Evaluation of information flows in the RAS-MAPK system using transfer entropy measurements

Nobuhisa Umeki[a], Yoshiyuki Kabashima[b,c], Yasushi Sako[a,*]

[a]Cellular Informatics Laboratory, RIKEN, Cluster for Pioneering Research
2-1 Hirosawa, Wako 351-0198, Japan
[b]Institute for Physics of Intelligence, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Toko 113-0033, Japan
[c]Trans-Scale Quantum Science Institute, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Toko 113-0033, Japan

[*]To whom correspondence may be addressed. Email: sako@riken.jp

The RAS-MAPK system plays an important role in regulating various cellular processes, including growth, differentiation, apoptosis, and transformation. Dysregulation of this system has been implicated in genetic diseases and cancers affecting diverse tissues. To better understand the regulation of this system, we employed information flow analysis based on transfer entropy (TE) between the activation dynamics of two key elements in cells stimulated with EGF: SOS, a guanine nucleotide exchanger for the small GTPase RAS, and RAF, a RAS effector serine/threonine kinase. TE analysis allows for model-free assessment of the timing, direction, and strength of the information flow regulating the system response. We detected significant amounts of TE in both directions between SOS and RAF, indicating feedback regulation. Importantly, the amount of TE did not simply follow the input dose or the intensity of the causal reaction, demonstrating the uniqueness of TE. TE analysis proposed regulatory networks containing multiple tracks and feedback loops and revealed temporal switching in the reaction pathway primarily responsible for reaction control. This proposal was confirmed by the effects of a MEK inhibitor on TE. Furthermore, TE analysis identified the functional disorder of a SOS mutation associated with Noonan syndrome**,** a human genetic disease, of which the pathogenic mechanism has not been precisely known yet. TE assessment holds significant promise as a model-free analysis method of reaction networks in molecular pharmacology and pathology.

Significance

Intracellular signal processing relies on intricate regulatory networks composed of various molecular reactions. Ordinary differential equation (ODE) modeling, which proposes the reaction formula and parameter values, is a conventional approach to analyzing the mechanism of reaction networks. In contrast, statistical analyses focus on the correlations among the reactions and complement ODE modeling. Among them, TE analysis is advantageous in dealing with the regulation of the reaction by capturing the dynamics of causality between the reactions. In this study, we analyzed the TE in the cell signaling pathway. The results demonstrate the usefulness of TE assessment as a model-free network analysis method.

Keywords

RAS/MAPK system / Signal processing / Single-cell analysis / Transfer entropy

Introduction

Assessing the functions of intracellular reaction networks is a central focus of contemporary molecular cell biology. Significant progress has been made through biochemical and genomic analyses, revealing the networks' elements and their interconnections. Analytical models based on coupled ordinary differential equations (ODEs) are widely used to explore the most probable network structure and determine reaction parameter values. These models offer mechanistic insights into the network and facilitate predictions of molecular activation dynamics (Omony, 2014; Travassoly et al., 2018). However, due to the networks' molecular diversity and structural complexity, the credibility of analytical models remains partially uncertain in many cases. It would be beneficial to have a method to confirm the results of ODE modeling from a different perspective. In addition, it may be important to note that the average concentration of activated molecules estimated in ODE models does not necessarily imply the ability to control downstream molecules. Control capability refers to how differences in the upstream components are faithfully read out by the downstream in the presence of large noises in the reactions and environments. Unfortunately, ODE models typically do not account for reaction fluctuations.

Statistical models complement analytical ones (Waltermann and Klipp, 2011). Statistical analyses are fundamentally model-free, making them robust when information about the network structure and reaction formulas is limited. Results obtained from information analyses based on statistics are expressed in universal units, such as bits or nats, making them suitable for comparison. These units are independent of the specific properties and functions of individual molecular elements within the network. Mutual information (MI) and transfer entropy (TE) are examples of measures used in information analyses (Schreiber, 2000; Gencaga et al., 2015). TE, in particular, possesses the power to quantify the information flow, which reflects the causal relationships between different reaction time courses (Granger, 1969; Barnett et al., 2009). When estimating the future intensity of the reaction from its past intensity, the estimation accuracy could be improved by referring to the past intensity of another reaction. TE implies this improvement. (See Supplementary text for a more detailed explanation.) Since TE primarily means the control capability of upstream behavior on downstream components, the TE time course provides insights into the dynamics of signal transduction from the downstream effect perspective.

In practice, though MI analysis has found various applications in examining snapshots of cell signaling reactions (Uda et al., 2013; Miyagi et al., 2021; Fuchsberger et al., 2023), the use of TE assessment for studying the dynamics of real intracellular reaction networks has been significantly limited up to the present time. This is primarily because conventional TE calculation methods require extensive simultaneous data sequences and substantial computational power. The use of TE in biology is generally limited to simulation studies (Pahle et al., 2008; Kamberaj and van der Vaart, 2009; Ursino et al., 2020) and neural response analysis, which have high reproducibility and low noise (Wibral et

al., 2011; Stetter et al., 2012). There have been only a few applications of TE to study noisy intracellular chemical reactions (Pahle et al., 2008; Ito and Sagawa, 2015). However, we recently have reported a method for TE assessment applicable to relatively small numbers of time sequences (Imaizumi et al. 2022). This new approach involves utilizing a Gaussian approximation of the response distributions, enabling the rapid evaluation of TE using covariance matrices (Fig. S1). Additionally, we have developed a method for result evaluation using bootstrapping. By applying these techniques, we detected significant levels of TE between SOS and RAF proteins following epidermal growth factor (EGF) stimulation, utilizing several hundred single-cell simultaneous response time courses.

SOS and RAF are essential components of the EGFR-RAF-MAPK system, which plays a crucial role in regulating cell fate (Lemmon and Schlessinger, 2010). Upon activation of the EGF receptor (EGFR), tyrosine phosphorylations occur in its cytoplasmic domain, leading to the translocation of a protein complex comprising GRB2 and SOS from the cytoplasm to the plasma membrane, where they recognize the EGFR phosphorylations. SOS functions as a guanine-nucleotide exchanger for RAS, a small GTPase located on the cytoplasmic side of the plasma membrane. The translocation of SOS induces the activation of RAS, which, in turn, results in the translocation of RAF, a serine/threonine kinase and an effector of RAS, from the cytoplasm to the plasma membrane. These intracellular translocations of SOS and RAF can be visualized using a total internal fluorescence microscope with fluorescently labeled SOS and RAF molecules, specifically observing near the basal surface of cells (Nakamura et al., 2017; Yoshizawa et al., 2021).

Given the significance of RAS as a key regulator of various cellular responses, understanding the signal transduction from SOS to RAF, from the upstream to downstream events involved with RAS activation, is of interest. Human cancer (Cai et al., 2019) and genetic diseases, such as Noonan syndrome (Aoki and Matsubara, 2013; Aoki et al., 2016), involve various mutations in SOS, which are expected to lead to the hyperactivation of the RAS-MAPK system. However, the specific molecular mechanisms underlying these SOS mutations remain largely unknown. In a previous study (Imaizumi et al., 2022), we observed that one of the Noonan mutations of SOS (R1131K) affected the information flows from RAF to SOS. This finding raised the possibility that TE between SOS and RAF might provide insights into the functions of the RAS-MAPK system, yet detailed analysis is lacking. In the present study, we aim to investigate the properties of TE between SOS and RAF more precisely to demonstrate the potential of TE analysis for studying intracellular reaction networks and gain deeper insights into SOS/RAF signaling dynamics.

Results

Simultaneous measurement of SOS and RAF responses in single cells

To visualize intracellular translocations of SOS and RAF subtypes, hSOS1 and CRAF, hSOS1-Halo and GFP-CRAF (hereafter SOS and RAF, respectively) were expressed in HeLa cells, in which the intrinsic wild-type hSOS1 gene was knockout (Fig. 1A). Upon EGF stimulation, the phosphorylation of both intrinsic and GFP-fused RAF at S338 increased, indicating their activation (Fig. 1B). After the tetramethylrhodamine (TMR)-ligand conjugation to Halo fused to SOS, cells were stimulated with EGF, and changes in fluorescent SOS and RAF densities on the basal plasma membrane were quantified simultaneously in single cells through epi and total internal reflection (TIR) fluorescence microscopies (Fig. 1C and Movies S1 and S2). The intensities of translocation responses varied with the EGF dose (ranging from 0 to 100 ng/ml; Fig. 1D). To obtain the full data set, six independent measurements were performed for each EGF dose. For most measurements, the variation in daily means was less than 30% of the total cellular response range (Fig. S2). No significant translocation of SOS and RAF was observed in cells following vehicle stimulation (Figs. 1C black and S3A). To quantify SOS and RAF responses, the integration of response time profiles during 0-60 minutes of EGF stimulation was measured (Fig. 1E). The RAF response exhibited higher saturation with varying EGF doses compared to that of SOS.

