

Diverse roles of guanine nucleotide exchange factors in regulating collective cell migration

Assaf Zaritsky,^{1,2*} Yun-Yu Tseng,^{3*} M. Angeles Rabadán,³ Shefali Krishna,³ Michael Overholtzer,³ Gaudenz Danuser,^{1,2} and Alan Hall³

¹Department of Cell Biology and ²Department of Bioinformatics, University of Texas Southwestern Medical Center, Dallas, TX 75390

³Cell Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065

Efficient collective migration depends on a balance between contractility and cytoskeletal rearrangements, adhesion, and mechanical cell-cell communication, all controlled by GTPases of the RHO family. By comprehensive screening of guanine nucleotide exchange factors (GEFs) in human bronchial epithelial cell monolayers, we identified GEFs that are required for collective migration at large, such as SOS1 and β -PIX, and RHOA GEFs that are implicated in intercellular communication. Down-regulation of the latter GEFs differentially enhanced front-to-back propagation of guidance cues through the monolayer and was mirrored by down-regulation of RHOA expression and myosin II activity. Phenotype-based clustering of knockdown behaviors identified RHOA-ARHGEF18 and ARHGEF3-ARHGEF28-ARHGEF11 clusters, indicating that the latter may signal through other RHO-family GTPases. Indeed, knockdown of RHOC produced an intermediate between the two phenotypes. We conclude that for effective collective migration, the RHOA-GEFs \rightarrow RHOA/C \rightarrow actomyosin pathways must be optimally tuned to compromise between generation of motility forces and restriction of intercellular communication.

Introduction

Collective cell migration involves intercellular mechanical communication through adhesive contacts (Tambe et al., 2011; Weber et al., 2012; Zaritsky et al., 2015). In migrating monolayers, such communication is initiated by cells at the monolayer boundary and gradually transmitted to cells at the back of the group (Ng et al., 2012; Serra-Picamal et al., 2012; Zaritsky et al., 2014, 2015; Ladoux et al., 2016; Mayor and Etienne-Manneville, 2016). Effective cell-cell communication requires balanced control of contractility and cell-cell and cell-matrix adhesions (Hidalgo-Carcedo et al., 2011; Weber et al., 2012; Cai et al., 2014; Bazellières et al., 2015; Das et al., 2015; Hayer et al., 2016; Notbohm et al., 2016; Plutoni et al., 2016). Coordination between these processes is regulated, among several pathways, by signaling activities of the Rho-family GTPases (Wang et al., 2010; Hidalgo-Carcedo et al., 2011; Timpson et al., 2011; Omelchenko and Hall, 2012; Cai et al., 2014; Omelchenko et al., 2014; Reffay et al., 2014; Plutoni et al., 2016). Rho-family GTPases are spatially and temporally modulated by complex networks of upstream regulators, including 81 activating guanine nucleotide exchange factors (GEFs), 67 deactivating GTPase-activating proteins, and 3 guanine dissociation inhibi-

tors (Jaffe and Hall, 2005; Omelchenko and Hall, 2012). The networks are composed of many-to-one and one-to-many interaction motifs; that is, individual GTPases are regulated by multiple GEFs, and one GEF often acts upon multiple GTPases. Moreover, some GEFs are effectors of GTPases, leading to nested feedback and feedforward interactions (Schmidt and Hall, 2002; Jaffe and Hall, 2005; Cherfils and Zeghouf, 2013; Hodge and Ridley, 2016). Such pathway “design” permits an enormous functional specialization of transient signaling events, at specific subcellular locations and with precise kinetics.

Our long-term goal is to disentangle these signaling cascades in the context of collective cell migration. Although the roles of GEFs and their interactions with Rho GTPases are widely studied for single-cell migration (Goicoechea et al., 2014; Pascual-Vargas et al., 2017), less is known about how they regulate collective migration (Hidalgo-Carcedo et al., 2011; Omelchenko et al., 2014; Plutoni et al., 2016). Here, we report a comprehensive and validated, image-based GEF screen that identified differential roles of GEFs. By design of quantitative measures that encode the collective dynamics in space and time, we were able to identify a surprising role of RHOA, RHOC, and a group of four up-

Dr. Hall died on May 3, 2015.

*A. Zaritsky and Y.-Y. Tseng contributed equally to this paper.

Correspondence to Gaudenz Danuser: gaudenz.danuser@utsouthwestern.edu; or Michael Overholtzer: overhom1@mskcc.org

Abbreviations used: 16HBE, human bronchial epithelial cells from the 16HBE14o line; GEF, guanine nucleotide exchange factor; PC, principal component; PCA, PC analysis; qRT-PCR, quantitative RT-PCR.



stream GEFs in modulating collective migration via efficient long-range communication.

