Figure 3
Fig. 4

(a) 

Ras-GTP
Ras-GTP-signaling nanocluster

EGF

(b) 

Ras nanoclusters (number)

EGF input

(c) 

Nanocluster Output (ppERK)

0.4 s

1
0.3
0.03
0

(d) 

System Output (ppERK)

increase clustered fraction

EGF input
Signalling ballet in four dimensions

Boris N. Kholodenko1,2*, John F. Hancock3, Walter Kolch1
1Systems Biology Ireland, University College Dublin, Belfield, Dublin 4, Ireland, 2Department of Pathology, Anatomy and Cell Biology, Thomas Jefferson University, Philadelphia, USA, 3Department of Integrative Biology and Pharmacology, University of Texas Medical School at Houston, Houston, Texas, USA

*Corresponding author

Preface

We are amassing extensive catalogues of signalling network components. However, our understanding of the spatio-temporal control of emergent network structures has lagged behind. Dynamic behaviour is starting to be explored genome-wide, but analysis of spatial behaviours is still confined to individual proteins. The grand challenge is to reveal how cells integrate the temporal and spatial information to determine specific biological functions. Key findings are the discovery of molecular signalling machines, such as Ras nanoclusters, spatial activity gradients and flexible network circuitries, involving transcriptional feedback. They reveal design principles of spatio-temporal organisation that are critical for network function and cell-fate decisions.

Introduction

Signal transduction was once viewed as a collection of linear information transporting pipelines that relate extracellular cues to specific genes. However, subsequent studies showed that different receptors often activate the same pathways and downstream effectors. Owing to pathway crosstalk, signals propagate through a tangled network of interconnecting proteins and cascades, rather than through independent linear routes. Furthermore, we find ourselves caught in a dilemma pushed into the limelight by the genome projects: the number of genes is far fewer than the number of biological processes. Hence, the concept that the specificity of biological processes is generated on the gene or even protein level erodes.

The idea of isolated pathways has given way to the concept of signalling networks which allow a limited number of components to generate an exponentially larger number of outcomes due to combinatorial interactions. Although we now can describe parts of these complex network topologies in great detail, we still do not understand how they operate to generate biological specificity. It is like trying to plan a journey with an incomplete railway network map lacking train time schedules. Even the simple transport of a signal requires a network map, a time schedule and a notion of connectivities, that is, spatio-temporal coordination. However, biological networks not only transport, but also process and integrate signals. Critical cellular decisions, including whether to undergo proliferation, apoptosis, and differentiation, are governed by the temporal dynamics and spatial distribution of key signalling effectors1-4. This realisation provided a strong impetus to explore the emergent properties of signalling networks that are encoded by spatial and temporal dynamics.

This review highlights current efforts towards understanding the spatial (three dimensional) and temporal (fourth dimension) dynamics of signalling networks and discusses the challenges ahead. We summarize how temporal signalling dynamics control cellular behaviour, and describe how the cellular localisation of protein interactions and spatially distributed signalling processes add to the regulation of phenotypic responses. We conclude with a discussion on how temporal and spatial controls cooperate to
choreograph the four-dimensional dynamics of signalling networks that specify complex biological processes. Due to space constraints we use selected examples to highlight the principles. This is an evolving field with sometimes competing hypotheses, which we discuss with a view of reconciling where possible and highlighting differences that may stimulate further research.

**Temporal network dynamics govern cell fate decisions**

The sharing of components between signalling pathways makes the function of these components context-dependent. How is context defined? Temporal specification is a versatile way to define context. In the railway metaphor, all trains use common components and modules, but shifting the time when they run can determine which connections are enabled or disabled.

**ERK signalling duration controls phenotypic responses**

First insights into how different cell-surface receptors employ shared pathways to generate specific cellular outcomes came from observations linking the duration of extracellular signal-regulated kinase (ERK) activation to cell fate decisions in rat pheochromocytoma PC-12 cells. In these experiments, transient ERK activation by epidermal growth factor (EGF) induced cell proliferation, whereas sustained ERK activation by nerve growth factor (NGF) induced differentiation. Subsequent work extended the connection between ERK activation kinetics and cellular outcome to other cell types and conditions. For instance, in human breast adenocarcinoma MCF-7 cells transient ERK activation by EGF induced proliferation, while sustained ERK activation by heregulin induced differentiation. In squamous cell carcinoma SCC-12F cells sustained ERK activation by EGF and scatter factor/hepatocyte growth factor stimulated migration, whereas transient ERK activation by keratinocyte growth factor and insulin-like growth factor 1 stimulated proliferation. These results highlight that temporal dynamics can specify distinct cell behaviours.

The ERK cascade comprises a three-tiered kinase module where the first kinase phosphorylates and activates the second kinase, which phosphorylates the third kinase in a two-step, non-processive reaction. This theme is implemented by nature in several variations commonly summarized as mitogen-activated protein kinase (MAPK) pathways, that currently include 15 known MAPKs, the most prominent being the ERK1/2, c-Jun N-terminal kinases (JNKs) and p38MAPK pathways. Computational modelling shows that distinct modes of ERK spatio-temporal dynamics can emerge from different feedback wiring. Depending on the feedback topologies and kinetic parameters, a MAPK cascade can display dramatically different temporal responses to an identical constant stimulus: a monotone, sustained response (panel A); a transient, adaptive response including near-perfect adaptation (B,C); damped oscillations (D); sustained oscillations (E); and a switch-like, bistable response where two stable steady states, “Off” and “On”, coexist (F). Different ERK temporal responses are also observed experimentally. Although some details remain unclear, plausible mechanisms for the dynamic control of MAPK signalling in mammalian cells have been proposed. They include ERK-induced feedback phosphorylation of upstream kinases, for instance Raf-1, whose ability to activate MEK is impeded by ERK-mediated phosphorylation, and their regulators, such as the Raf kinase inhibitor protein (RKIP) and the Ras activating guanine nucleotide exchange factor (GEF) SOS. Note that the core Ras family of small GTPases has three members K-Ras, H-Ras and N-Ras, and unless specified we use Ras to refer to the family. In combination, these MAPK feedbacks can produce complex temporal activity patterns. In fact, oscillations in MAPK cascades were first theoretically predicted to occur as a result of negative feedback ERK to SOS or Raf-1 and ultrasensitivity of the ERK responses to changes in the input and later were discovered experimentally. Oscillations of ERK activity induced by negative feedback are enhanced by switch-like ERK activation caused by positive conditions.
(double-negative) feedback arising from ERK-mediated inhibitory phosphorylation of RKIP\textsuperscript{16}. Interestingly, both bistable and oscillatory MAPK dynamics can also arise from double phosphorylation of ERK and mutual sequestration when a kinase at an upper level, such as MEK, forms a complex with a kinase (ERK) at the next cascade level\textsuperscript{18,19}.

Inferring connections within signalling networks that underlie the observed complex dynamics is an emerging challenge in cell biology. It is not obvious how to capture interactions between individual signalling nodes, since any activating or inhibitory stimulus applied to a particular node rapidly propagates through a network, causing widespread changes. An approach to untangle unknown network topologies, modular response analysis (MRA) groups multiple components into functional modules and infers their connections by measuring system-wide responses to systematic perturbations to all network modules\textsuperscript{20,21}. Exploiting MRA, Santos et al found that distinct temporal profiles of active ERK stimulated by EGF and NGF indeed emerge from differential feedback wiring of the ERK cascade, with EGF eliciting negative feedback and NGF inducing positive feedback\textsuperscript{22}. Thus, networks are not hardwired, but can respond to different inputs by reconfiguring themselves. It will be interesting to study whether this plasticity just involves feedback loops or whether connections between network core structures are also subjected to dynamic remodelling.