Transfer entropy between SOS and RAF dynamics

The TE between the response sequence of SOS and RAF was evaluated both in the directions from SOS to RAF (forward; fw) and from RAF to SOS (backward; bw). In the fw (bw) TE calculation, we refer to the SOS (RAF) sequence as the causal sequence, X(t), and the RAF (SOS) sequence as the result sequence, Y(t) (Fig. 2A). The TE from a fixed causal time point (Tc) to the result sequence was calculated along t, concerning the distribution of Y at a fixed reference time point (Tr). In this study, TE calculation was performed for Tc = Tr < t. The 3D distribution of SOS and RAF at Tr, Tc, and t was unimodal and fairly symmetric in its main bodies (Fig. S4), allowing the Gaussian distribution as a first approximation.

TE time courses were calculated varying Tc (=Tr) from 1–35 min, using all measurements at EGF concentrations of 0.1-100 ng/ml (Fig. 2B). Statistically significant levels of TE were detected in both fw and bw directions. Statistically insignificant TE values were close to zero in our cases and were removed from the TE time course plots. Values of MI between X(Tr) and X(t) and Y(Tc) and X(t) were statistically significant for all time points Tc (=Tr) and t, and their significance thresholds were smaller than the TE values (Fig. S5). The detections of TE were not due to the limited number of data sequences, and the bootstrap mean values of TE were stable to the data number (Fig. S6). TE values obtained after vehicle stimulation were one order of magnitude lower than those obtained after EGF stimulation (Fig. S2B). The detection of significant levels of bw-TE suggests the presence of at

least one feedback loop from RAF with a significant ability to control the SOS response and/or the presence of a common input to SOS and RAF. (See 'Discussion' and 'Stochastic ODE simulations' in the supplement and Fig. S11 for this point.)

The time courses of TE show that TE from a Tc in the early stages of the cell response was transmitted as multiple peaks. For example, in the fw direction, TE from Tc = 1 min formed two peaks at t = 4 and 25-35 min. Between these two peaks (t ~ 15 min), fw-TE almost disappeared. Also in the bw direction, TE from Tc = 1~10 min formed two peaks before 20 min and after 30 min. The second peaks were not large but statistically significant. The values of bw-TE were insignificant or small during t = 20–30 min. Such separation of TE peaks from a single time point implies indirect information transfer resulting in the second peak (Fig. 2C). In a reaction cascade without branches, information from the upstream reaction is conveyed to the downstream reaction as a single peak (Fig. 2Ca). No TE should be observed after the first information flow has ended (t > Te). Multiple TE peaks indicate that the information from Tc was once transferred to some hidden states and passed to the late periods in the result sequence as the later peaks (Fig. 2Cb). During the period between the end of the first peak and the beginning of the second peak from the hidden state ($t = t_1–t_2$), TE disappears. The hidden states can be generated through a feedback loop (Fig. 2Cc). The exchanges of fw- and bw-TE between SOS and RAF were observed in practice. The fw (or bw)-TE from the peak time of the previous bw (or fw)-TE time course generated a peak from which the next round of fw (or bw)-TE peaks were generated (Fig. S7). This finding supports that at least a portion of TE was alternatively transferred between SOS and RAF using a feedback loop.

In our calculation, at least three and two TE peaks were observed at different periods in the fw and bw directions, respectively. We named the three fw peaks as early, middle, and late fw peaks, and the two bw peaks as early and late bw peaks. To briefly capture the characteristics of the multiple-peak TE time courses, we use the sum of TE amounts (TEsum) during the periods of TE peaks in the typical calculation conditions. The representative Tc (= Tr) and the range of t for each peak used in the evaluations of TEsum are shown in Figure 2D. The TEsum values were estimated as the average of 100 bootstrap runs. Although the accuracy of TEsum estimation depends on the number of data sets, estimation errors can be reduced by taking the average of bootstrapping (Fig. S6B). The average converged to nearly identical values (< 4% differences) for > 500 data sets.

Dependency on the EGF dose

We examined the EGF dose dependencies of the SOS and RAS response intensities, as well as the TEsum values resulting from these responses, by analyzing the reaction time courses obtained under varying EGF concentrations (Fig. 1D). The intensity of the SOS and RAF responses was defined as the integration of the molecular translocation during the periods of TE peaks, as used for the TEsum calculations. The response intensities at the early fw- and bw-peaks exhibited saturable and sigmoidal

profiles to the EGF dose in a log scale, respectively, as is commonly observed for intracellular reactions (Fig. 3A). The dose dependencies at the later peaks were bell-shaped with a faint peak at 10 ng/ml EGF.

The time course of TE for each EGF dose examined consisted of multiple peaks (Fig. S8) as observed in the calculations using the mixed data (Fig. 2B). The TEsum values calculated at each peak in the TE time courses showed a sharp bell-shaped EGF dose dependency in every case (Fig. 3B). This is in contrast to the EGF dose dependencies of the response intensities (Fig. 3A). The maximum values in the TEsum were often observed at the EGF doses distinct from those that elicited the maximum response intensity, even though the response and TE originated from the same causal reaction. This result shows that TE analysis provides different information than what is obtained through typical reaction response analyses.

To investigate the effect of EGF on TE between SOS and RAF, we compared the TEsum amounts conditioned and non-conditioned to the EGF dose (Vakorin et al., 2009). First, TEsum was calculated for each EGF dose, and then the conditional TEsum was calculated as the average of the TEsum values at different doses (Fig. 3C). The TEsum calculated using the mixed data under all EGF concentrations represents the non-conditional TEsum (Fig. 2D). To evaluate the significance of differences between conditional and non-conditional TEsums, we conducted 100 time bootstrapping trials, calculated the difference between the TEsum for each random data sampling, and obtained the average and 5-95% confidence range (Fig. 3D). Statistically significant differences were observed in the early fw- and bw-peaks. In these peaks, conditional TEsum exceeded non-conditional TEsum. This result was unexpected because, in the expected pathway of the tandem reaction from the EGF/EGFR complex to RAF via SOS, the non-conditional TE that includes the uncertainty of the most upstream element (EGF) should be equal to or greater than the TE that is conditioned to EGF, which mostly rejects the uncertainty of the EGF dose (Supplement text 'TE and conditional TE'). The larger conditional TEsum suggests that there was a bypass involving EGF/EGFR in the regulation of RAF response outside the well-known EGFR/SOS/RAF pathway (Fig. 3Ea left and Supplement text 'Stochastic ODE simulations and Fig. S11). Similarly, the SOS response was regulated by EGF/EGFR outside of the bw pathway from RAF that did not involve EGF/EGFR (Fig. 3Eb left). These bypasses mainly functioned in the early phase (< 15–20 min) of EGF signaling. Another possibility is that SOS and RAF were regulated by EGF/EGFR through parallel pathways, causing a pseudo-correlation between the SOS and RAF reaction time courses (right in Fig. 3Ea, b).

Response/transfer entropy relationships

TE in our calculation was derived from the SOS (for fw-TE) or RAF (for bw-TE) response at Tc after EGF stimulation. To investigate the correlation between the intensity of the SOS or RAF response and the amount of TE directly, we classified the reaction time courses according to the response intensity

at Tc. The classification used the same bin size to avoid the effect of the dynamic range in the causal response. As a result, the number of data sets was not identical for each response class. To ensure reasonable numbers (> 150) of data sets for the calculation, data in the larger and smaller ends of the response distributions were not used (Fig. 4A). We then examined the response (Fig. 4B) and TE (Fig. 4C) profiles against the causal response intensity. We estimated the error range for the bootstrapped mean of TEsum from the calculations using different numbers of random samples from all datasets (Fig. S6B), and expected the means to be within 63–107% (early fw peak) or 83–104% (other peaks) range of the true values, when > 150 data sets were used.

The response intensity in the result sequence was either linearly or saturably increased with the response intensity at the causal time responsible for it (Fig. 4B). In contrast to the dose/response profiles (Fig. 3A), these response/response profiles showed no intermediate peak. This could be because the EGF effect is not visibly included in the response/response profiles, although we cannot exclude the possibility that they were affected by the truncation of the data sets. The uniqueness of TE analysis was again observed in the difference between the response/response profiles and the response/TE profiles. The response/TE profiles were also largely different from the EGF dose/TE profiles (Fig. 3B). Interestingly, in the late phase, larger causal responses produced only smaller TE in both fw and bw directions, even though the larger causal response induced a larger result response. This could be due to saturation in the response intensities.

Effects of a MEK inhibitor and a SOS mutation

Previous studies have reported the involvement of negative feedback loops in the EGFR-RAS-MAPK system from ERK to EGFR and SOS (Corbalan-Gracia et al., 1996; Saha et al., 2012). The detection of significant levels of bw-TE suggests that these feedback loops play a role in the information flow. These feedback loops may be the reactions producing the bw-TE and the multiple TE pathways (Fig. 3E). To investigate these possibilities further, TE analysis was conducted in cells pretreated with a MEK inhibitor, trametinib. MEK is a protein kinase that mediates activations of RAF and ERK. ERK phosphorylation (activation) caused by MEK activity was examined by changing the concentration of trametinib (Fig. 5A). In this experiment, 10 ng/ml EGF was used for cell stimulation. The EGF-induced ERK activation was inhibited to the basal level in cells pretreated with 10 nM MEK inhibitor. However, when translocations of SOS and RAF were measured in cells preincubated with 10 nM MEK inhibitor, the SOS and RAF responses were only minimally affected in comparison with the responses after vehicle preincubation (Figs. 5B and S5A). TE was calculated between SOS and RAF responses in these conditions (Figs. 5C and S5B). The MEK inhibitor significantly increased early bw-TE (Fig. 5D). This result indicates that the feedback from MEK suppresses bw-TE under normal conditions, thereby reducing the control capability of RAF on the SOS response in the early phase of signal transduction.