Results and discussion

Quantification of monolayer cell migration in space and time

Collective cell migration emerges from the individual motility of cells in an interacting group: an action of one cell affects its neighbor and can propagate over time to eventually coordinate distant cells (Zaritsky et al., 2015). To identify molecules implicated in this mechanism, we performed live-cell imaging of the wound-healing response of human bronchial epithelial cells from the 16HBE14o (16HBE) line (Fig. 1 A and Video 1). Cells formed apical junctions and maintained epithelial markers and group cohesiveness before scratching the monolayer, as assessed by the localization of E-cadherin and the tight-junction protein ZO1 at the lateral cell–cell contact areas (Fig. 1 B). Upon scratching, the monolayer transitioned over ~2 h from a nonmotile phase to an acceleration phase to steady-state wound closure (Fig. 1 C). The acceleration phase was associated with a gradual transition of cells from unorganized local movements to a faster and more organized motility. Cells at the wound edge underwent this transition first, followed by a wave of coordinated motility propagating away from the wound edge (Fig. 1 A, insets). The propagation is thought to be driven by mechanical cell–cell communication (Matsubayashi et al., 2011; Ng et al., 2012; Serra-Picamal et al., 2012; Zaritsky et al., 2014, 2015; Notbohm et al., 2016).

To screen for GEFs implicated in the migration-initiating process, we devised an automated and robust pipeline quantifying the spatiotemporal dynamics of motility activation in the monolayer. The core of this pipeline relies on robust detection of the wound edge via a segmentation algorithm that considers both image texture and segmentation consistency over time (Materials and methods). We then applied particle image velocimetry at subcellular resolution to measure the speed of cells in the monolayer as a function of their distance to the edge over time. Spatiotemporal dynamics were quantified and visualized in kymographs, in which every bin records the mean speed in a band of constant distance from the wound edge at a particular time point (Fig. 1 D). As expected, the speed kymograph showed a backward-propagating wave (Fig. 1 E, top left).

Using this assay, we first measured the effects of depletion of the canonical Rho GTPases Cdc42, Rac1, and RhoA. shRNAs were designed to selectively knock down each GTPase (Fig. 1 F). As expected (Simpson et al., 2008; Vitorino and Meyer, 2008; Omelchenko and Hall, 2012; Plutoni et al., 2016), knockdown of Cdc42 and Rac1 reduced cell speeds throughout space and time (Fig. 1, E and H, wound-healing rate; confirmed by a second distinct hairpin, Fig. 1 F and Fig. S1 A). We also confirmed inhibition of cell speed and monolayer migration under shRNA-mediated knockdown of the Rac1- and Cdc42-GEF β -PIX (Omelchenko et al., 2014; Fig. S1, B–D).

Knockdown of the Rho-GTPase RhoA induced near concurrent transition of cells at the front and back of the monolayer from a nonmotile to a motile state (Fig. 1, E and G). Importantly, RhoA-depleted cell monolayers reached the same wound-healing rates as unperturbed monolayers (Fig. 1 H). These data suggest that RhoA is critically implicated in the communication chain required to synchronize cell behavior throughout the monolayer after wound infliction.

Information encoded in a wound-healing experiment

Given the differential effects observed for the knockdown of Rho GTPases, we expected that down-regulating GEFs might also differentially alter the dynamics of motility activation. To capture the shifts in these behaviors, we established a concise representation of a live wound-healing experiment between wound infliction and steady-state wound closure, which could be compared across many experiments. We defined a 12-dimensional feature vector to encode the information contained in a speed kymograph over the first 180 μ m of the monolayer and 200 min after scratching by averaging the speed values in bins of 60 μ m and 50 min, respectively (Fig. 2 A; Materials and methods). Accordingly, features 1–4 encode the acceleration of cells at the monolayer front, features 5–8 the acceleration of cells 60–120 μ m behind the monolayer edge, and so on. Features 1, 5, and 9 encode spatial variations in migration shortly after wounding. In the example of Fig. 2 A, cells at the front have already developed motility (feature 1), but cells further back in the monolayer (feature 9) have not. Features 2, 6, and 10; features 3, 7, and 11; and features 4, 8, and 12 encode the same spatial variations at later times. The pattern of 12-dimensional feature vectors was replicated for 402 control experiments without shRNA treatment (Fig. 2 B, top). The broad variation in collective migration behavior, even for constant experimental conditions, was visualized by normalizing each feature across the population of all 402 experiments (Fig. 2 B, bottom; see Materials and methods).

