.
- Bouabe, H., Fässler, R., and Heesemann, J. (2008). Improvement of reporter activity by IRES-mediated polycistronic reporter system. *Nucleic Acids Res.* 36, e28.
- Daigle, T.L., Madisen, L., Hage, T.A., Valley, M.T., Knoblich, U., Larsen, R.S., Takeno, M.M., Huang, L., Gu, H., Larsen, R., et al. (2018). A suite of transgenic driver and reporter mouse lines with enhanced brain-cell-type targeting and functionality. *Cell* 174, 465-480.e22.
- Evdokimov, A.G., Pokross, M.E., Egorov, N.S., Zaraisky, A.G., Yampolsky, I.V., Merzlyak, E.M., Shkorporov, A.N., Sander, I., Lukyanov, K.A., and Chudakov, D.M. (2006). Structural basis for the fast maturation of Arthropoda green fluorescent protein. *EMBO Rep.* 7, 1006-1012.
- Fisher, A.C., and DeLisa, M.P. (2008). Laboratory evolution of fast-folding green fluorescent protein using secretory pathway quality control. *PLoS One* 3, e2351.
- Hafner, A., Stewart-Ornstein, J., Purvis, J.E., Forrester, W.C., Bulyk, M.L., and Lahav, G. (2017). p53 pulses lead to distinct patterns of gene expression albeit similar DNA-binding dynamics. *Nat. Struct. Mol. Biol.* 24, 840-847.
- Hellen, C.U., and Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* 15, 1593-1612.
- Jang, S.K., Kräusslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., and Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J. Virol.* 62, 2636-2643.

- Kim, D.Y., Woo, K.C., Lee, K.H., Kim, T.D., and Kim, K.T. (2010). hnRNP Q and PTB modulate the circadian oscillation of mouse Rev-erb alpha via IRES-mediated translation. *Nucleic Acids Res.* *38*, 7068-7078.
- Kim, J.H., Lee, S.R., Li, L.H., Park, H.J., Park, J.H., Lee, K.Y., Kim, M.K., Shin, B.A., and Choi, S.Y. (2011). High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* *6*, e18556.
- Komar, A.A., and Hatzoglou, M. (2011). Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. *Cell Cycle* *10*, 229-240.
- Li, X.L., Li, G.H., Fu, J., Fu, Y.W., Zhang, L., Chen, W., Arakaki, C., Zhang, J.P., Wen, W., Zhao, M., et al. (2018). Highly efficient genome editing via CRISPR-Cas9 in human pluripotent stem cells is achieved by transient BCL-XL overexpression. *Nucleic Acids Res.* *46*, 10195-10215.
- Liu, Z., Chen, O., Wall, J.B.J., Zheng, M., Zhou, Y., Wang, L., Ruth Vaseghi, H., Qian, L., and Liu, J. (2017). Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci. Rep.* *7*, 2193.
- Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* *450*, 56-62.
- Martinez-Salas, E. (1999). Internal ribosome entry site biology and its use in expression vectors. *Curr. Opin. Biotechnol.* *10*, 458-464.
- Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E., and Hayakawa, T. (2000). IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Mol. Ther.* *1*, 376-382.
- Park, K., Jeong, J., and Chung, B.H. (2014). Live imaging of cellular dynamics using a multi-imaging vector in single cells. *Chem. Commun. (Camb.)* *50*, 10734-10736.
- Pelletier, J., and Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* *334*, 320-325.
- R Development Core Team. (2010). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).
- Ryan, M.D., King, A.M., and Thomas, G.P. (1991). Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *J. Gen. Virol.* *72* (Pt 11), 2727-2732.
- Sage, D., Unser, M., Salmon, P., and Dibner, C. (2010). A software solution for recording circadian oscillator features in time-lapse live cell microscopy. *Cell Div.* *5*, 17.
- Sarnow, P. (1989). Translation of glucose-regulated protein 78/immunoglobulin heavy-chain binding protein mRNA is increased in poliovirus-infected cells at a time when cap-dependent translation of cellular mRNAs is inhibited. *Proc. Natl. Acad. Sci. U. S. A.* *86*, 5795-5799.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* *9*, 676-682.
- Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* *22*, 1567-1572.
- Storch, K.F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F.C., Wong, W.H., and Weitz, C.J. (2002). Extensive and divergent circadian gene expression in liver and heart. *Nature* *417*, 78-83.
- Sugihara, G., May, R., Ye, H., Hsieh, C.H., Deyle, E., Fogarty, M., and Munch, S. (2012). Detecting causality in complex ecosystems. *Science* *338*, 496-500.
- Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., and Naef, F. (2011). Mammalian genes are transcribed with widely different bursting kinetics. *Science* *332*, 472-474.
- Szymczak, A.L., Workman, C.J., Wang, Y., Vignali, K.M., Dilioglou, S., Vanin, E.F., and Vignali, D.A. (2004). Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat. Biotechnol.* *22*, 589-594.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* *126*, 663-676.
- Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (New York: Springer-Verlag New York).
- Yunger, S., Rosenfeld, L., Garini, Y., and Shav-Tal, Y. (2010). Single-allele analysis of transcription kinetics in living mammalian cells. *Nat. Methods* *7*, 631-633.
- Zhang, H., Pu, W., Tian, X., Huang, X., He, L., Liu, Q., Li, Y., Zhang, L., He, L., Liu, K., et al. (2016). Genetic lineage tracing identifies endocardial origin of liver vasculature. *Nat. Genet.* *48*, 537-543.
- Zhang, R., Lahens, N.F., Ballance, H.I., Hughes, M.E., and Hogenesch, J.B. (2014). A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 16219-16224.