We also conducted a study on the effects of the R1131K mutation in SOS, which is found in Noonan syndrome patients, with and without the MEK inhibitor. The R1131K mutation had minimal impact on the sensitivity to the MEK inhibitor for ERK phosphorylation (Fig. 5A) and translocations of SOS and RAF (Figs. 5B and S5A) upon stimulation with EGF. Additionally, it had no significant effect on EGF-induced RAF activation (Fig. 1B). However, R1131K had a significant impact on the amounts of bw-TE (Fig. 5C, D). Similar to the MEK inhibitor, R1131K significantly increased the early bw-TE peak. In addition, R1131K specifically decreased the late bw-TE peak. The increase in the early bw-TE peak in cells with R1131K SOS was observed in both the presence and absence of the MEK inhibitor, while the decrease in the late bw-TE peak was observed only in the absence of the MEK inhibitor, meaning a complex relationship between the effects of the MEK inhibitor and R1131K.

Pathway and regulation of the information flow in the EGFR-RAS-MAPK system

The MEK inhibitor and R1131K SOS mutation affect the feedback from MEK/ERK and the SOS function, respectively. These two factors regulated bw-TE amount in combination (Fig. 6A). We observed that the regulation system changed according to the time of cell stimulation. In the early phase (< 20 min), dysfunction in at least one of the regulation mechanisms, i.e., the MEK inhibitor or the R1131K mutation, increased the bw-TE compared to the control condition. Whereas, in the late phase (> 25 min), only in the absence of the MEK inhibitor, the R1131K mutation decreased the bw-TE.

There could be many different reaction pathways for realizing these types of regulations. Here, we suggest two simple pathways for the early and late phase regulations, respectively (Fig. 6B). In the early phase, the TE level was normal only when the feedback regulation and the SOS protein were both normal, suggesting a tandem pathway of the two regulations. Based on our prior knowledge about the EGFR-RAS-MAPK system, the MEK/ERK should be upstream of SOS in the feedback path. There are two regulation points between MEK/ERK and the output from SOS affected by the MEK inhibitor and the R1131K mutation, respectively. Since the R1131K mutation affects the SOS function directly, its action point should be downstream. It is also known that the MEK/ERK feedback reduces SOS activity. Therefore, the MEK inhibitor increases the SOS activity. In Noonan syndrome patients, the activity of the RAS-MAPK system is generally enhanced, suggesting that the R1131K mutation also increases the SOS activity. The abnormal regulation at these two points by the MEK inhibitor and the R1131K mutation affects the output from SOS to increase the bw-TE. However, it should be noted that a higher SOS activity does not necessarily result in a higher TE level (as shown in Fig. 4). The amount of TE is only given in the correlation of the two reactions.

The late-phase regulation may seem odd, but it has a reasonable explanation: In the wt SOS cells, the feedback pathway is mostly disconnected in the late phase, so the MEK inhibitor has little effect on the TE values. The R1131K mutation reconnects the pathway, which restores the effect of

the MEK/ERK feedback to decrease the bw-TE amount. However, in the presence of the MEK inhibitor, the feedback loop gets strongly inhibited, and therefore R1131K does not significantly change the TE value. According to this model, the late-phase regulation operates similarly to the early-phase regulation. The only difference between them is the strengths of the connectivity.

The results of the conditional TE analysis to the EGF dose, suggest that there were multiple pathways for the TE. Some pathways include the EGF/EGFR complex, as well as SOS and RAF (Fig. 3E). The involvement of the EGF/EGFR complex in TE pathways is also supported by the effects of the MEK inhibitor and the R1131K SOS mutation. We constructed a multitrack reaction network that carries bw-TE, including the MEK/ERK feedback loop and activation state of the EGF/EGFR complex (Fig. 6D). The network for the early phase is consistent with that shown on the left side of Figure 3Eb. These pathways are discussed in the “Discussion” section.

Discussion

We conducted an assessment of transfer entropy (TE) between SOS and RAF activations in living cells stimulated with EGF. We have observed a significant amount of TE in both directions, from SOS to RAF and from RAF to SOS. There are two possibilities for the generation of bw-TEs: The existence of feedback loops from RAF to SOS and/or the existence of parallel input paths from hidden states to both the SOS and RAF. In both cases, the EGF/EGFR complex is likely to play an important role (Fig. 3Eb). It is important to note that TE not only signifies the existence of the reaction pathways but also their control capability. The EGFR-RAS-MAPK system is known to involve several feedback loops (Lake et al., 2016). For instance, ERK activation downstream of RAF induces EGFR threonine 669 phosphorylation, enhancing EGFR internalization and suppressing EGFR kinase activity (Li et al., 2008). ERK also phosphorylates several serine residues in the GRB2-binding domain of SOS (Rojas et al., 2011), weakening its interaction with GRB2 and reducing SOS translocation to the plasma membrane upon EGF stimulation (Corbalan-Gracia et al., 1996; Saha et al., 2012). Furthermore, desensitization of RAF to RAS signal has been reported as a result of multiple serine phosphorylations on RAF by ERK (Wartmann et al., 1997). These feedback loops likely carry the bw-TE. Our finding that the MEK inhibitor affected the amount of early bw-TE (Fig. 5) supports this likelihood. Besides the mentioned negative feedback loops, other factors would complicate the feedback pathway and contribute to the TE amounts, such as multiple membrane binding domains of RAF (Ghosh et al., 1996; Hibino et al. 2011), RAF phosphorylation on the plasma membrane for activation (Lavoie and Therrien, 2015), and the SOS/RAS positive feedback loop (Margarit et al., 2003).

After examining the time courses of TE, we found that they had several peaks in both forward and backward directions (Fig. 2). The second peak from a single causal time must be mediated by hidden states, which created bypasses for TE outside the pathway for the first peak. The conditional TE analysis also suggested that there were multiple tracks of information (Fig. 3E). One possibility is that some of the multiple tracks pass through EGF/EGFR complex. However, in the fw direction, signaling from SOS to EGFR skipping the MEK/ERK feedback pathway (Fig. 3Ea, left) has not been proposed as far as we know. A recent study by Lin et al. (2022) suggests that large complexes consisting of EGFR, GRB2, and SOS are formed by a liquid-liquid phase separation. This may provide a potential pathway for SOS to signal to EGFR. Further research is needed to investigate this possibility. Another possibility is that the EGF/EGFR complex regulates SOS and RAF in parallel. One of the domains in RAF, known as the cysteine-rich domain, has been found to interact with membrane lipids (Ghosh et al., 1996). The activation of EGFR can cause a membrane remodeling around EGFR (Abe et al. bioRxiv 2024), which could affect the association of RAF with the plasma membrane. The phosphorylation of RAF by kinase molecules on the plasma membrane (Lavoie and Therrien, 2015) could also be regulated by EGFR and affect the translocation of RAF. At the present time, the parallel network (Fig. 3Ea right) is more likely for the fw-TE.

The involvement of the EGF/EGFR complex in the bw-TE pathway was supported by the effects of the MEK inhibitor on TE (Figs. 5 and 6). A negative feedback from ERK to EGFR has been reported (Li et al. 2008). The combination of the MEK inhibitor and R1131K SOS mutation suggested a tandem regulation pathway for the bw-TE (Fig. 6C). The reaction network containing this regulation pathway (Fig. 6D, upper) is consistent with the left reaction network in Figure 3Eb. A comparable network appears to be utilized for the late bw-TE (Fig. 6D, lower). However, in the wt SOS cells, the regulation from the EGF/EGFR complex would be weakened in the late phase (as we will discuss later) and would not be detectable in the conditional TE analysis against the EGF dose. The decrease in the middle peak of fw-TE after the MEK inhibitor pretreatment in R1131K cells (Fig. S9) also suggests that there was a bw-TE pathway via EGF/EGFR in the early phase. There is another feedback pathway exists from ERK to SOS directly (Corbalan-Gracia et al., 1996; Saha et al., 2012). The MEK inhibitor and R1131K might affect this pathway regardless of the activity of EGF/EGFR. However, considering the time dependency found in the effect of EGF dose on the bw-TE (Fig. 3), it seems less likely that this regulation was primarily functional.