To identify core features that would capture the relevant variation of these data over noise, we applied principal component analysis (PCA; Jolliffe, 2002) to orthogonally transform the set of possibly correlated 12-dimensional features to a smaller set of linearly uncorrelated features denoted as the principal components (PCs). Indeed, when using PCs, the first three components captured 92% of the variability in the population of control experiments, allowing a fourfold dimensionality reduction of the feature space. The variation of PC1 between experiments was highly correlated with the variation in wound-healing rate (Fig. 2 C). The same applied to a lesser extent to PC3, whereas PC2 was uncorrelated. We further analyzed the coefficients that map the original 12 features into the scalar value of the respective PC. For PC1, all 12 coefficients have a similar value. Hence the mapping corresponds to an averaging of the speed values across the kymograph (Fig. 2 D). For PC2, the coefficients are positive for features 1, 5, and 9 (i.e., the speed values in the first time window) and negative for features 4, 8, and 12 (i.e., the speed values in the last time window). Thus, PC2 captures the reverse temporal gradient of the speed values. For PC3, the coefficients are positive for features 1–4 (i.e., the speed values at the wound edge) and negative for features 9–12 (i.e., the speed values at the back of the monolayer). Thus, PC3 captures the spatial gradient of the speed values in the direction of wound closure.

Knockdown of Cdc42 and Rac1 decreased PC1, reflecting a significant overall reduction in speed (Fig. 2 E). PC2 was somewhat increased for shCdc42 but unaltered for shRac1, whereas PC3 strongly decreased for both. This shows that the global reduction in speed under these conditions is accompanied by flatter spatiotemporal gradients (Fig. 2 E; see Materials and methods). Knockdown of RhoA most prominently altered PC3, with no effect on PC1 (Fig. 2 E), reflecting an alteration in the spatial gradient, in accordance with Fig. 1 E.