**ERK signalling induces transcriptional negative feedback**

Phosphorylation of MAPK cascade kinases is reversed by serine/threonine phosphatases, tyrosine phosphatases and dual specificity phosphatases (DUSPs). DUSPs are also known as MAPK phosphatases (MKPs), because they dephosphorylate ERK1/2, JNKs, and p38. Many DUSPs are immediate early genes (IEGs) induced by activated MAPKs\textsuperscript{23,24}. ERK-mediated DUSP induction tightly controls ERK activity, and such transcriptional negative feedback appears to be a common design principle of nearly all eukaryotic signalling pathways\textsuperscript{25,26}. Differential induction and localisation of DUSP isoforms in the cytoplasm and nucleus raises the intriguing possibility of different temporal dynamics for cytoplasmic and nuclear pools of ERK, and adds to the repertoire of signalling responses that determine cell-fate decisions.

Transcriptional negative feedback generated by ERK-mediated DUSP expression can bring about ERK and DUSP oscillations that develop on a longer time scale than oscillations arising from immediate negative feedback in the cytosol\textsuperscript{8,27}. In haploid *Saccharomyces cerevisiae* yeast cells, pheromones secreted by a cell activate MAPK signalling in cells of the opposite mating type, which results in the formation of mating-projections. Recent study demonstrated that in these cells, oscillations of Fus3 MAPK activity depend on transcriptional induction of the MAPK phosphatase Msg5 and the negative regulator of pheromone receptor signalling, Sst\textsuperscript{28}. The Fus3 oscillations are observed on the same time scale (two to three hours) as the periodic formation of additional mating projections\textsuperscript{28}. Thus, oscillations in MAPK signalling activity have a clear physiological role to control gene expression on long time scales.

**Discrete, digital outputs determine cell decisions**

Cells in an organism are immersed in an ocean of growth factors and hormones. In addition, protein concentrations vary between individual cells\textsuperscript{29} adding intrinsic stochastic noise. How do cells discriminate between signal and noise? One possibility is that graded, analogue signals from receptors are converted into discrete, digital outputs, such as the all-or-none responses of signalling and gene expression cascades. Indeed, theoretical studies have demonstrated that cell signalling circuits can act as analogue-to-digital converters, generating abrupt switches, multistable dynamics, excitable pulses, and oscillations, and these distinct outputs facilitate signal discrimination\textsuperscript{30-32} (Box 1). Recent experimental
work has shown that in different cell types and organisms, conversion of graded signals to digital outputs occurs at different levels. For instance in *Xenopus laevis* oocytes, graded progesterone and sorbitol stimuli are converted into switch-like, all-or-none responses of the Mos/MEK/p42 MAPK and JNK cascades in the cytosol, whereas in Swiss 3T3 fibroblast cells analogue ERK activation is converted into digital IEG responses in the nucleus.

Ultrasensitive and bistable responses (Fig.1F) can be elicited by long positive or short positive feedback loops. An example for the former is RKIP inactivation by ERK. The latter arises when RasGTP (produced when SOS catalyzes the exchange of GDP for GTP in the nucleotide binding pocket of Ras) binds to the SOS allosteric pocket causing a significant increase in the activity of this Ras-activating GEF, thereby stimulating further Ras activation. Signal digitalisation is also observed in the nucleus. In Swiss 3T3 cells EGF elicits transient ERK activity and negligible induction of the c-Fos IEG, whereas platelet-derived growth factor (PDGF) elicits sustained ERK activity and a significant c-Fos response. This is a result of cooperative expression and phosphorylation. EGF stimulates c-fos transcription, but the c-Fos protein is rapidly degraded. PDGF induces sustained activity of ERK and its substrate, p90 ribosomal S6 kinase 2 (RSK), which together stabilize nascent c-Fos by phosphorylation on multiple sites. Initial phosphorylation of c-Fos in the C-terminus exposes an ERK docking site, resulting in its further phosphorylation on Thr325 and Thr331. We hypothesize that multi-site phosphorylation of c-Fos, and nested feed-forward stabilization loops from ERK and RSK to c-Fos, result in the observed switch-like expression response. Thus, c-Fos serves as a binary cellular sensor of the duration and threshold intensity of ERK signalling.

**Spatio-temporal control of information transfer**

Besides the temporal kinetic specification of network function, the spatial control plays a major and complementary role. Our trains may leave at the same time but go to different destinations, or leave at different times heading for the same destination. While the spatial component is less studied than the temporal one, emerging evidence highlights its importance.

**Spatio-temporal organisation controls network function**

Recent systems biology work has elucidated how differences in the dynamic ERK interactome instigate cell fate decisions in PC12 cells. Quantitative proteomics showed that ~ 20% of ERK interaction partners bind to ERK in a growth factor regulated fashion, with ~30% of these interactions being differentially regulated by EGF and NGF, and impacting on differentiation. Interestingly, the majority of these interactions regulated spatial and temporal aspects of ERK signalling. Cytosolic interactions involved preformed complexes, from which stimulation caused release of ERK, whereas nuclear interactions were induced. NGF stimulated the nuclear translocation of ERK required for differentiation. In the nucleus ERK phosphorylated transcription factors and prompted export of transcriptional inhibitors. Interfering with these processes blocked differentiation. Thus, part of the NGF specific transcriptional programme is initiated by the physical removal of inhibitors from the nucleus. In the cytosol, the regulated ERK interactions included fine spatial compartmentalisation processes. For instance, NGF induced the sustained dissociation of the Ras-GAP NF1 from a Ras-ERK complex extending Ras and ERK activation. NGF also induced the sustained release of ERK from the anchor protein PEA-15, which sequesters ERK in the cytosol. Whenever tested, ERK interaction partners were also found to be ERK substrates, suggesting that ERK is anchored in the cytosol by its different
substrates. This mechanism ensures rapid phosphorylation by ERK, but also localises ERK signalling to different micro-compartments determined by the expression and distribution of substrates.

It will be interesting to explore whether upstream ERK activators are also part of these preformed complexes, and if they reach them by random diffusion or active recruitment. The existence of binary and ternary scaffolds that bind different components of the Ras-Raf-MEK-ERK cascade suggests that both partially and fully scaffolded complexes occur. For instance, MAPK organizer-1 (MORG1) interacts with Raf, MEK and ERK, but also with MP1\textsuperscript{38}, which scaffolds MEK and ERK\textsuperscript{39}. Thus, MORG1 can function as scaffold of a scaffold that may allow the modular assembly of a combinatorial variety of signalling complexes with distinct input – output functions. It is tempting to speculate that MORG1’s ability to regulate ERK activation only in response to selected stimuli is related to this supra-scaffolding function. The participation of individual proteins in different signalling complexes will generate competing partitioning between the complexes and may accentuate the specification of functional outputs. This may serve as mechanism to coordinate signalling specificity, although this issue is unexplored at present.

Another example of output regulation by spatio-temporal coordination is the NFκB pathway. The NFκB transcription factor is inactive in the cytosol when tethered to IκB inhibitor proteins. Phosphorylation-induced degradation of IκB releases NFκB to translocate into the nucleus and activate transcription, which can follow by NFκB transport back to the cytoplasm\textsuperscript{40}. The oscillation cycle of nucleo-cytoplasmic shuttling is finely tuned by several pathway inhibitors, including A20 and IκB isoforms, whose expression is stimulated by NFκB. IκB proteins can bind to NFκB in the nucleus and force export to the cytosol, where they retain NFκB. A20 blocks NFκB activation by inducing the degradation of receptor interacting protein (RIP), an adaptor protein that mediates IKK activation\textsuperscript{41}. Single cell analysis and modelling shows that different oscillation frequencies are determined by differentially-timed stimulation pulses and sequential induction of inhibitors\textsuperscript{42}. In this study fluorescent protein labelled fusion proteins were expressed at near physiological levels\textsuperscript{42}. There is a general caveat that exogenous expression may disturb the behaviour of signalling networks. However, recent results show that endogenous protein levels in individual cells vary extensively\textsuperscript{29} making artefacts resulting from mild overexpression unlikely. A triple feedback model considering stochastic transcription of the IκBα and A20 and delayed expression of IκBε genes predicts that IκBα and A20 cause oscillations, while IκBε increases response heterogeneity between cells\textsuperscript{42}. This is due to the delayed transcriptional induction of IκBε that adds stochastic noise\textsuperscript{42}. The biological benefit of adding noise is not well understood, but can be related to desynchronizing oscillations in different cells. The cellular heterogeneity generated by noise could provide an advantage where a cell population needs to react in a highly adaptive and selective way, such as in the immune response. Similar conclusions were derived from extensive stochastic simulations of the NFκB pathway\textsuperscript{43}. Importantly, the oscillation frequency modulates the specificity of gene expression\textsuperscript{42}. These studies demonstrate that the spatial arrangement of signalling proteins is subject to dynamic regulation, and vice versa that spatial organisation can specify kinetic activity profiles (Fig. 2).