Next, we consider the time dependency in TE regulation. The extra pathways involving EGF/EGFR were effective in the early phase when the early fw- and bw-TE peaks were detected. A change in TE regulation occurred between 20–30 minutes of EGF stimulation (Fig. 6A), during which the bw-TE mostly disappeared (Fig. 2). During this period, the time profiles of SOS and RAF translocation were altered from the initial peak phase to the later plateau phase (Fig. 1D). The early phase in the SOS and RAF activation is responsible for transient MEK/ERK activation following EGF signaling (Marshall, 1995; Kao, 2001; Yoshizawa et al., 2021). In the late phase, dephosphorylation and/or endocytosis of EGFR reduces signaling from EGFR (Monast et al. 2012; Fig. S10). This decrease in the EGFR signal should be the reason for the insensitivity of the TE levels to the MEK inhibitor in the late phase of wt SOS cells (Fig. 5). The roles of negative feedback from ERK to EGF/EGFR would be less important after the reduction of EGFR activity.

The information flow was perturbed in Noonan cells carrying the R1131K mutation. The mutation had opposite impacts on the early and late phases in cells without the MEK inhibitor, i.e., it led to an increase in the early phase and a decrease in the late phase of bw-TE (Fig. 6A). These seemingly controversial effects can be explained by considering the time dependence in the EGFR signaling intensity, as mentioned above (Figs. 6D, E, and S10). Despite suggestions of the involvement of EGF/EGFR complex in the early fw-TE pathway (Fig. 3), R1131K did not have a significant effect on the early fw-TE (Fig. 5). Moreover, R1131K did not show any significant impact on the late fw-TE (Fig. 5). Assuming that the function of R1131K is to sustain the EGFR signaling to SOS by counteracting the negative feedback effect, it will not have any impact before the feedback. In the late phase, R1131K may not be strong enough to boost the weak EGFR signal and alter the quantity of fw-TE (Fig. 6E). The cell line used in this study expresses SOS2, which has a similar function to SOS1.

The function of SOS2 must cause part of the RAF translocation. Our results suggest that R1131K SOS1 has a dominant negative effect on SOS2 and that wt SOS1 at least partially regulates RAF translocation.

The molecular mechanism of onset for most mutants found in Noonan syndrome, including SOS R1131K, is still unknown. After conducting TE analysis, the R1131K mutation is highly possible to affect the EGFR to SOS pathway instead of the SOS to RAF pathway. In other words, the R1131K mutant appears to be defective in receiving input but not in sending output. Furthermore, it was observed that the abnormality caused by R1131K changed with time after cell stimulation. During the early phase, the increase in bw-TE in the R1131K cells competes with fw regulation in the SOS/RAF system. This will prevent the SOS response from precisely following the EGF input. Conversely, during the late phase, R1131K decreases the bw-TE, which relatively strengthens fw regulation. In the late phase, the SOS/RAF feedback loop appears to operate autonomously, outside of the EGF control (Fig. 3). If inhibition of SOS activation is dominant as the bw regulation, the positive fw and negative bw regulations for the response intensities likely form an adaptation cycle that stabilizes SOS and RAF activation levels (Ito and Sagawa, 2015). This type of network could explain the quasi-steady states observed during the late phase (Fig. 1D). The R1131K mutation disrupts the mutual regulation between SOS and RAF, which can cause the system being out of control. Such abnormality in the information flows caused by R1131K was difficult to detect using conventional biochemical analysis. However, the TE analysis makes it possible to detect bidirectional TE in the steady state without perturbing the system. The phase-dependent effect of R1131K could provide some insight into the medical treatment of Noonan syndrome. The R1131K mutation disrupted feedback regulation. MEK inhibition rescued the abnormality in the late phase but not in the early phase.

This study demonstrated that TE assessment allows for the analysis of intracellular reaction networks from the perspective of reaction control capability. It is important to note that the amount of TE is not necessarily determined by the strength of molecular activation, but rather, by the correlation between the two reactions in single cells. As a result, TE analysis provides a unique perspective of the network that is difficult to obtain with conventional analysis of reaction strengths. In this study, TE analysis successfully predicted the presence of feedback loops and multiple pathways combining SOS and RAF response dynamics without any perturbation to the system and free from any mathematical models. This means that TE analysis is hopeful to be a complement to the conventional ODE analysis. Additionally, TE analysis identifies disorders in reaction regulation caused by a chemical inhibitor and a pathogenic mutation. This shows the applicability of TE analysis to molecular pharmacology and pathology. It is essential to note that the behavior of TE is complex, and gaining more experience with TE analysis will further improve its application in understanding intracellular systems.

412 **Materials and Methods**

413 **Sample Preparations:**

414 Methods for the construction of SOS and RAF plasmids and the establishment of SOS1 knockout
415 HeLa cells have been previously described (Nakamura et al., 2017). HeLa cells were cultured in D-
416 MEM supplemented with 10% FBS at 37°C, 5% CO_2. Overnight, cells were transferred onto 25 mm-
417 ϕ coverslips for microscopy or on a 35 mm-ϕ cell culture dish for biochemistry and then transfected
418 with the plasmids using Lipofectamine 3000 (Thermo Fisher). The cells were cultured in MEM
419 containing 1% BSA for one day before being used for the experiments. In some experiments, cells
420 were pretreated with a MEK inhibitor, trametinib (ChemScene) in DMSO (final concentration 0.05%)
421 for 1 hr before EGF stimulation.

422

423 **Microscopy:**

424 Halo-SOS in cells on a coverslip was conjugated with 300 nM TMR Halo-tag ligand for 1 hr, washed,
425 and SOS and RAF translocations in cells were observed in MEM, 1% BSA, 10 mM HEPES (pH 7.4).
426 An in-house total internal reflection (TIR) fluorescence microscope system, based on IX83 inverted
427 fluorescence microscope (Olympus), was used for imaging. The microscope is equipped with a
428 multicolor laser diode (LDI, Chroma Technology) for fluorophore excitation, a 60x NA 1.5 oil-
429 immersion objective (UPLAPO60XOHR, Olympus), a dual-color image splitting system (W-VIEW
430 GEMINI-2C, Hamamatsu), and two cMOS cameras (ORCA-Flush4.0, Hamamatsu). Using a 2-
431 dimensional piezo mirror scanner (S-335, PI) in the excitation beam path, this system can alternatively
432 switch the excitation mode between TIR illumination with a circularly rotating laser beam at 40 Hz
433 and epi-illumination, covering a field number of 19 (320 μm in diameter under a 60x objective). At
434 each time point of the measurement, a set of TIR and epi-illumination images were acquired
435 successively at the same field of view under the excitation laser wavelengths of 470 nm for GFP and
436 555 nm for tetramethyl-rhodamine (TMR) simultaneously. The exposure time was 125 ms and 25 ms
437 for each frame in the TIR- and epi-illumination modes, respectively. Image acquisitions were repeated
438 with 1-minute intervals. A triple-band dichroic mirror (ZT405/470/555rpc, Chroma) was used in the
439 microscope. In the image splitting optics, the GFP and TMR signals were separated by a dichroic
440 mirror (T550lpxr-UF2, Chroma), and emission filters for GFP (ET510/40m, Chroma) and TMR
441 (ET575/50m, Chroma). Just before the 4th time of image acquisition (time 0), cells were stimulated
442 by adding 550 μl of observation medium containing EGF (Pepro Tech) to the 450 μl of the same
443 medium in the observation chamber. All measurements were performed at 25°C. With 10~13
444 observation fields, we obtained time-lapse movies of 20~100 cells in a single measurement.
445 Representative responses of cells are shown in Supplemental Movies S1-S4. Response curves in the
446 cells expressing wt SOS under 100 ng/ml EGF without MEK inhibitor were obtained under the same
447 conditions as in Imaizumi et al. (2022), but from independent measurements.

Image processing

After the shading and background correction, the outline of single cells in the epi-illumination images was determined using the "cellpose" algorithm (https://github.com/mouseland/cellpose), and the average intensities per pixel in the single cell areas were measured in each video frame. Data from cells with the lower 20% of signal intensities at time 0 were removed. From the single-cell time courses of the average intensity, bleed-through between the GFP and TMR fluorescence was subtracted, drift in the basal intensity with time was removed, then the TIR signal was normalized to the epi signal. The difference in the normalized TIR signal to the 3-frame average before stimulation was plotted as the single-cell response. Image processing was performed using Image J (NIH). The response dynamics data used in this study have been deposited to https://github.com/transer_entropy_2/.

Transfer entropy calculations

Transfer entropy (TE) values from SOS to RAF and RAF to SOS were calculated along the time courses of the cell responses. The source code for the calculations has been deposited to https://github.com/transfer_entropy_2./, which is a modification of https://github.com/kabashiy/transfer_entropy/. The method used in these programs is precisely described by Imaizumi et al. (2021).