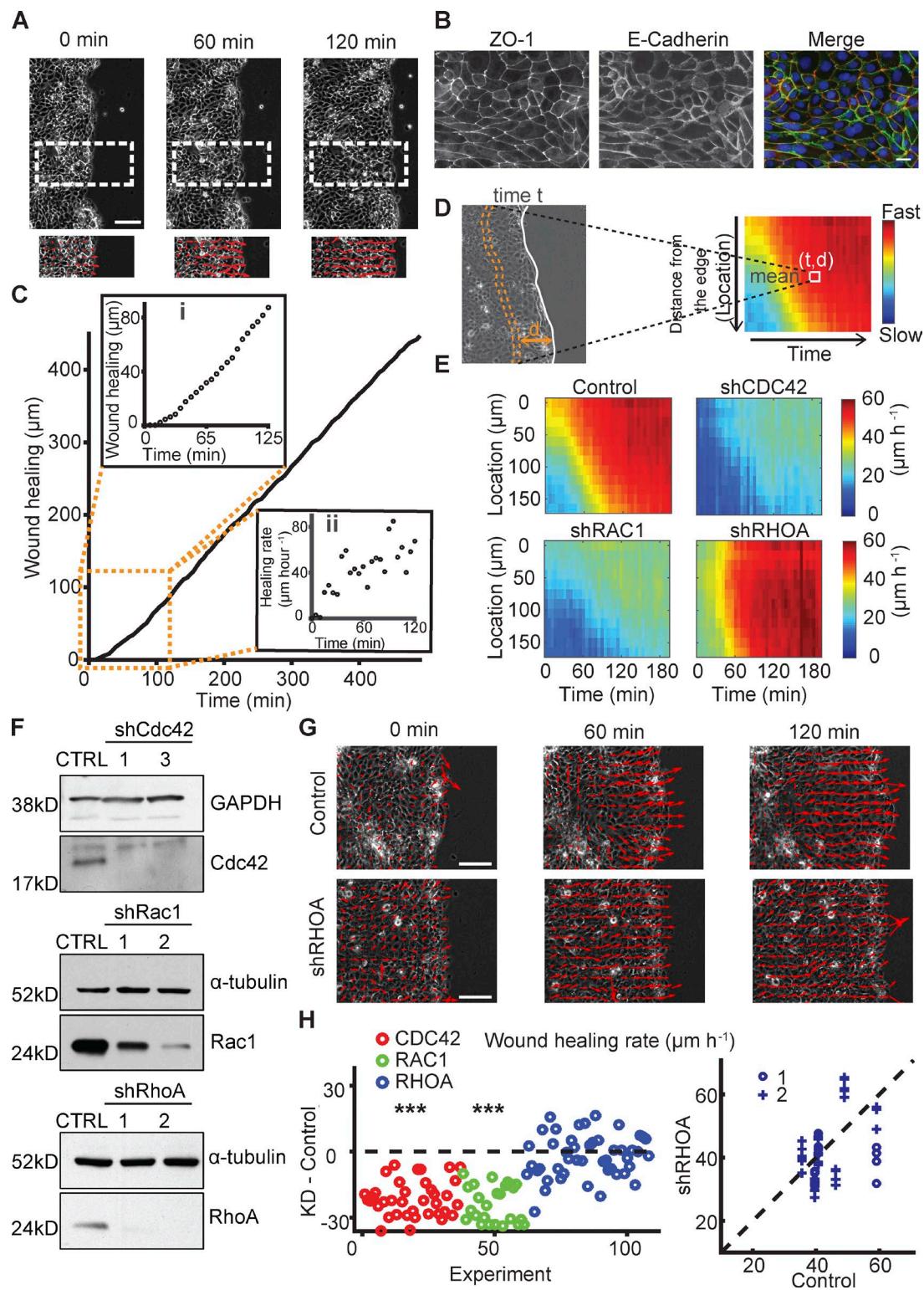


Figure 1. Effects of GTPase knockdown on collective cell migration. (A) Example of a wound-healing experiment. (insets) Velocity vectors showing that front cells begin to migrate before deeper cells. Bar, 100 μm . (B) Immunofluorescent staining of E-cadherin and ZO1 before scratching the monolayer. Bar, 20 μm . (C) Monolayer edge evolution over 500 min. (inset top) Edge evolution during the first 125 min. (inset bottom) Increase of wound-healing rate during the first 125 min. (D) Construction of a kymograph of a wound-healing experiment: mean speed of cells at different distances from the monolayer edge over time. (E) Speed kymographs for control cells and under depletion of Cdc42, Rac1, and RhoA. All kymographs are averages of four locations in a well (similar to Kim et al., 2013). (F) Western blots of control (CTRL) and RhoA-, Rac1-, and Cdc42-depleted cells. (G) Monolayer velocity over time (left to right) for control (top) and RhoA-depleted cells (bottom). Bars, 100 μm . (H) Wound-healing rate. Each point was calculated as the difference between the migration rate in one location and the mean of the same day's control experiment. $N = 6–9$ d, $n = 24–47$ locations (CDC42: $N = 9$, $n = 37$; RAC1: $N = 6$, $n = 24$; RHOA: $N = 9$, $n = 47$). Statistics via Wilcoxon signed rank test: ***, $P \leq 0.0001$. The right panel shows the migration rate in one location as a function of the mean daily control for the two RhoA hairpins.

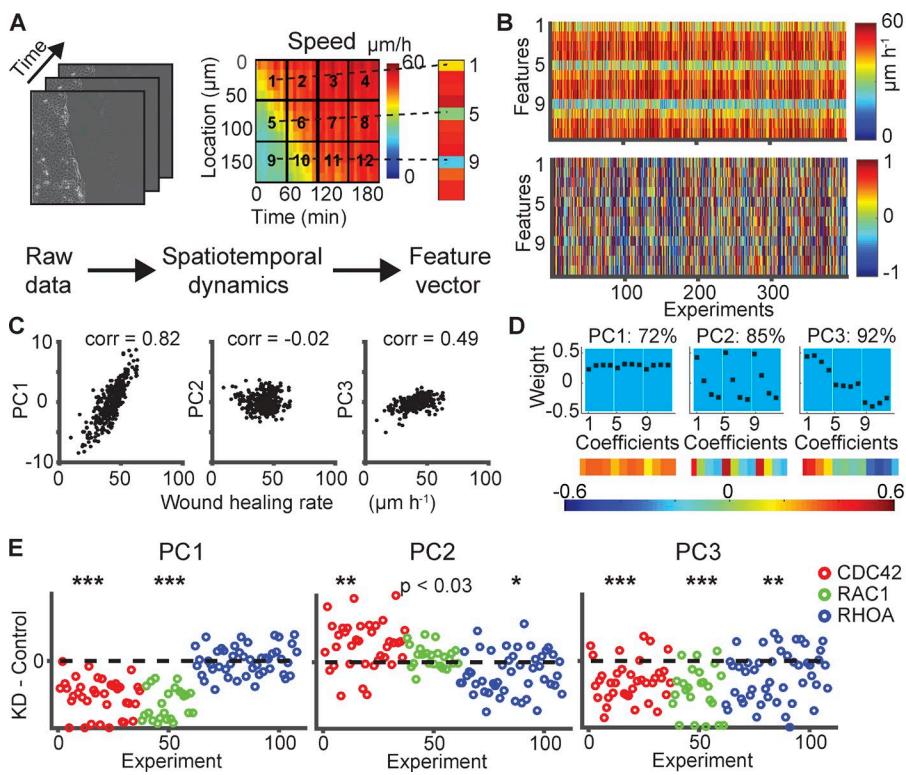


Figure 2. Information encoded in a wound-healing experiment. (A) Reducing a wound-healing experiment to a 12-dimensional feature vector encoding spatiotemporal dynamics. (B) Speed feature vectors for 402 control experiments. The bottom panel displays the normalized features. (C) Association of principal components (PCs) 1–3 with wound-healing rate. corr, correlation. (D) Coefficients of PCs 1–3, together encoding 92% of the speed variance in control experiments. (E) Effect on PCs by depletion of Cdc42, Rac1, and RhoA. Statistics via Wilcoxon signed rank test: *, $P \leq 0.01$; **, $P \leq 0.001$; ***, $P \leq 0.0001$.