**Scaffolds: managers of spatio-temporal organisation**

The marriage of spatial and temporal orchestration is embodied in scaffolding proteins. Scaffolds are hallmarked by their ability to simultaneously bind two or more signalling proteins that typically have an enzyme-substrate relationship. The physical co-localisation generates interesting properties, such as insulating signalling modules by physically tying them together; reducing reaction kinetics to zero order; enabling immediate feedback; and anchoring protein complexes to distinct subcellular sites. Importantly, scaffolds allow the re-use of enzymes for different functions in a highly context dependent manner,
providing a simple solution to the dilemma of possessing fewer genes than processes. These properties make scaffolds ideally suited to operate as organising principles in both synthetic and metabolic signalling networks. The first examples of scaffolds were the A-Kinase Anchoring Proteins (AKAPs), an impressive family of wide functional diversity. As AKAPs have been extensively reviewed we will simply point out that they often bind kinases, their phosphatases, upstream activators and downstream effectors. They assemble signalling platforms that orchestrate input-output relationships through physically coupled activation-deactivation cycles at highly-localised sites in cells. The function of scaffolds in other systems, such as the MAPK pathway, is increasingly being appreciated and is also covered by recent reviews. Here, we focus on open questions and try to delineate the design principles of how scaffolds contribute to spatio-temporal organisation of signalling networks. To illustrate this, we use two examples

The first is β-arrestin, a scaffold that coordinates the activation of multiple signalling pathways downstream of G-Protein Coupled Receptors (GPCRs). GPCRs activate ERK via two spatially and temporally separated pathways. Rapid ERK activation emanating from the plasma membrane is transient, β-arrestin independent, and permits ERK translocation to the nucleus. Sustained ERK activation is triggered from an endosomal Raf-MEK-ERK module that is assembled by β-arrestin scaffolding and restrains ERK signalling to the cytoplasm. Originally described as a protein that desensitizes signalling from GPCRs, β-arrestin has emerged as a multivalent scaffold protein that binds a bewildering array of signalling molecules. Some of these molecules are downstream effectors that propagate signals into the cell, such as ERK, JNK3, p38, AKT, whereas the others are enzymes that deactivate second messengers and inhibit GPCR-induced signalling, such as phosphodiesterases that degrade cyclic AMP and diacylglycerol kinases that degrade diacylglycerol. Typically, β-arrestin scaffolds whole modules, such as Raf-MEK-ERK or ASK1-MKK4-JNK3 (ASK1 and MKK4 are the upstream kinases of JNK3). This raises the question of how binding to the scaffold is regulated, and how the correct specific assemblies are generated. The latter may be explained in part by the interaction sites that exist in individual components of the specific kinase cascades and lead to preformed modules. A scaffold, in this case β-arrestin, would stabilize the preformed assemblies but also regulate the specificity, efficiency, and amplitude of signal propagation. However, β-arrestin still needs to pick up the appropriate module and this may be regulated by spatial distribution and dynamics. For instance, as described below, the activation of the Raf-MEK-ERK cascade occurs at the plasma membrane in distinct Ras nanoclusters using KSR, a scaffold for the Raf-MEK-ERK module. GPCR-mediated ERK activation can be dissected into two phases: an early phase that is β-arrestin independent and may correspond to nanocluster activation at the membrane, and a late phase that extends ERK activity and is β-arrestin dependent. Interestingly, the early phase permits ERK nuclear translocation, while the late phase is triggered by endocytosed GPCRs and confines ERK signalling to the cytosol. This mechanism diversifies ERK function in GPCR signalling using spatial and temporal separation of activation, conceptually resembling signal splitters in electronic circuits. Thus, scaffolding can orchestrate signalling by defining the sequence of events in time and space.

The second paradigm illustrates that the specificity of ERK substrate phosphorylation may be determined by the localisation of upstream signals. Again scaffolds may play prominent roles by directing ERK phosphorylation to distinct substrates in the cytosol or nucleus, depending on the subcellular structure on which Ras is activated. Ras activation is not confined to the plasma membrane but also can occur at endomembranes. To phosphorylate cytosolic phospholipase A2 (cPLA2), ERK activated at the plasma membrane employs the KSR1 scaffold, while ERK activated at the endoplasmatic reticulum employs Sef-1. For feedback phosphorylation of the EGFR, ERK uses the IQGAP1 scaffold. Thus, depending on the input, differential scaffolding and subcellular
targeting can allow a kinase module to signal to different effectors, and importantly to do this in parallel. We still know very little about the molecular mechanisms, but new optical approaches will bridge this gap (Box 3) and recent advances have started to clarify some of the membrane structures that organise signalling.

An interesting new aspect emerged with the discovery that scaffolds can function as allosteric regulators of their client kinases. KSR can activate its client B-Raf by side-to-side dimerisation presumably by an allosteric mechanism. Finer mechanistic details were elaborated in the yeast scaffold Ste5, which functions in the pheromone mating pathway. Ste5 has two docking sites for its client kinase Fus3. The strong docking site stimulates Fus3 to phosphorylate and downregulate mating signalling through the Ste5 pathway. The weak binding site assists the activation of Fus3 by allosterically improving its accessibility to phosphorylation by the upstream kinase Ste7 thereby enhancing pheromone signalling. In addition, the localisation of Ste5 is critical for the quality of the signal output. In its natural localisation at the cell membrane, Ste5 generates a graded output, while confining Ste5 to the cytosol enhances the inherent ultrasensitive activation of Fus3.

**Organelle apposition facilitates signal transfer.**

Similarly as scaffolds bring together signalling molecules to facilitate their interactions, the apposition of two cellular organelles can create spatial highways for exchanging signalling molecules. For instance, close location of mitochondria to the endoplasmic reticulum (ER) Ca^{2+}-releasing channels overcomes the low affinities of mitochondria Ca^{2+}-transporters, allowing for a rapid calcium accumulation in the mitochondrial matrix following the opening of ER calcium stores. Recently, the components of the macromolecular tether between the ER and mitochondria were identified as the integral ER membrane protein Mmm1 and the outer mitochondria membrane proteins Mdm10, Mdm12 and Mdm34. This large tether complex, termed ERMES (ER-mitochondria encounter structure), contains multiple copies of these proteins. In addition, ERMES was shown to facilitate direct phospholipid exchange between the ER and mitochondria.

Depletion of ER calcium stores triggers the opening of store-operated Ca^{2+}-channels in the plasma membrane, allowing extracellular Ca^{2+} to entry the cell. ER-plasma membrane junctions appear to be key regulators of Ca^{2+} influx, whereas the ER protein STIM1 serves as a Ca^{2+}-sensor. STIM1 that spans the ER membrane aggregates into oligomeric complexes upon store depletion and translocates to the plasma membrane to activate Ca^{2+}-release-activated Ca^{2+} (CRAC) channels in non-excitable cells. Like in ER-mitochondria contacts, an assembly of protein complexes that contain both ER and plasma membrane proteins is required for tunnelling signals at ER-plasma membrane junctions. The STIM oligomers directly bind the CRAC channel protein Orai1, which opens the CRAC channel. Subsequently, the smooth ER Ca^{2+}-ATPase pump (SERCA) replenishes the ER Ca^{2+} stores.