In brief, when assessing the uncertainty of the stochastic time sequence of X at time t (X_t) by the statistical entropy $H(X_t)$, the amount of the uncertainty reduction after obtaining the past sequence of X at the reference time point Tr, $\boldsymbol{X}_{Tr}$ is $I(X_t;\boldsymbol{X}_{Tr}) = H(X_t) - H(X_t|\boldsymbol{X}_{Tr})$. Here, $H(X_t|\boldsymbol{X}_{Tr})$ is the conditional entropy, and I is mutual information. The TE from Y to X is defined as $TE_{Y\to X}^{Tc,Tr}(t) = I(X_t;\boldsymbol{X}_{Tr},\boldsymbol{Y}_{Tc}) - I(X_t;\boldsymbol{X}_{Tr})$, which represents the reduction of uncertainty about X_t, given the additional information from Y at the causal time point Tc, $\boldsymbol{Y}_{Tc}$, beyond what is already provided by $\boldsymbol{X}_{Tr}$. The calculation of TE from real experimental data is usually difficult due to insufficient data numbers to accurately estimate the multi-dimensional distribution of the generating function. Even if the data number is sufficient, a large amount of computational power is required. However, in cases where the joint distribution of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is multivariate Gaussian, TE can be analytically evaluated using covariance matrices for the Gaussian model as

$$TE_{Y\to X}^{Tc,Tr}(t) = \frac{1}{2}\ln\left(\frac{\sum(X_t|\boldsymbol{X}_{Tr})}{\sum(X_t|\boldsymbol{X}_{Tr}\oplus\boldsymbol{Y}_{Tc})}\right).$$

Here, $\sum(X)$ and $\sum(X,Y)$ denote the covariance matrix of X and the cross-covariance matrix of X and Y, respectively, and $\sum(X_t|\boldsymbol{X}_{Tr}) = \sum(X_t) - \sum(X_t,\boldsymbol{X}_{Tr})\sum(x_{Tr})^{-1}\sum(X_t,\boldsymbol{X}_{Tr})^{\mathrm{T}}$. $X\oplus Y$ refers to the concatenation of the matrices X and Y. Procedure of TE calculation from the measured data set is shown in Fig. S1. Gaussian approximation is the first approximation but the only practical way to realize TE calculation from noisy real data sets. Even in non-Gaussian cases, it is known that non-zero

TE values under the Gaussian approximation mean non-zero values of the true TE (Marinazzo et al. 2008).

To determine the significance threshold, we first generate a sample set by random sampling and calculate the TE with the null hypothesis, TE = 0. This means that we use a setting in which all elements of the cross-covariance between $X_t \oplus \boldsymbol{X}_{Tr}$ and $,\boldsymbol{Y}_{Tc}$ are vanished. Then, we obtain the distribution of TE values by resampling the data 1000 times. From this distribution, we adopt the upper 1% percentile point as the significant threshold. Finally, if the TE value exceeds this threshold, we conclude that the TE value is statistically significant. This procedure was repeated 100-times for the randomly sampled data, allowing duplication (bootstrapping), to obtain the average and 5-95% range. The number of sampled data was the same as the number of original data number unless otherwise noted.

Western blot analysis

Cells under the described conditions were solubilized and separated on 10% polyacrylamide gels, and transferred to polyvinylidene difluoride membranes (BD Biosciences). The membranes were incubated with an anti-SOS1 (#5890, Cell Signaling Technology), anti-RAF (#610151, BD Bioscience), anti-RAF pS338 (05-538, Millipore), anti-EGFR pY1068 (3777S, Cell Signaling Technology), or anti-β-action (#5441, Sigma), and a secondary antibody conjugated with horseradish peroxidase (#7074, #7076, Cell Signaling). Antibody bindings were visualized with ECL Prime Western Blotting Detection Reagent kit (Cytiva) and imaged by using ImageQuant LAS500 (GE Healthcare).

Numerical simulations

Stochastic simulations of reaction cascade models were performed using the program described in Shindo et al. (2018), where Langevin ODE models were solved using the 4th-order Runge-Kutta method. The Supplementary Text provides details of these models and the results of TE analyses of the models. The purpose of analytical modeling is to assess the behavior of TE conceptually, not to reproduce the real intracellular reaction network.

Abbreviations

bw-TE, backward TE; EGF, epidermal growth factor; EGFR, EGF receptor; fw-TE, forward TE; MI, mutual information; ODE, ordinary differentiation equation; TE, transfer entropy; TEsum, sum of the significant TE; TIR, total internal reflection; TMR, tetramethylrhodamine.

Acknowledgments

The authors thank Mutsumi Nakanishi and Hiromi Sato for technical assistance, Sinsuke Uda at Kyushu University, Atsushi Hatano at Nigata University, and Tomoyuki Obuchi at Kyoto University for helpful discussions and suggestions. The SOS knockout Hela cell was established by Mitsuhiro Abe at RIKEN. YS and YK were funded by JST CREST (JPMJCR1912). YS and YK were funded by Grants-in-Aid for Scientific Research, MEXT (19H05647) and (22H05117), respectively.

Author contributions: NU and YS designed the research; NU and YS performed the research; YK contributed analytic tools; NU, YK, and YS wrote the paper.

The authors declare no competing interest.

References

M. Abe, M. Yanagawa, M. Hiroshima, T. Kobayashi, Y. Sako, Bilateral regulation of EGFR activity and local PI dynamics observed with superresolution microscopy. bioRxiv/2024/578337 (2024).

Y. Aoki, Y. Matsubara, Ras/MAPK syndromes and childhood hemato-oncological diseases. *Int J Hematol* **97**, 30–6 (2013).

Y. Aoki, T. Niihori, S. Inoue, Y. Matsubara, Recent advances in RASopathies. *J Hum Genet* **61**, 33–9 (2016).

L. Barnett, A. B. Barrett, A. K. Seth, Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys. Rev. Lett.* **103,** 238701 (1-4) (2009).

D. Cai, P. S. Choi, M. Gelbard, M. Meyerson, Identification and characterization of oncogenic *SOS1* mutations in lung adenocarcinoma. *Mol. Cancer Res.* **17,** 1002-1012 (2019)

S. Corbalan-Garcia, S. S. Yang, K. R. Degenhardt, D. Bar-Sagi, Identification of the mitogen-activated protein kinase phosphorylation sites on human Sos1 that regulate interaction with Grb2. *Mol. Cell Biol.* **16**, 5674–82 (1996).

F. F. Fuchsberger, D. Kim, N. Baranova, H. Vrban, M. Kagelmacher, R. Wawrzinek, C. Rademacher, Information transfer in mammalian glycan-based communication. *eLife* **2023,** e69415 (1-32) (2023).

D. Gencaga, K. H. Knuth, W. B. Rossow, A recipe for the estimation of information flow in a dynamical system. *Entropy*, **17,** 438-470 (2015).

C. W. J. Granger, Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, **37,** 424–438 (1969).

S. Ghosh, J. C. Strum, V. A. Sciorra, L. Daniel, R. M. Bell, Raf-1 kinase possesses distinct binding domains for phosphatidylserine and phosphatidic acid. Phosphatidic acid regulates the translocation of Raf-1 in 12-O-tetradecanoylphorbol-13-acetate-stimulated Madin-Darby canine kidney cells. *J Biol. Chem.* **271,** 8472-8480 (1996).

K. Hibino, T. Shibata, T. Yanagida, Y. Sako, Activation kinetics of RAF protein in the ternary complex among RAF, RasGTP, and the kinases on the plasma membrane of living cells. *J. Biol. Chem.* **286,** 36460-36468 (2011).

T. Imaizumi, N. Umeki, R. Yoshizawa, T. Obuchi, Y. Sako, Y. Kabashima, Assessing transfer entropy from biochemical data. *Phys. Rev. E.* **105,** 034403 (2021).

S. Ito, T. Sagawa, Maxwell's demon in biochemical signal transduction with feedback loop. *Nat. Comm.* **6,** 7498 (2015).

H. Kamberaj, A. van der Vaart, Extracting the causality of correlated motions from molecular dynamics simulations, *Biophysical J.* **97,** 1747-1755 (200).

S. Kao, R. K. Jaiswal, W. Kolch, G. E. Landreth, Identification of the mechanisms regulating the differential activation of the mapk cascade by epidermal growth factor and nerve growth factor in

PC12 cells. *J. Biol. Chem.* **276,** 18169-18177 (2001).

D. Lake, S. A. L. Corrêa, J. Müller, Negative feedback regulation of the ERK1/2 MAPK pathway. *Cell. Mol. Life Sci.* **73,** 4397-4413 (2016).

H. Lavoie, M. Therrien, Regulation of RAF protein kinases in ERK signalling. *Nat. Rev. Mol. Cell Biol.* **16**, 281-298 (2015).

M. A. Lemmon, J. Schlessinger, Cell signaling by receptor tyrosine kinases. *Cell* **141,** 1117-1134 (2010).

D. Marinazzo, M. Pellicoro, S. Stramaglia, Kernel method for nonlinear granger causality. *Phys. Rev. Lett.* **100**, 144103 (2008).

X. Li, Y. Hung, J. Jiang, S. J. Frank, ERK-dependent threonine phosphorylation of EGF receptor modulates receptor downregulation and signaling. *Cell Signal.* **20,** 2145-2155 (2008).

C-W. Lin, Laura, M. Mocka, B. L. Stinger, J. B. DeGrandchamp, L. J. Nugent Lew, S. Alvarez, H. T. Phan, Y. Kondo, J. Kuriyan, J. T. Groves, A two-component protein condensate of the EGFR cytoplasmic tail and Grb2 regulates Ras activation by SOS at the membrane. *Proc. Natl. Acad. Sci. USA* **119,** e2122531119 (2022).