When a monolayer migrates collectively, cells within the monolayer eventually have to migrate toward the monolayer edge. This directional cue is transmitted via backward propagation of a strain gradient (Zaritsky et al., 2014). Thus, the spatial propagation of migration directionality is a measure related to intercellular communication. To capture this aspect, we defined for every kymograph bin directionality as the absolute ratio between the mean velocity component perpendicular to the monolayer edge and the velocity component parallel to the monolayer edge. The directionality kymograph displayed a backward-propagating wave pattern, as previously observed for other cell systems (Ng et al., 2012; Zaritsky et al., 2014; Fig. S1 E). Analogous to the analysis of the speed kymographs, we binned directionality kymographs into a feature vector of 12 values (Materials and methods; Fig. 2 A) and applied PCA. The first three PCs resembled the PCs for speed (Fig. S1 F), suggesting a functional coupling between speed and directionality establishment in the process of activating collective migration. Similar to the speed analysis, PC1 of the directionality correlated strongly with the wound-healing rate, but PC2 and PC3 were uncorrelated (Fig. S1 G). Depletion of the GTPases Cdc42 and Rac1 primarily affected the overall magnitude of directionality, whereas RhoA depletion reduced the spatial gradient, indicative of rapid long-range communication between the monolayer front and back after wound infliction (Fig. S1 H). Altogether, our data establish the first three PCs of a feature space derived from kymographs as measures for the detection of alterations in magnitude and temporal or spatial gradients in speed and directionality and underline the roles RhoA plays in long-range communication across the monolayer.

Comprehensive GEF screen

Equipped with our image-based assay for collective migration, we targeted the 81 Rho-family GEFs known in the human genome by shRNA (Fig. 3 A). A custom library of pSUPER

shRNA retroviruses (three hairpins per GEF) was used for the screening, targeting 80 of 81 GEFs. 75 of the remaining 80 GEFs were confirmed to be expressed in the 16HBE cell line by Western blotting or quantitative RT-PCR (qRT-PCR) in cases in which no reliable antibody was available (Materials and methods). Western blots and qRT-PCR were used to evaluate the knockdown efficiency of every hairpin, and a cutoff of 50% depletion was selected for analysis of wound-healing experiments (Fig. 3 B). 16 GEFs had knockdown efficiency of less than 50% for all three hairpins and thus were eliminated from the screen. Within the remaining group of 59 GEFs, the knockdown of 11 GEFs could not be validated by qRT-PCR, because of failures in the production of efficient primers. Nonetheless, we included those genes in our imaging-based screen and labeled them “unknown knockdown efficiency.” A total of more than 2,600 time-lapse movies from the screen and the follow-up experiments were analyzed (Materials and methods). Control videos showed broad day-to-day variability in wound-healing rates (Fig. 3 C).

To score the alteration of a knockdown well, we quantified in multiple image locations the differences from the control well on the same plate. For every location, we extracted PCs 1–3 for speed and directionality. For each of the six PCs, we quantified the separation of target well and control well using three different distance metrics that balance intra- and inter-experiment variability (Materials and methods). For each combination of PC and distance metric, we used 24 experiments with no measurable knockdown (KD efficiency = 0%) as a baseline to calculate z -scores (Materials and methods).

To identify high-confidence hits, we estimated an objective z -score threshold by balancing between false-positive and false-negative rates. As positive controls we grouped together known motility reducers in Cdc42, Rac1, and β -PIX. Experiments with no measurable knockdown of the target were used as off-target controls. Their mean and standard deviation were