**Nano and microscale signalling domains**

The plasma membrane is a major platform for signal transduction. Plasma membrane architecture in turn can dictate network properties, by sequestering signalling proteins in space and time. The plasma membrane is an asymmetric lipid bilayer, comprising >7000 lipid species. It is compartmentalized into ultrafine compartments by the engagement of transmembrane proteins with the submembrane cortical actin mesh. Classical diffusion occurs within individual compartments, but long range diffusion across multiple compartments is impeded by compartment boundaries. Inhomogeneities in the lipid bilayer also exist as a result of incomplete lipid mixing. These lipid assemblies include complexes of glycosphingolipid and cholesterol, conventionally called lipid rafts, which transiently exist on ~20ms time scales and <20nm length scales. Rafts can be stabilized by engagement with lipid anchored
proteins or cross-linking\textsuperscript{66-68}. In consequence, the plasma membrane comprises a complex, non-random, dynamic array of lipids and proteo-lipid complexes on many different length and time scales. The extent to which specific proteo-lipid and lipid-lipid complexes are gainfully employed by cells has been debated\textsuperscript{64,67}. Yet, there is clear evidence for proteo-lipid and lipid based sorting platforms functioning in endocytic and exocytic trafficking, membrane curvature formation, cell migration and polarity\textsuperscript{68}.

**Assembling protein-lipid nanodomains**

Plasma membrane spatio-temporal dynamics may regulate signalling complexes that are permanently tethered, or transiently recruited to the plasma membrane. One example is the formation of transient, nanoscale protein clusters (nanoclusters) that operate as temporary signalling platforms or reaction chambers. These clusters may contain mixtures of kinases, phosphatases and other signalling proteins, anchored directly to the membrane or by lipid anchored regulatory proteins or subunits. Localisation of proteins to such nanoscale domains can accelerate reaction rates of signalling events that require specific protein–protein or protein–lipid interactions\textsuperscript{69}.

**Fine control of input-output by signalling nanocircuits**

Among the best-characterized membrane proteolipid complexes are those formed by Ras GTPases. The three Ras isoforms, H-, N- and K-Ras share a common G-domain (guanine nucleotide binding domain) but use different C-terminal anchors for membrane binding: H- and N-Ras are farnesylated and acylated, whereas K-Ras is only farnesylated and requires an adjacent polybasic domain for stable anchoring\textsuperscript{70}. Ras proteins are spatially arrayed on the plasma membrane as a combination of nanoclusters and freely diffusing monomers\textsuperscript{71}. A nanocluster comprises \(\sim 7\) Ras proteins, has a radius of \(\sim 9\text{nm}\) and an estimated lifetime of \(0.5-1\text{s}\)\textsuperscript{72,73} (Fig.4a). The formation of highly dynamic nanoclusters involves a complex interplay between the Ras lipid anchor, plasma membrane elements, amino acids in the C-terminal Ras hypervariable region and its G-domain, and ancillary scaffold proteins such as galectins. In consequence, N-, H- and K-Ras assemble into spatially distinct, non-mixed clusters, with further segregation between GTP-loaded and GDP-loaded proteins. For example, H-RasGDP clusters are cholesterol and actin dependent, H-RasGTP clusters are actin and cholesterol independent, K-RasGTP clusters are weakly actin dependent and cholesterol independent\textsuperscript{72}. K-RasGTP and H-RasGTP clusters require specific scaffold proteins to promote assembly: these include galectin-1 or galectin-3 that are selectively recruited from the cytosol by H-RasGTP or K-RasGTP, respectively\textsuperscript{74-76}. The scaffold protein stabilizes the G-domain of a RasGTP monomer in a signaling competent orientation with respect to the plasma membrane\textsuperscript{76,77}.

The subset of RasGTP proteins found in nanoclusters is termed the clustered fraction (~40% in fibroblasts). Importantly, currently available data suggest that RasGTP nanoclusters might be the sole sites of Raf recruitment and ERK activation on the plasma membrane; if Ras nanoclustering is abolished, ERK activation emerging from the membrane fails\textsuperscript{71,73,76,78}. Conversely, increasing the K-RasGTP clustered fraction enhances ERK signaling\textsuperscript{75,76}. RasGTP nanoclusters recruit Raf and KSR-MEK-ERK complexes from the cytosol for activation. Due to scaffolding, the ERK pathway module in a nanocluster generates the same activated ERK (ppERK) output for a wide range of Raf kinase inputs, operating as a low threshold switch\textsuperscript{78-80} (Fig.4c). The generation of ppERK is terminated by spontaneous disassembly of the nanocluster without requirement for biochemical deactivation of the kinase cascade. This is a novel use of spatio-temporal dynamics that within the digital nanodomain eliminates the potential problem of hysteresis arising from dual non-processive ERK phosphorylation\textsuperscript{81}.

Spatio-temporal dynamics are also responsible for a cellular ppERK response that is analogue with respect to EGF stimulation. As a result of non-equilibrium kinetics the RasGTP clustered fraction is
approximately constant over a wide range of RasGTP levels. Consequently there is a linear relationship between RasGTP levels and the number of nanoclusters generated on the plasma membrane (Fig. 4b). Thus, although each nanocluster delivers a brief, quantal ppERK output (Fig. 4c), the total output of ppERK from the plasma membrane is analogue (Fig. 4d). The Ras nanocluster circuitry therefore allows the plasma membrane to operate as analogue-digital-analogue (ADA) converter, digitizing the EGF analogue input signal for transmission across the plasma membrane by generating an appropriate number of Ras nanoclusters, and then regenerating the analogue EGF signal as a matched ppERK output into the cytosol. The system operates with high fidelity, as expected of a digital signalling system.

A similar ADA converter operates via the glycosylphosphatidylinositol-anchored protein (GPI-AP) CD59 nanocircuitry regulating immune cell activation. Signalling is triggered by the formation of CD59 clusters. The Src family kinase, Lyn and the heterotrimeric G-protein subunit, Gαi, are recruited to the clusters and activated, in turn activating PLCγ. The signal output from each cluster is a digital pulse of inositol-1,4,5-tris-phosphate (InsP3), produced by PLCγ-catalysed hydrolysis of phosphatidylinositol bis-phosphate, which is released into the cytosol. As with Ras nanoclusters, the duration of the InsP3 pulse is time-limited and on the same time-scale as the output of ppERK from a Ras nanocluster. The overall system response is analogue, because the InsP3 outputs from the individual clusters are summed in the cytosol and converted into an analogue Ca2+ signal by activation of InsP3 receptors on intracellular calcium stores. The short lifetime of the Ras and CD59 clusters is critical for high fidelity signal transmission because it allows for a high sampling rate of the analogue input signal. Indeed, computation shows that as the cluster lifetime increases, fidelity is lost and the system response becomes progressively digital, being determined by the biochemical kinetics and not plasma membrane spatio-temporal dynamics.

Given the similarity between the Ras and GPI-anchored nanocluster systems, it is tempting to speculate that this type of ADA circuitry may represent a general mechanism for high fidelity signal transmission by lipid-anchored signalling proteins. In common with Ras proteins, GPI-APs also exhibit a fixed monomer/cluster distribution that violates simple mass action kinetics. Recent work has shown that this distribution is actively maintained and critically dependent on cortical actin dynamics. More broadly the role of the unique architecture of the plasma membrane in supporting assembly of ADA converters brings into focus membrane spatio-temporal dynamics as a novel regulator of signal transmission. Signal response is determined by the prevailing nanoclustered fractions of key signalling molecules. Since the clustered fractions are dependent, inter alia, on the state of the cortical actin cytoskeleton and lipid content of the plasma membrane, these and other inputs, such as cellular growth and metabolic state, migratory information and cell contact data, can be integrated to set the gain control of the ADA converter (Fig. 4d).

Activation of the ERK cascade on the Golgi by H- or N-Ras, or in the cytosol, is dependent on Raf kinase input. Therefore, the ppERK output from Ras nanoclusters if they operate on Golgi membranes is analogue not digital. Thus, subcellular localisation critically determines how the ERK module output is wired. These different system outputs from different compartments are biologically relevant. In the developing mouse immune system, antigens that drive high strength activation of the Ras/ERK pathway from the plasma membrane lead to clonal deletion of T-cells, whereas antigens that drive low strength activation from the Golgi result in clonal expansion. Thus, spatial organisation can directly determine the quality and biological effect of signalling output.