S. M. Margarit, H. Sondermann, B. E. Hall, B. Nagar, A. Hoelz, M. Pirruccello, D. Bar-Sagi, J. Kuriyan, Structural evidence for feedback activation by Ras·GTP of the Ras-specific nucleotide exchange factor SOS. *Cell* **112**, 685–695 (2003).

C. J. Marshall, Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* **80,** 179-185 (1995).

C. S. Monast, C. M. Furcht, M. J. Lazzara, Computational analysis of the regulation of EGFR by protein tyrosine phosphatases. *Biophys. J.* **102,** 2012-2021 (2012).

H. Miyagi, M. Hiroshima, Y. Sako, Cell-to-cell diversification in ERRB-RAS-MAPK signal transduction that produces cell-type specific growth factor responses. *Biosystems* **19,** 104293 (1-10) (2021).

Y. Nakamura, N. Umeki, M. Abe, Y. Sako, Mutation-specific mechanisms of hyperactivation of Noonan syndrome SOS molecules detected with single-molecule imaging in living cells. *Sci. Rep.* **7,** 14153 (2017).

J. Omony, Biological Network Inference: A Review of Methods and Assessment of Tools and Techniques. *Annu. Res. Rev. Biol.* **4**, 577–601 (2014).

J. Pahle, A. K. Green, C. J. Dixon, U. Kummer, Information transfer in signaling pathways: A study using coupled simulated and experimental data. *BMC Bioinformatics* **9,** 139 (2008).

J. M. Rojas, J. L. Oliva, E. Santos, Mammalian son of sevenless guanine nucleotide exchange factors: old concepts and new perspectives. *Genes Cancer* **2**, 298–305 (2011).

M. Saha, A. Carriere, M., Cheerathodi, X. Zhang, G. Lavoie, J. Rush, P. P. Roux, B. A. Ballif, RSK phosphorylates SOS1 creating 14-3-3-docking sites and negatively regulating MAPK activation.

Biochem. J. **447**, 159–66 (2012).

T. Schreiber, Measuring information transfer, *Phys. Rev. Lett.* **85,** 461-464 (2000).

Y. Shindo, Y. Kondo, Y. Sako, Inferring a nonlinear biochemical network model from a heterogeneous single-cell time course data. *Sci. Rep.* **8**, 6790 (2018).

O. Stetter, D. Battaglia, J. Soriano, T. Geisel, Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Comp. Biol.* **8** e1002653 (2012).

I. Travassoly, J. Goldfarb, R. Iyngar, Systems biology primer: The basic methods and approaches. *Essays Biochem.* **62,** 487-500 (2018).

S. Uda, T. N. Saito, T. Kudo, T. Kokaji, T. Tsuchiya, H. Kubota, Y. Komori, Y.-I. Ozaki, S. Kuroda, Robustness and compensation of information transmission of signaling pathways. *Science* **341,** 558-561 (2013).

M. Ursino, G. Ricci, E. Magosso, Transfer entropy as a measure of brain connectivity: A critical analysis with the help of neural mass models. *Front. Compt. Neurosic,* **14,** 45 (2020).

V. A. Vakorin, O. A. Krakovska, A. R. McIntosh, Confounding effects of indirect connections on causality estimation. *J. Neurosci. Meth.* **194,** 152-160 (2009).

C. Waltermann, E. Klipp, Information theory based approaches to cellular signaling. *Biochim. Biophys. Acta* **1810**, 924–932 (2011).

M. Wartmann, P. Hofer, P. Turowski, A. R. Sltiel, N. E. Hynes, Negative modulation of membrane localization of the Raf-1 protein kinase by hyperphosphorylation. *J. Biol. Chem.* **272,** 3915-3923 (1997).

M. Wibral, B. Rahm, M. Rieder, M. Lindner, R. Vicente, J. Kaiser, Transfer entropy in magnetoencephalographic data: Quantifying information flow in cortical and cerebellar networks *Prog. Biophys. Mol. Biol.* **103,** 80 (2011).

R. Yoshizawa, N. Umeki, A. Yamamoto, M. Okada, M. Murata, Y. Sako, p52Shc regulates the sustainability of ERK activation in a RAF-independent manner. *Mol. Biol. Cell* **32,** 1838-1848 (2021).

Figures

Figure 1. SOS and RAF responses to cell stimulation

A. Expression levels of SOS and RAF probes. Halo-SOS (wild type, wt or R1131K mutant, R/K) and GFP-RAF were expressed in SOS knockout HeLa cells. **B.** RAF phosphorylation before and after 10 min of 10 ng/ml EGF stimulation of cells. Results of representative Western blotting and quantification of three independent experiments. **C.** Translocations of SOS (magenta) and RAF (green). Epi-fluorescence (upper) and TIR-fluorescence (lower) images at the indicated times after 100 ng/ml EGF stimulation. Bar, 50 µm (Supplemental Movies S1 and S2.) **D.** Time course analysis of SOS and RAF translocation observed in the total 1467 cells after vehicle (black: 150 cells) or EGF (blue: 0.1 ng/ml, 300 cells; green: 1 ng/ml, 310 cells; orange: 10 ng/ml, 370 cells; red: 100 ng/ml, 337 cells) stimulation. Changes in the molecular density on the cell surface (TIR) relative to the whole cell signal (epi) are plotted. The lower panels show the average with SD of the single cell data in the upper panels. **E.** EGF dose dependence of the SOS (red) and RAF (green) responses. Response intensities were determined as the sum of the molecular translocation during 0–60 min normalized to the maximum. Inset shows the relationship between SOS and RAF responses. Cell-to-cell averages are shown with SE.

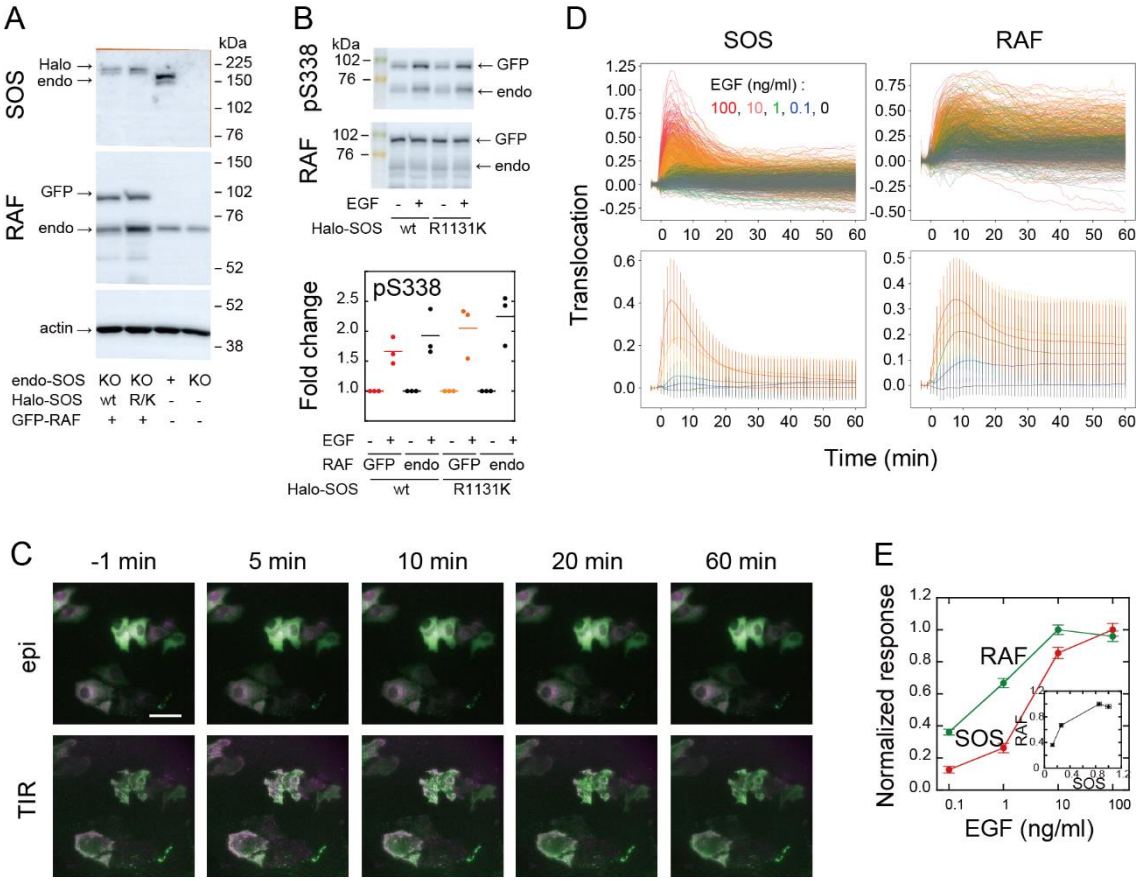
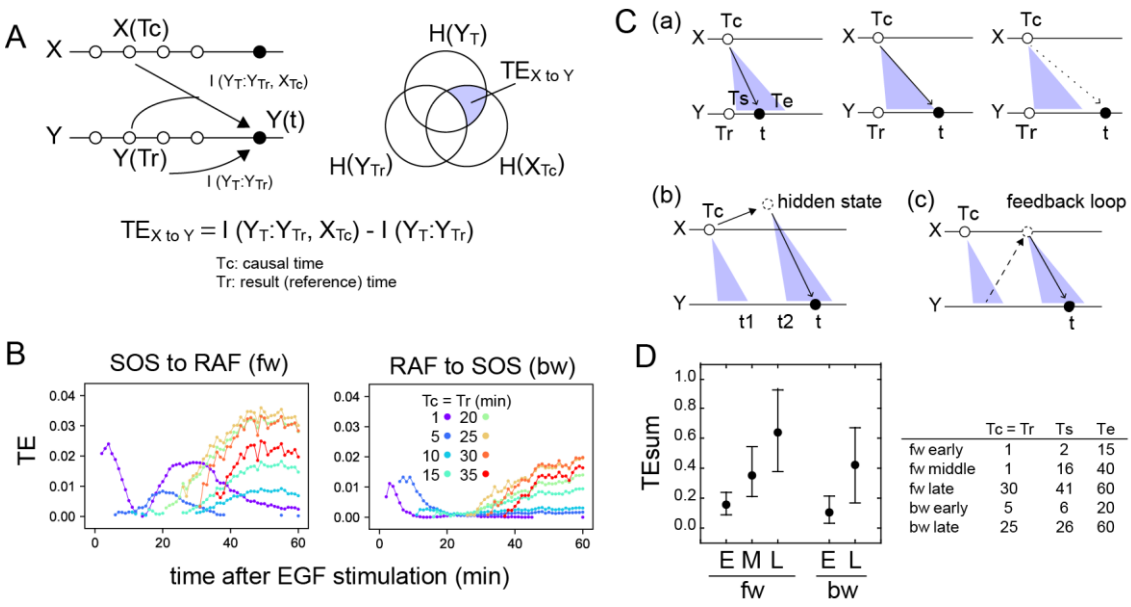