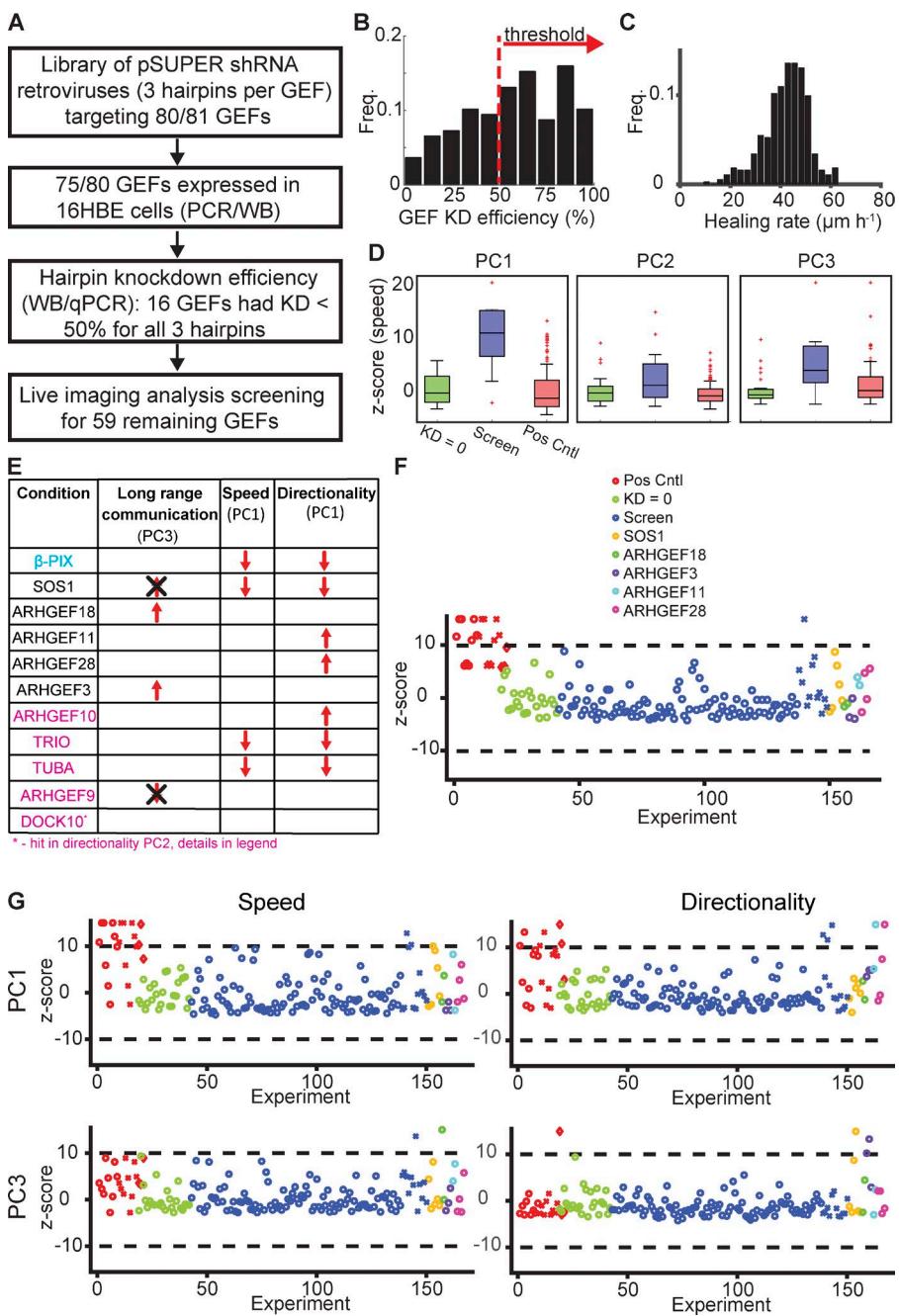


Figure 3. Comprehensive GEF screen. (A) Flowchart of screen. WB, Western blot. (B) Histogram of knockdown efficiency of all hairpins used in the screen. Freq., frequency. Only hairpins with depletion greater than 50% were considered for the screen (red). N (total number of hairpins) = 125. (C) Histogram of wound-healing rate. Mean \pm SD = $42.2 \pm 8.8 \mu\text{m h}^{-1}$. (D) Boxplots of z-scores for off-target controls (green, knockdown efficiency = 0%, $n = 24$), positive controls (Pos Cntl; red, CDC42, RAC1, and β -PIX, $n = 21$), and the screened GEFs (blue). Box represents 25%–75%, whiskers 5%–95% of the data, assuming a normal distribution. Z-scores were calculated relative to the off-target control. (E) Summary of GEFs identified as hits in the screen and their phenotypes. PC3 in speed or directionality signifies long-range communication. β -PIX was a GEF with known effects on collective migration and used as a positive control (cyan). Magenta labels hits that were not followed up. Black cross indicates that these phenotypes were excluded (see Fig. S2 for details). (F and G) Visualization of measures that were altered by the hits identified in the screen and that were validated. Shown are z-scores for each well. Positive controls (red): CDC42 (circle), RAC1 (cross), and β -PIX (diamond), off-target controls (knockdown = 0, green), GEFs that were screened (blue) but not identified (circle) or validated (cross) as hits, and validated hits. (F) Wound-healing rate. (G) PC1 (top) and PC3 (bottom).

used to calculate z-scores for all experiments. Distributions of z-scores in the speed PC measures are shown in Fig. 3 D. As expected, the positive controls were most discriminative for PC1. Thus, we exploited the known motility reduction of the positive controls together with the corresponding distribution of off-target controls to estimate a threshold appropriate to identify hits (Materials and methods). The threshold was estimated as 9.8. Because the threshold was standardized relative to the off-target controls, it was transferrable to the other PC-measures.

Screen summary

In addition to β -PIX, which has previously been reported as required for collective cell migration (Omelchenko et al., 2014), 10 new GEFs were identified with significant knockdown effects. Among those, we focused on SOS1, ARHGEF18, ARHGEF28,

ARHGEF11, and ARHGEF3 for validation (Fig. 3 E, black). The other five (Fig. 3 E, magenta; and Fig. 3, F and G, blue crosses) were not followed up for reasons discussed in Fig. S2 (A–E).