Chemical reactions and diffusion form spatial domains
Not long ago a cell was considered a bag of enzymes, where biochemical reactions proceeded within the well-stirred, spatially uniform milieu of an enzymologist’s test tube. Yet, Alan Turing’s theoretical work showed that biochemical reactions can impose spatial order and break the symmetry of an initially homogeneous medium. The fundamental novelty of Turing’s idea is that diffusion, which intuitively seems a process eliminating spatial heterogeneity, can generate periodic spatial patterns in the initially uniform environment, if the diffusion coefficients of interacting species are different. This laid the foundation of the physico-chemical theory of morphogenesis. If two morphogens, typically an activator and inhibitor, have different diffusivities, and the activator autocatalytically reproduces itself and stimulates its inhibitor, then the spatially uniform distribution can become unstable, and this instability drives the formation of heterogeneous spatial patterns. In order to prevent autocatalytic explosion, the inhibition process should be faster and its diffusion coefficient should be larger than that of the activator. Spatial ranges of activation and inhibition are determined by the diffusion coefficient and the half-life of the species. Short-range, local activation and long-range inhibition govern periodic increases in the activator concentration and the spacing between repeating peaks. Similar spatial patterns occur when long-range inhibition is substituted by depletion of a substrate that is required to produce an activator and is consumed by activation.

A different mechanism to generate positional information and spatial patterns exploits pre-existing heterogeneity in a cell. Signal transduction proceeds through cycles of reversible covalent modification of target proteins, catalyzed by an activator and deactivator, such as a kinase and phosphatase for a phospho-protein, or a GEF and GAP for a small GTPase. Owing to the presence of cellular structures, such as membranes, cytoplasm, organelles and chromosomes, opposing activator and deactivator enzymes are often spatially segregated. In other words, the intracellular environment for reactions and diffusion is initially inhomogeneous and does not resemble the uniform media traditionally considered for Turing-type models. For a protein phosphorylated by a membrane-bound kinase and dephosphorylated by a cytosolic phosphatase, a precipitous phosphorylation gradient with high phosphoprotein concentration close to the membrane and low concentration in the cell interior was predicted. Provided that the phosphatase is far from saturation, the stationary phosphorylation profile decays almost exponentially with the distance from the membrane. The characteristic decay length \( L_{\text{grad}} \) is determined by the diffusion coefficient \( D \) and the apparent first-order rate constant \( k \) of the phosphatase (deactivator) and does not depend on the kinase (activator) kinetics, \( L_{\text{grad}} = \sqrt{D/k} \). This simple spatial pattern is stable and occurs because of the pre-existing separation of opposing enzymes. Activity gradients of this type have been discovered experimentally for the small GTPase Ran, yeast MAPK Fus3, phosphatase PTP1B, Aurora B kinase, and yeast protein kinase Pom1.

If an active protein form associates with other proteins, generating multi-protein complexes, or rapidly, reversibly binds to cytoskeleton elements, the apparent diffusion coefficient \( D^* \) of this form becomes smaller than the diffusion coefficient \( D \) of a free inactive form. Then, stable intracellular gradients of the total protein abundance arise from the spatial separation of activator and deactivator enzymes. These total protein gradients \( \text{Grad}_{\text{total}} \) are less precipitous than gradients \( \text{Grad}_a \) of the active form, \( \text{Grad}_{\text{total}} / \text{Grad}_a = (1 - D^*/D) < 1 \). Thus, stable stationary patterns of different signalling activities and protein abundances in diverse subcellular domains can arise from the spatial separation of opposing enzymes and diffusion.

**Intricate concentration landscapes in single cells**

Intricate landscapes of steady-state protein activities arise from different spatial localisation of kinases/phosphatases and GEFs/GAPs on cell membranes, chromatin structures or in the cytoplasm. For
many kinase cascades, such as MAPK cascades, the first level kinase is activated on the plasma membrane in response to external stimulation. When this occurs in a fully scaffolded environment, such as nanoclusters, the whole activation process is kinetically confined by spatial colocalisation. However, when the cascade components are allowed to diffuse, either after the rapid disassembly of nanoclusters or alternative modes of activation in different compartments, the phosphorylation level and activity of the first kinase sharply decrease during a diffusion journey in the cytoplasm. Since only the phosphorylated kinase fraction can stimulate a downstream kinase, a progressive reduction of the stimulation efficacy down the cascade occurs. This raises the question as to under what conditions signals that emanate from the membrane robustly propagate into the cell interior. The assumption that both kinases and phosphatases are far from saturation, and that the apparent first-order rate constants do not change from layer to layer, allows us to formulate a simple criterion of signal propagation. If the ratio of the phosphatase (deactivator) and kinase (activator) rate constants is much smaller than one, activation signals readily spread from the plasma membrane into the cell interior. The stationary activation profiles of successive kinases down the cascade display long, flat plateaus, which abruptly decay at the spatial locations following each other at almost constant space intervals. These intervals can be much larger than the characteristic decay length of the spatial activation profile for the initial kinase, spreading phosphorylation signals deep into a cell. If the ratio of the deactivator and activator rate constants is greater than one, the signal propagation stalls as the spatial profiles of successive activated proteins rapidly decay, confining the signal to the membrane. More complex stable spatial patterns can arise from multisite phosphorylation of kinases in signalling cascades. For instance, if a kinase at each level has two phosphorylation sites and the condition considered above for signal propagation is fulfilled, the activation profiles of dually phosphorylated forms at successive cascade levels display similar long, flat plateaus, whereas mono-phosphorylated kinases exhibit non-monotonous, transient concentration profiles. Their peaks are localized close to the places where the stationary fronts of dually phosphorylated kinases rapidly decay.

In GTPase cascades, an active GTPase can positively or negatively control GEFs or GAPs at multiple levels. For instance, a ballet of small GTPases controls cytoskeleton dynamics, whereby Cdc42 activates Rac and possibly Rho, whereas Rac and Rho inhibit each other. Theoretical considerations showed that the spatial separation of GEFs and GAPs can lead to complex patterns of GTPase activities in a cell, where activities can decrease, increase or exhibit peaks with an increase in the distance from the cellular structure (such as a membrane or chromosome) where the initial GEF is localised. Such complex, non-monotonic, stable concentration profiles were recently reported for a chromosome-dependent RanGTPase-importin cascade, coupled with a secondary phosphorylation network. All these patterns are brought about by the spatial separation of opposing enzymes in activation-deactivation cycles of protein modification, rather than by the spatial symmetry breaking and instability of the spatially uniform distribution that occurs in the Turing mechanism.

Positive and negative feedback loops in protein cascades can bring dynamic instabilities in time and space. For instance, traveling waves can occur in bistable MAPK cascades where diffusion coefficients of components are assumed to be the same (Box 2). Incorporation of the existing spatial heterogeneity brought about by cellular structures into Turing-type models appears to be promising to account for many intricate dynamic processes within single cells. For instance, it recently was shown that nonlinear interactions between prototypic activator and inhibitor on the plasma membrane can lead to the emergence of Turing’s spatial patterns even for equal diffusion constants, provided that the exchange rates between the membrane and cytoplasm or decay rates in the cytoplasm are different for the activator and inhibitor. Likewise, a Turing-type model was shown to account for the spontaneous initiation of cell polarization by the small GTPase Cdc42, its GEF Cdc24 and the effector protein Bem1, which
shuttle between the membrane and cytoplasm\textsuperscript{102}. Alternative models of cell polarization based on bistable kinetics of the Cdc42/Rac/Rho network have also been proposed\textsuperscript{103}. These discrepancies may reflect the infancy of our understanding or simply the fact that Nature has evolved more than one solution.