Figure 2. Transfer entropy time course analysis

A. Concept of TE between two reaction time courses. TE from X(Tc) to Y(t) was calculated using Y(Tr) as the reference (Tc = Tr < t). **B.** TE time courses for the total 1317 pairs of translocation trajectories with 0.1-100 ng/ml of EGF stimulation. Calculations were performed for the indicated Tc = Tr from 1 to 35 min after cell stimulation. Only statistically significant TE values (see Materials and Methods) are plotted. **C.** Appearance of the separated TE peaks from single Tc. (**a**) In a single-track reaction pathway, TE from Tc in the X time series appears during a single continuous period (Ts to Te) and disappears after $t \geq$ Te. (**b**) The second TE peak is indirect information flow through a hidden state outside the pathway of the first TE peak. (**c**) The second TE peak can be caused by a feedback loop. **D.** The sum of the significant TE (TEsum) in the TE peaks. Average and 5–95% range in 100-time bootstrapping are shown. TE from Tc (= Tr) was integrated during $t_s \sim t_e$ as TEsum. Tc, t_s, and t_e used in the calculations are shown in the list. E: early, M: middle, L: late.



	Tc = Tr	Ts	Te
fw early	1	2	15
fw middle	1	16	40
fw late	30	41	60
bw early	5	6	20
bw late	25	26	60

Figure 3. Conditional transfer entropy analysis to the EGF dose

A. EGF dose dependency of SOS and RAF responses at the TE peaks. Response time courses (Fig. 1D) were integrated during the TE peak periods listed in Figure 2D. The averages for single cells were plotted with SD. **B.** EGF dose dependency of TEsum for each TE peak. **C.** TEsum conditioned to the EGF dose. **D.** Difference between TEsums conditioned (**C**) and non-conditioned (Fig. 2D) to the EGF dose. **bootstrapping $p < 0.05$. **B–D.** The averages obtained from 100-time bootstrapping are shown with the 5–95% interval. **E.** Multiple TE tracks involving EGF/EGFR complex suggested form (**D)** (see text for details). Two possible cases are presented for both early fw-TE (a) and early bw-TE (b) peaks.

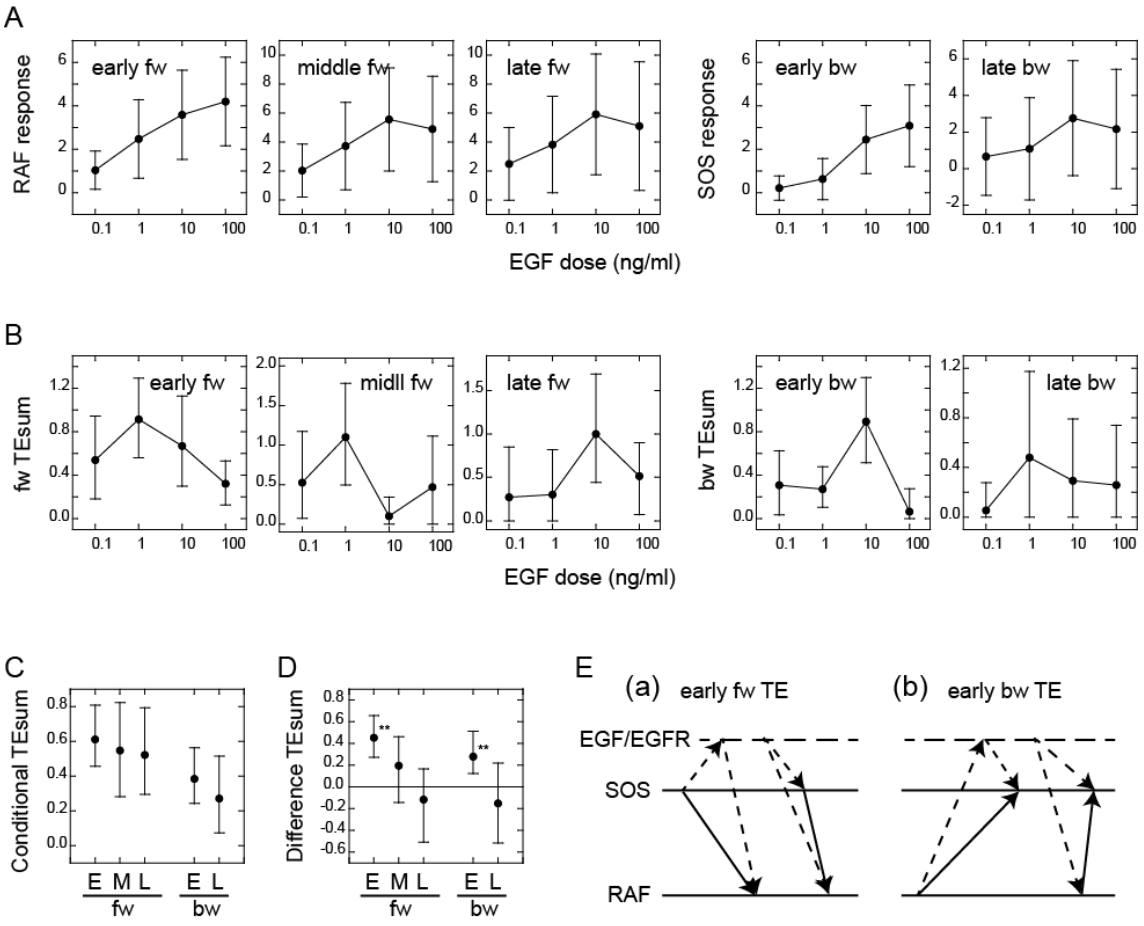


Figure 4. Response/transfer entropy relationships

Four groups of single-cell response time courses were extracted from the total data sets (Fig. 1D) according to the causal response intensity at Tc. **A.** Parameter list for data extraction. Each group had fixed bin sizes with > 150 data sets. Data in the low and high ends of distributions were not used. min, max: the minimum and maximum response intensities, respectively. n < min, n > max: the excluded data numbers in the low and high end for calculation, respectively. d: bin size. n1–n4: data numbers contained in each bin. nT: the total data number. **B.** Relationships between the SOS and RAF responses. The response intensities during the period of TE peaks in the result sequence are plotted as a function of the causal response intensity at Tc. Cell-to-cell averages are plotted with SD. **C.** Relationships between the causal response intensity and the corresponding TEsum. Averages from 100-time bootstrapping are plotted. Error bars show the estimated error range for the bootstrap average for n > 150 (Fig. S2).