SOS1-RAS pathway is required for monolayer migration

We first investigated SOS1, a dual GEF for RAC1 and RAS (Nimnuan et al., 1998) that was found to regulate epithelial tight-junction formation in 16HBE cells through the MEK/ERK pathway (Durgan et al., 2015). Knockdown of Sos1 by two different hairpins induced a reduction in wound-healing rate, cell speed, and directionality (Fig. 3, F and G; and Fig. S2 F), consistent with its role as a Rac1 activator. Following the experiments in Durgan et al. (2015), we then blocked MEK and ERK activity directly by small-molecule inhibitors and concluded that the SOS1-RAS pathway (Fig. S2 G; Materials and

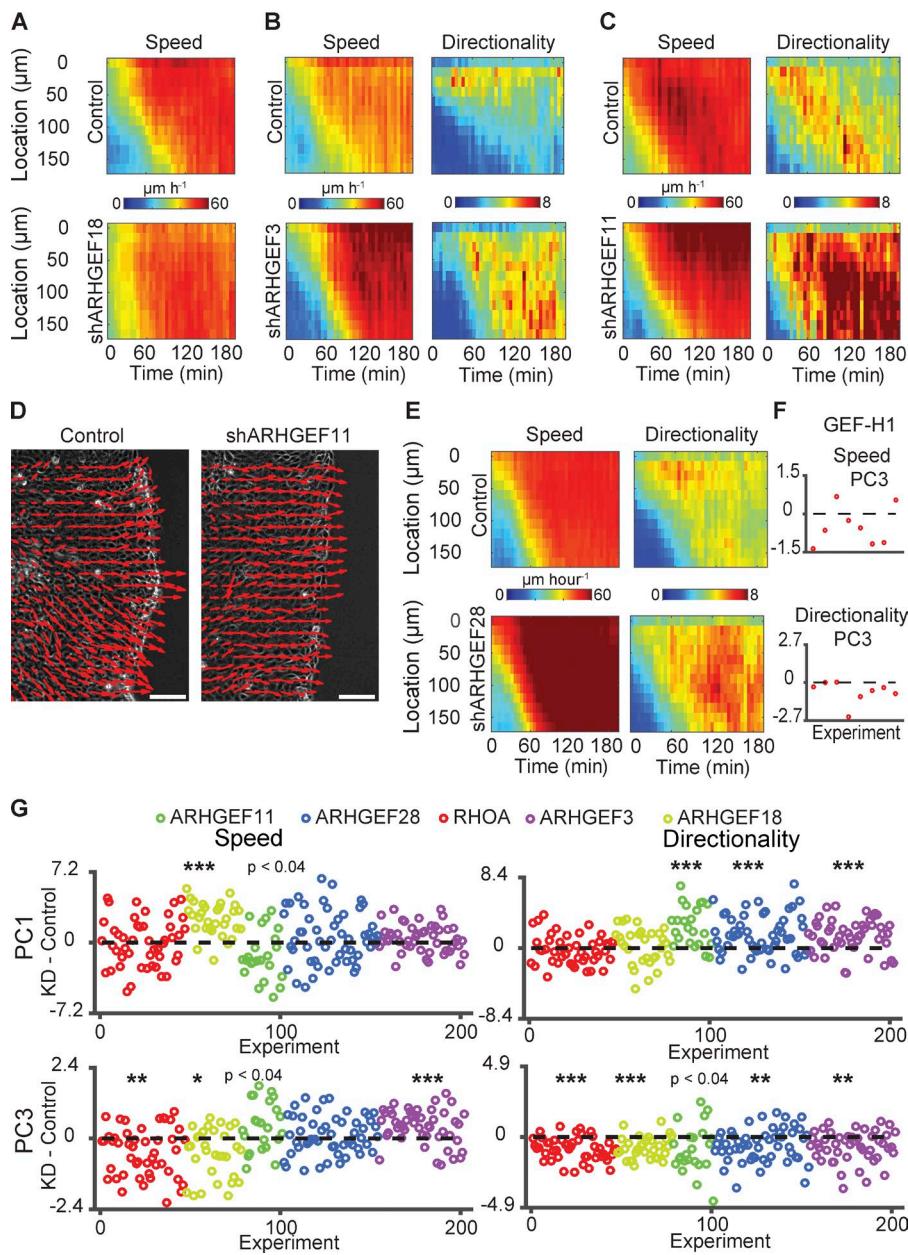


Figure 4. Effects of RhoA GEFs depletion on intercellular communication. (A–C) Kymographs of speed and directionality for control versus ARHGEF18, ARHGEF3, and ARHGEF11 knockdown. (D) Vector fields 100 min after wound infliction for control versus ARHGEF11 knockdown. Bars, 100 μm . (E) Kymographs of speed and directionality for control versus ARHGEF28 knockdown. (F) Difference in PC3 for speed (top) and directionality (bottom) between GEF-H1 knockdown (two hairpins with 30% and 45% depletion) and control. Each point represents the difference in one well location and the mean of the same day's control experiment. $N = 1$ d, $n = 8$ locations. (G) Validation of screen hits in RhoA GEFs. Each point represents the difference in PC of the indicated variable between GEF knockdown in one well location and the mean of the daily control; N , number of wells; n , number of locations. RHOA: $N = 9$, $n = 47$; ARHGEF18: $N = 6$, $n = 31$; ARHGEF11: $N = 6$, $n = 23$; ARHGEF28: $N = 10$, $n = 53$; ARHGEF3: $N = 9$, $n = 48$. Statistics via Wilcoxon signed rank test: *, $P \leq 0.01$; **, $P \leq 0.001$; ***, $P \leq 0.0001$.