**Conclusions**

We have changed our perception of signalling pathways from linear pipelines to networks. We also have begun to rationalise how these network structures can determine the kinetics of distinct biochemical processes with high fidelity to translate them into specific biological responses. Along this way we have realised that specificity is generated by combinatorial assemblies and spatio-temporal dynamics rather than by a large number of genes with specific functions. We now face the challenge to explain why evolution chose combinatorial assemblies over single pathway deterministic solutions. An obvious advantage of the former is that successful designs can be recycled and adapted for new purposes. Spatial and temporal separation may be convenient means to specify signalling functions. This suggests that cell shape plays an important role as it defines the spatial coordinates. Intriguing first glimpses were provided by work showing that cell shape controls the dynamics of localised biochemical activities\textsuperscript{104, 105}. This also highlights the need for new approaches, both conceptually and technologically, to move hand in hand for developing the insights and tools that allow us to survey the complex landscape of cell signalling.

**Acknowledgement.**

We thank Mikhail Tsyganov, Javier Muñoz García, Anatoly Kiyatkin and Nikolai Kaimachnikov for discussions. This work was supported by Science Foundation Ireland under Grant No. 06/CE/B1129 and NIH grants GM059570, GM066717. We apologize that we could not cite many pertinent contributions to the field because of space limitations.
Box 1. Temporal signalling dynamics.

Understanding the temporal dynamics of signalling networks is facilitated by using kinetic schemes and ordinary differential equations (ODEs). Each signalling species is produced and consumed in particular reactions. The ODE left-hand side is the time derivative of a species concentration, and the ODE right-hand is the algebraic sum of the reaction rates, which produce and consume that species. Since ODEs do not consider spatial dimension, this kinetic description implies a well-mixed, homogeneous reaction medium. This simplification facilitates the analysis of the effects of multiple inputs, feedback loops and pathway crosstalk on the dynamics of complex signalling networks. Remarkably, already simple signalling motifs display intricate temporal dynamics. For instance, it was recently shown that the basic activation/deactivation cycle of the Src tyrosine kinase can exhibit complex signalling dynamics that includes oscillations, toggle switches and excitable behaviour. Src kinase can exist in four different states (see the figure, part a). In the basal autoinhibited conformation (S₀), Src is phosphorylated on the inhibitory Tyr in the C-terminus and unphosphorylated on the activatory Tyr in the catalytic domain. Both of these Tyr residues are unphosphorylated in the partially active form (S). In the first fully active conformation (S₁) the inhibitory Tyr is dephosphorylated and activatory Tyr is phosphorylated, and in the second fully active form (S₂) both Tyr residues are phosphorylated. A critical nonlinearity is brought about by intermolecular autophosphorylation of the activatory Tyr (shown by green lines). Importantly, the complex Src dynamics do not require imposed external feedback loops and can occur at constant activities of Src inhibitors (such as C-terminal Src kinase (Csk)) and Src activators (such as protein tyrosine phosphatase 1B (PTP1B) and receptor-type protein tyrosine phosphatases (RPTP)). In different ranges of activities of these Src regulators, Src kinase activity can exhibit hysteresis (bistability) (see figure part b), oscillations (see figure part c), and excitable responses of active Src kinase fractions, s₁ and s₂ (see figure part d). Modified/Reproduced with permission from Ref.
Box 2. Spatially distributed signalling.

Ordinary differential equation-based models can account for spatial dynamics only by using coarse-grained compartmentalisation, for instance, by considering the cytoplasm and nucleus as two different, well-mixed compartments. More accurately, spatially-distributed signalling processes are modelled using partial differential equations (PDE), which contain derivatives with respect to both time and space and are referred to as "reaction-diffusion" equations. Reaction-diffusion PDEs describe the concentration changes due to both the production and consumption of a species in chemical reactions and its redistribution in space caused by diffusion and transport (see, e.g.,\(^{87,100}\)). Solutions to reaction-diffusion equations are generally obtained by numerical integration, although for linear or saturated kinetics, stationary concentration profiles were obtained analytically (see, e.g.\(^{88,108}\)). In spatially distributed signalling networks, such as the β-adrenergic receptor pathway\(^{105}\) or MAPK cascades\(^{100}\), the initiating signals, cyclic AMP or Raf kinase activity, respectively, are generated at the plasma membrane and subsequently spread into the cell through the sequential activation of downstream proteins (see the figure, part a, which illustrates activation of successive kinases in the Raf/MEK/ERK cascade).

Importantly, the transmission of spatial information is controlled by feedback and feedforward network motifs and the cell shape\(^{104,109,110}\). For instance, bistability in a two-site ERK (de)phosphorylation cycle\(^{18}\) was shown to generate a phosphoprotein wave that propagates from the surface into the cell interior\(^{100}\). Figure, part b shows the simulated time-course of the propagation of MEK (dashed red) and ERK (blue) activation into the cytoplasm of a large cell, such as *Xenopus laevis* egg (cell radius is 50μm, nuclear radius is 20μm, modified from Ref.\(^{100}\)). This wave propagation is facilitated by feedback inhibition of phosphatases due to generation of reactive oxygen species (for detail see Ref.\(^{100}\)). Positive feedback from ppERK to cytoplasmic MEK further enhances the propagation span of the wave, making it possible to convey phosphorylation signals over exceedingly long distances. Such waves of protein phosphorylation/modification that travel with constant amplitude and velocity may transmit survival signalling along axons in developing neurons\(^{111,112}\).
Box 3

**Microscopic technologies to analyze spatiotemporal organization of signalling networks**

Assembling signaling networks in time and space requires cellular imaging. Different methods provide high-resolution spatial and temporal information by spatially mapping molecules with respect to a defined cellular structure or to another molecule. Each method has limitations and complete visualization of a network requires multiple approaches.

2 Dimensional imaging: although wide field fluorescence imaging cannot provide 3D resolution, this problem is in part mitigated by TIRF (Total internal reflectance microscopy) that limits imaging to a thin (~200nm) slice of the cell immediately adherent to the coverslip. Events occurring on, or in the vicinity of the basal plasma membrane can be exclusively observed.

3 Dimensional imaging: is achieved by confocal and two-photon microscopy, the latter can access deep into samples or tissues. Both methods are diffraction limited. Higher resolution imaging is feasible with 4-Pi and STED (STimulated Emission Depletion) microscopy that use non-linear de-excitation of fluorophores to bypass the resolution limit of diffraction. STED microscopy can visualize lipid rafts in live cells.

**FRET methods:** detect proximity of molecules on length scales of 1-10nm. Live cell FRET imaging can quantify and localize specific molecular interactions to cellular structures. Fluorescence lifetime imaging (FLIM) is predominantly used as a robust method to measure FRET. FRET between donor and acceptor fluorophores reduces the fluorescence lifetime of the donor, which can be used to generate a FRET image (see figure). Fluorescence anisotropy microscopy is used to identify homo-FRET between identical fluorophores, it has been used recently to map spatiotemporal dynamics of GPI-AR nanoclustering.

**Single molecule and spectroscopic techniques:** SPT (single particle tracking), SFVT (single fluorophore video tracking), track the diffusion of single molecules with high spatial and temporal resolution. SPT does not rely on fluorescent excitation and emission and yields higher temporal resolution. Fluorescence correlation spectroscopy (FCS) and fluorescence cross-correlation spectroscopy (FCCS) measure fluorescent fluctuations in small confocal volumes. The data can be used to derive diffusion constants and in the case of FCCS infer and quantify molecular interactions without relying on FRET.

**Example of a FLIM-FRET analysis**

BHK cells expressing mGFP-KrasG12V alone or in combination with mRFP-Raf1 were imaged in a wide field FLIM microscope. The lifetime of mGFP-KrasG12V is significantly decreased when co-expressed with mRFP-Raf1 due to FRET between mGFP and mRFP after recruitment of Raf-1 to Ras nanoclusters. Pixel-by-pixel fitting of two lifetimes and calibration using a mGFP-mRFP fusion protein allows the fraction of mGFP molecules undergoing FRET to be calculated (~30%)
**Figure Legends**

**Fig. 1. Versatile MAPK dynamics.** Each panel schematically displays the Raf-MEK-ERK mitogen-activated protein kinase (MAPK) cascade with no feedback (panel A), negative feedback (shown red, panels B-E) or positive feedback (green, panel F) from extracellular signal-regulated kinase (ERK) to Raf, which is the initial MAPK cascade kinase activated by RasGTP. The different temporal responses of active, dually-phosphorylated ERK (ppERK) to a constant RasGTP stimulus are obtained by changing parameters of ERK-mediated feedback. A kinetic description of MAPK cascade reactions (rate equations) is given in Supplementary Table 1, where parameter values correspond to panel A. Parameters $F$ and $K_f$ describe the feedback regulation ($F = 1$, if there is no feedback, $F < 1$ for negative feedback, $F > 1$ for positive feedback, see Supplementary Table 1). 