A

Tc	min	n < min	max	n > max	d	n1	n2	n3	n4	nT
SOS 1 min	-0.1	2	0.26	226	0.09	175	558	202	154	1089
SOS 30 min	-0.06	90	0.14	112	0.05	200	401	336	178	1115
RAF 5 min	-0.03	21	0.37	204	0.1	325	300	278	189	1092
RAF 25 min	-0.04	32	0.36	158	0.1	193	441	341	152	1127

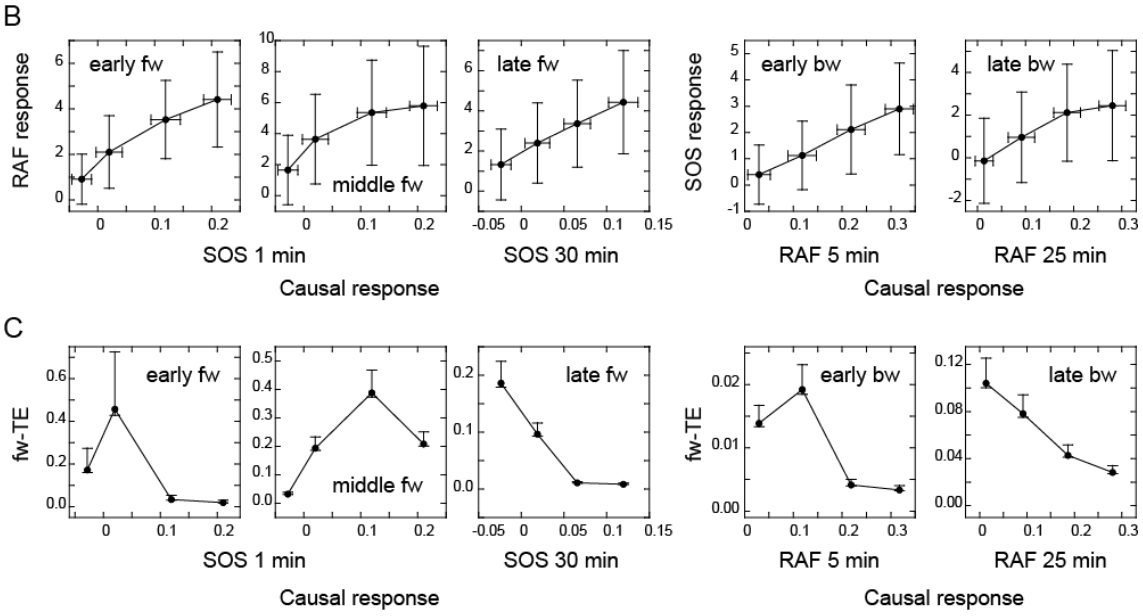


Figure 5. Effects of a MEK inhibitor and a Noonan SOS mutation

A. Inhibition of ERK phosphorylation in cells treated with a MEK inhibitor, trametinib. Typical results from western blotting (left) and the inhibitor dose dependency in four independent experiments (right) are shown. Bars indicate the averages. **B.** SOS and RAS responses in cells stimulated with 10 ng/ml EGF. Cells were expressing wt (upper) or R1131K mutant (lower) of SOS and pretreated with 10 nM MEK inhibitor (red) or vehicle solution (DMSO; blue). The numbers of cells were 704 (wt, MEK inhibitor), 744 (wt, vehicle), 832 (R1131K, MEK inhibitor), and 732 (R1131K, vehicle). Averages of cells are plotted with SD. See Fig. S5A for the single-cell time courses. **C.** TEsum for each peak in the indicated experimental condition. **D.** Difference between TEsum values under indicated condition and control condition (wt cells with vehicle stimulation). Positive values indicate smaller TEsums in the control condition. **bootstrapping $p < 0.05$. **C, D.** Averages from 100-time bootstrapping are shown with the 5–95% interval.

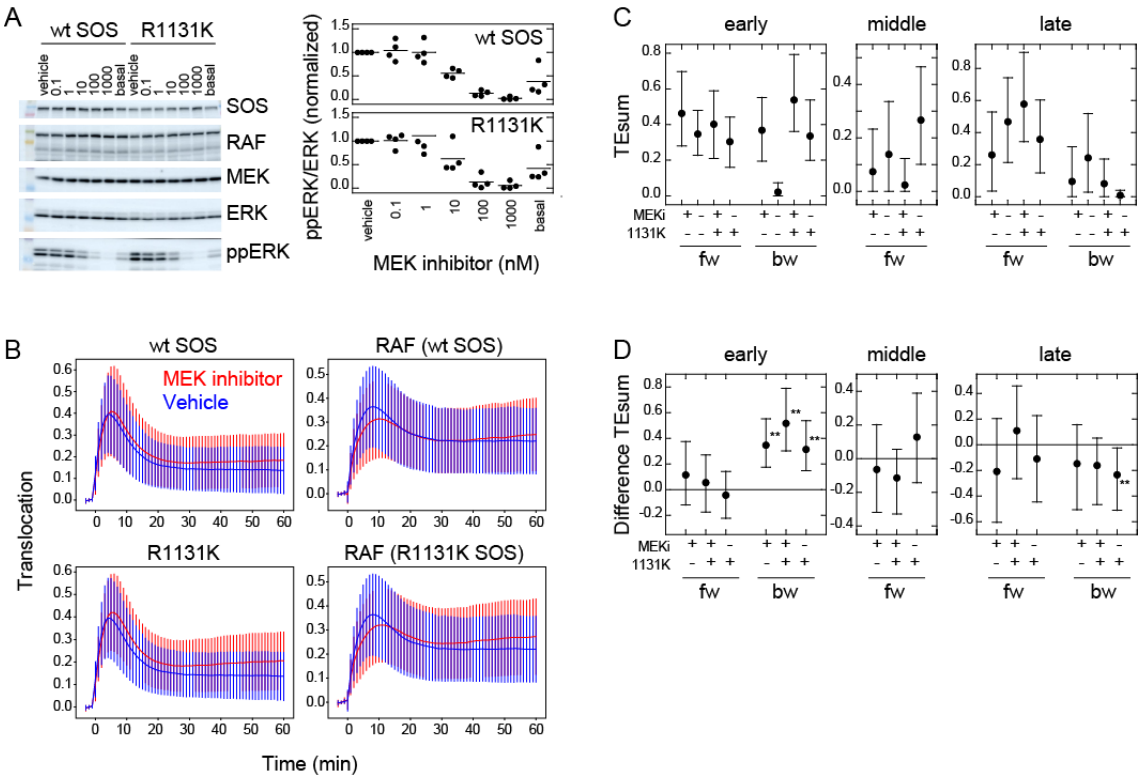
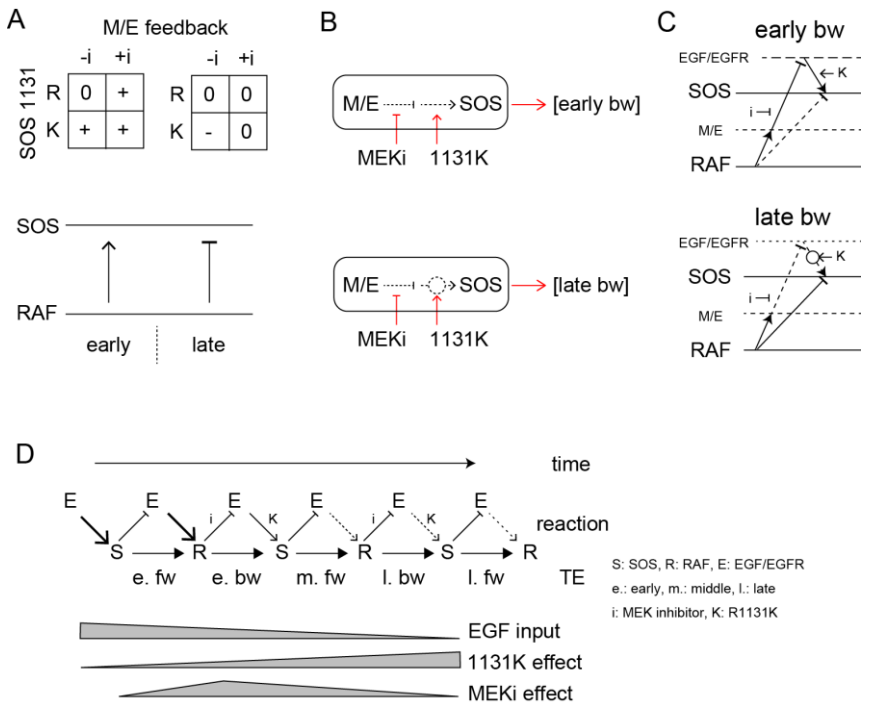


Figure 6. TE Regulation by the feedback loop and SOS function

A. 2D matrices showing the effects of the MEK inhibitor and R1131K mutation to the early (left) and late (right) bw-TE. The absence (-i) and presence (+i) of the MEK inhibitor, and the difference in the amino acid residue at SOS 1131 (R or K) caused an increase (+), maintenance (0), or decrease (-) in the TE amount compared to the control condition (-i, R). These 2D regulations increased or decreased bw-TE, respectively. **B.** Simple reaction networks to realize NAND and inversed ANDN gates. The black solid, black dotted, and red arrows indicate TE, reaction, and regulation, respectively. M/E: MEK and ERK. **C.** Possible reaction networks carrying the bw-TE in the early (upper) and the late (lower) peaks. EGF/EGFR and M/E are the hidden states. These networks are consistent with one of the networks carrying bw-TE suggested in Fig. 3E. See text for details. **D.** TE exchange between SOS and RAF involving the EGF/EGFR complex. The role of EGF/EGFR decreases with time after the cell stimulation. The action points of the MEK inhibitor and the R1131K SOS mutation were indicated. See text for details.