<table>
<thead>
<tr>
<th>Panel</th>
<th>$F$</th>
<th>$K_f$</th>
<th>$V_{\text{max}}^\gamma$</th>
<th>$k_{\text{cat}}^\gamma$</th>
<th>$V_{\text{max}}^{\text{II}}$</th>
<th>$k_{\text{cat}}^{\text{II}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>25nM</td>
<td>0.175 nM$^{-1}$s$^{-1}$</td>
<td>0.025 s$^{-1}$</td>
<td>0.025 nM$^{-1}$s$^{-1}$</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.34</td>
<td>25nM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.01</td>
<td>1nM, 0.175 nM$^{-1}$s$^{-1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.27</td>
<td>25nM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.01</td>
<td>25 nM, 100 nM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>5</td>
<td>100 nM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Depending on the initial conditions (pre-existing activity level of the cascade), ppERK either descends to the low activity state (blue curves) or approach the high-affinity state (red curves); dashed line indicates a threshold.

**Fig. 2. Scaffolds and spatial organisation.** (a) G-protein coupled receptors (GPCRs) activate ERK through two spatially and temporally separated pathways. Mechanistic details are omitted for the sake of clarity and were reviewed previously. Rapid ERK activation emanating from the plasma membrane through protein kinase C (PKC), Src, and receptor tyrosine kinase (RTK) stimulation is transient and β-arrestin independent. It permits ERK translocation to the nucleus. Sustained ERK activation is triggered from an endosomal β-arrestin dependent Raf-MEK-ERK module and restrains ERK signalling to the cytosol. The blue lines represent the integrated ppERK that results from combining the nuclear (red) and cytosolic, β-arrestin scaffolded (green) ppERK levels. (b) Ras activated at different subcellular compartments uses different scaffolding proteins to target ERK substrates. IQGAP1 (on the cytoskeleton) mediates negative ERK feedback phosphorylation of the epidermal growth factor (EGF) receptor, while KSR (at the plasma membrane) and Sef1 (in the Golgi) facilitate phosphorylation of cytoplasmic phospholipase A2 (cPLA2) by ERK activated at the plasma membrane or endomembranes, respectively. It is suggested that Ras activated at the ER stimulates Sef1-bound ERK at the Golgi. However, as Ras also can be activated at the Golgi directly, we have included this possibility. In either case ERK phosphorylation activates cPLA2 to generate arachidonic acid which is a precursor to signalling molecules including leukotrienes and prostaglandins.

**Fig. 3. Intrinsic transcriptional feedback inhibition of NFκB.** In resting cells the NFκB transcription factor is rendered inactive because IκBs retain it in the cytosol. Upon activation, e.g by TNFα, receptor protein complexes are assembled that initiate dual signalling pathways. One promotes apoptosis by stimulating caspase 8 activation. Another counteracts apoptosis by activating NFκB. This activation is enabled by the stimulus induced degradation of IκBs, which releases NFκB from its cytosolic anchor proteins to translocate to the nucleus. However, nuclear NFκB also induces the transcription of its own inhibitors. IκBs can bind to nuclear NFκB and export it back to the cytosol, while A20 can interrupt receptor mediated NFκB activation by inducing the degradation of RIP, an adaptor protein that couples the TNFα receptor to the NFκB pathway.

**Fig 4. Ras nanoclusters digitize transmembrane signal transmission**
a. Activation of EGF receptors generates K-RasGTP on the plasma membrane. A fixed proportion of these K-RasGTP molecules (= the clustered fraction) assemble into signalling nanoclusters. Each cluster has a radius of ~9nm and contains ~7 K-RasGTP molecules. At higher EGF concentrations more nanoclusters form. b. Since K-RasGTP levels are directly proportional to non-saturating EGF doses, and the K-RasGTP clustered fraction remains constant as K-RasGTP levels increase, the number of K-RasGTP nanoclusters depends linearly on stimulating EGF concentration. c. After the recruitment of Raf and KSR/MEK/ERK complexes from the cytosol each nanocluster outputs a digital pulse of ppERK. The ppERK output is insensitive to Raf kinase input and is limited by disassembly of the nanocluster; a 2-log range of relative Raf kinase inputs is shown. d. In consequence of b and c, the total system response to EGF, which is the aggregated digital outputs from all of the transiently active nanoclusters, is analog (blue line). The gain of the response is increased if the Ras clustered fraction increases from 40% to higher values (orange and purple lines).
References


A conceptual break-through summarising multiple experimental observations that different durations of ERK activity may result in different phenotypic responses.


Provides insight into a full set of protein-protein interactions involving ERK and demonstrates how ERK partners control the ERK spatio-temporal dynamics and cell decisions.


Predicted sustained MAPK oscillations later discovered experimentally (see Refs. 12 and 14).


A working model of combinatorially complex interactions of multi-domain proteins that control PI3-K and ERK pathway crosstalk.


**Direct experimental determination of context-dependent and time-varying topology of dynamic connections between MAPK cascade components.**

**Direct experimental evidence of the Ras-SOS positive feedback**

**Demonstrated how a short and prolonged duration of ERK signalling can be sensed at the level of immediate early genes**

**Provides a combined experimental and mathematical analysis of the nucleo-cytoplasmic shuttling cycles of NF-kappaB and how they relate to specifying gene expression.**

**A kinetic model, which demonstrated that scaffold organisation of a kinase cascade dramatically changes the input-output relationships.**


**Provides insight in how Ras signalling from different membrane compartments utilises different scaffold proteins for the ERK pathway to selectively target downstream ERK substrates.**
Shows that Ras signalling can emanate from different subcellular membrane compartments and activate different downstream pathways.


**Demonstrated that in mammalian cells the MAPK cascade can operate as a switch with different sensitivity to the input signals from the plasma membrane and cytoplasm.**


**showed that diffusion can destabilize spatially uniform steady-state distribution resulting in heterogeneous spatial concentration patterns.**


**Demonstrated that the spatial separation of opposing enzymes in a protein-modification cycle brings about protein activity gradients and non-uniform spatial profiles.**


**Demonstrated that the propagation of phosphorylation signals solely by diffusion can be terminated by cytoplasmic phosphatases. In large cells, motor-driven trafficking of endosomes or scaffolds carrying phosphorylated kinases or assembled signalling complexes is required for signal transduction.**


**Showed the possibility of waves of protein phosphorylation travelling through the cytoplasm or long neuron axons.**


Signalling ballet in four dimensions

Supplementary Information
Boris N. Kholodenko\textsuperscript{1,2*}, John F. Hancock\textsuperscript{3}, Walter Kolch\textsuperscript{1}

\textsuperscript{1}Systems Biology Ireland, University College Dublin, Belfield, Dublin 4, Ireland, \textsuperscript{2}Department of Pathology, Anatomy and Cell Biology, Thomas Jefferson University, Philadelphia, USA, \textsuperscript{3}Department of Integrative Biology and Pharmacology, University of Texas Medical School at Houston, Houston, Texas, USA
Table I. Kinetic description of MAPK cascade reactions

The maximal rates ($V_i^{\text{max}}$), Michaelis ($K_m$) and catalytic ($k_i^{\text{cat}}$) constants are expressed in nM·s$^{-1}$, nM and s$^{-1}$, respectively. The total protein concentrations (the sums of differently phosphorylated forms) are assumed to be constant on the time scale considered, $Raf_{\text{total}} = 300$ nM; $MEK_{\text{total}} = 300$ nM; $ERK_{\text{total}} = 300$ nM. Input signal starts at $t = 0$ when $[\text{RasGTP}]$ changes from 0 to 10 nM, and resulting responses of ppERK are shown in Fig. 1 of the main text. The parameters $F$ and $K_f$ describe feedback regulation by ppERK; $F < 1$ for negative feedback, $F > 1$ for positive feedback and $F = 1$, if this feedback is absent.

<table>
<thead>
<tr>
<th>N</th>
<th>Reaction</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\text{Raf} \rightarrow \text{pRaf}$</td>
<td>$V_1 = \frac{k_1^{\text{cat}} [\text{RasGTP}][\text{Raf}]}{1 + [\text{Raf}]/K_{m_1} + [\text{pRaf}]/K_{m_2}} \cdot \frac{1 + F[\text{ppERK}]/K_f}{1 + [\text{ppERK}]/K_f}$ \hspace{1cm} $k_1^{\text{cat}} = 1$ \hspace{1cm} $K_{m_1} = 100, K_{m_2} = 200$</td>
</tr>
<tr>
<td>2</td>
<td>$\text{pRaf} \rightarrow \text{ppRaf}$</td>
<td>$V_2 = \frac{k_2^{\text{cat}} [\text{RasGTP}][\text{pRaf}]/K_{m_2}}{1 + [\text{pRaf}]/K_{m_1} + [\text{pRaf}]/K_{m_2}} \cdot \frac{1 + F[\text{ppERK}]/K_f}{1 + [\text{ppERK}]/K_f}$ \hspace{1cm} $k_2^{\text{cat}} = 0.25$</td>
</tr>
<tr>
<td>3</td>
<td>$\text{ppRaf} \rightarrow \text{pRaf}$</td>
<td>$V_3 = \frac{V_3^{\text{max}} [\text{ppRaf}]/K_{m_3}}{1 + [\text{ppRaf}]/K_{m_3} + [\text{pRaf}]/K_{m_4}}$ \hspace{1cm} $V_3^{\text{max}} = 2.5$ \hspace{1cm} $K_{m_3} = 50, K_{m_4} = 100$</td>
</tr>
<tr>
<td>4</td>
<td>$\text{pRaf} \rightarrow \text{Raf}$</td>
<td>$V_4 = \frac{V_4^{\text{max}} [\text{pRaf}]/K_{m_4}}{1 + [\text{pRaf}]/K_{m_3} + [\text{pRaf}]/K_{m_4}}$ \hspace{1cm} $V_4^{\text{max}} = 3.75$</td>
</tr>
<tr>
<td>5</td>
<td>$\text{MEK} \rightarrow \text{pMEK}$</td>
<td>$V_5 = \frac{k_5^{\text{cat}} [\text{ppRaf}][\text{MEK}]/K_{m_5}}{1 + [\text{MEK}]/K_{m_5} + [\text{pMEK}]/K_{m_6}}$ \hspace{1cm} $k_5^{\text{cat}} = 2.5, K_{m_5} = 250, K_{m_6} = 250$</td>
</tr>
<tr>
<td>6</td>
<td>$\text{pMEK} \rightarrow \text{ppMEK}$</td>
<td>$V_6 = \frac{k_6^{\text{cat}} [\text{ppRaf}][\text{pMEK}]/K_{m_6}}{1 + [\text{MEK}]/K_{m_5} + [\text{pMEK}]/K_{m_6}}$ \hspace{1cm} $k_6^{\text{cat}} = 0.5$</td>
</tr>
<tr>
<td>7</td>
<td>$\text{ppMEK} \rightarrow \text{pMEK}$</td>
<td>$V_7 = \frac{V_7^{\text{max}} [\text{ppMEK}]/K_{m_7}}{1 + [\text{ppMEK}]/K_{m_7} + [\text{pMEK}]/K_{m_8}}$ \hspace{1cm} $V_7^{\text{max}} = 3$ \hspace{1cm} $K_{m_7} = 250, K_{m_8} = 80$</td>
</tr>
<tr>
<td>8</td>
<td>$\text{pMEK} \rightarrow \text{MEK}$</td>
<td>$V_8 = \frac{V_8^{\text{max}} [\text{pMEK}]/K_{m_8}}{1 + [\text{pMEK}]/K_{m_7} + [\text{pMEK}]/K_{m_8}}$ \hspace{1cm} $V_8^{\text{max}} = 3.75$</td>
</tr>
<tr>
<td>9</td>
<td>$\text{ERK} \rightarrow \text{pERK}$</td>
<td>$V_9 = \frac{k_9^{\text{cat}} [\text{ppMEK}][\text{ERK}]/K_{m_9}}{1 + [\text{ERK}]/K_{m_9} + [\text{pERK}]/K_{m_{10}}}$ \hspace{1cm} $k_9^{\text{cat}} = 0.125$ \hspace{1cm} $K_{m_9} = 250, K_{m_{10}} = 250$</td>
</tr>
<tr>
<td>10</td>
<td>$\text{pERK} \rightarrow \text{ppERK}$</td>
<td>$V_{10} = \frac{k_{10}^{\text{cat}} [\text{ppMEK}][\text{pERK}]/K_{m_{10}}}{1 + [\text{ERK}]/K_{m_9} + [\text{pERK}]/K_{m_{10}}}$ \hspace{1cm} $k_{10}^{\text{cat}} = 0.125$</td>
</tr>
<tr>
<td>11</td>
<td>$\text{ppERK} \rightarrow \text{pERK}$</td>
<td>$V_{11} = \frac{V_{11}^{\text{max}} [\text{ppERK}]/K_{m_{11}}}{1 + [\text{ppERK}]/K_{m_{11}} + [\text{pERK}]/K_{m_{12}} + [\text{ERK}]/K_{m_{13}}}$ \hspace{1cm} $V_{11}^{\text{max}} = 3.75$ \hspace{1cm} $K_{m_{11}} = 120, K_{m_{12}} = 20$</td>
</tr>
<tr>
<td>( v_{12} )</td>
<td>( p\text{ERK} \rightarrow \text{ERK} )</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>( v_{12} = \frac{V_{12}^{\text{max}} [p\text{ERK}] / K_{m12}}{1 + [p p\text{ERK}] / K_{m11} + [p\text{ERK}] / K_{m12} + [\text{ERK}] / K_{m13}} )</td>
<td>( V_{12}^{\text{max}} = 5, \ K_{m13} = 300 )</td>
<td></td>
</tr>
</tbody>
</table>
Differential equation system that describes the MAPK dynamics

\[
\begin{align*}
\frac{d[Raf]}{dt} &= v_4 - v_1 \\
\frac{d[pRaf]}{dt} &= v_1 - v_2 + v_3 - v_4 \\
\frac{d[ppRaf]}{dt} &= v_2 - v_3 \\
\frac{d[MEK]}{dt} &= v_8 - v_5 \\
\frac{d[pMEK]}{dt} &= v_5 - v_6 + v_7 - v_8 \\
\frac{d[ppMEK]}{dt} &= v_6 - v_7 \\
\frac{d[ERK]}{dt} &= v_{12} - v_9 \\
\frac{d[pERK]}{dt} &= v_9 - v_{10} + v_{11} - v_{12} \\
\frac{d[ppERK]}{dt} &= v_{10} - v_{11}
\end{align*}
\]

The initial conditions at \( t=0 \) are the following. Figs. 1A-1E: \([Raf] = Raf_{total} = 300 \text{ nM} \); \([MEK] = MEK_{total} = 300 \text{ nM} \); \([ERK] = ERK_{total} = 300 \text{ nM} \); all phosphorylated forms are assumed to equal zero. Fig.1F. Blue curves (below threshold) \([ppERK](t=0) = 50, 100, 105 \text{ nM} \); Red curves (above threshold) \([ppERK](t=0) = 111, 120, 150 \text{ nM} \). The threshold value is about 108 nM. The values of \([ERK](t=0) \) are calculated as \([ERK] = 300 - [ppERK](t=0) \text{ (nM)} \); \([pERK](t=0) = 0 \).