methods) is also required for collective migration (Fig. S2 H), in addition to epithelial tight-junction formation.

RhoA GEFs regulate intercellular communication

Four paralog RHOA-GEFs were identified as hits (Fig. 3 E). Depletion of these GEFs did not dampen wound-healing rate, cell speed, or directionality but enhanced front-to-back propagation of motility. Specifically, depletion of Arhgef18, similarly to depletion of RhoA, accelerated front-to-rear long-range communication (Fig. 4 A). Depletion of Arhgef3, Arhgef11, and Arhgef28 somewhat accelerated long-range communication in speed and directionality and enhanced directionality overall (Fig. 3 G and Fig. 4, B–E). GEF-H1 (also called Lcf or ARHGEF2), a mechanoresponsive RhoA-specific GEF and paralog of the four GEFs identified in the screen (Birkenfeld et al., 2007; Chang et al., 2008; Guilluy et al., 2011b), was excluded from the list, because all hairpins produced a knock-

down <50%. Nonetheless, experiments using these relatively inefficient hairpins revealed a phenotype similar to depletion of RhoA (Fig. 4 F, compare with Fig. 2 E and Fig. S1 H for RhoA).

In a set of validation experiments, each RHOA-GEF hit was replicated with at least two different hairpins in 6–10 wells (a total of 23–53 locations per hit). When assessing multiple replicates, systematic alterations were recognized that were missed by the stringent criteria of the screen (Fig. 4 G). For example, faster long-range communication in directionality was a common outcome of depleting any RhoA-GEF, while increased directionality (Arhgef11, Arhgef28, and Arhgef3) and long-range communication in speed (Arhgef18) were GEF specific. This suggests differential roles among RhoA-GEFs in regulating functions downstream of RhoA signaling.

Transmission of motility guidance cues from cell to cell results in the formation of clusters of cells moving with coordinated trajectories (Zaritsky et al., 2014; Materials and methods). Cluster formation was quantified by recording the fraction

of the monolayer in which cells migrated coordinately. In control experiments, the fraction increased steadily over time (Fig. S3 A), because of an expansion of clusters from the front into the monolayer (Fig. S3, B and C). In analogy to speed and directionality, the spatiotemporal dynamics of cluster formation was captured by three PCs, which encoded magnitude, temporal, and spatial gradients (Fig. S3 D). Importantly, these PCs explicitly capture the strength and propagation of short-range communication. PC1, which was most associated with the wound-healing rate (Fig. S3 E), was increased upon depletion of Arhgef28, while depletion of RhoA, Arhgef18, or Arhgef28 increased PC3 in coordination (Fig. S3 F), showing that under these manipulations, the front and rear of the monolayer are coordinated rapidly after wounding.

Altogether, these results establish roles for RhoA and four of its activating GEFs in regulation intercellular communication.

Actomyosin contractility disturbs intercellular communication downstream of the ARHGEF18-RHOA pathway

Previous work showed that partial down-regulation of actomyosin contractility enhances passive force transmission through cells, whereas high levels of myosin activity “scramble” mechanical signals (Ng et al., 2015). Hence we speculated that the enhancement of long-range communication induced by RhoA and RhoA-GEF depletion is related to reduced myosin activity under these perturbed conditions. To test this we inhibited myosin II directly using blebbistatin or via ROCK inhibition using Y27632. To assess the effect of short versus long-term treatment, the drugs were applied without or with 24 or 48 h preincubation to wounding experiment.

Treatment with low doses of Y27632 (15 or 20 μ M) and blebbistatin (10 μ M) in general increased cell speed and coordination (Fig. S3, G and H). Preincubation for 48 h led to increased long-range communication in directionality and/or increased coordination similar to the behavior observed upon depletion of RhoA/RhoA-GEFs (Fig. S3, G–I). Inhibition of the formin pathway, which is also downstream of RhoA (Higashida et al., 2004; Fig. S3 J), with low doses of SMIFH2 (5 or 10 μ M; Fig. S3, J and K) increased speed, directionality, and coordination without changing long-range communication (Fig. S3 K). Treatment with higher doses of Y27632 (25 μ M for 48 h) or SMIFH2 (25 μ M) inhibited overall motility (Fig. S3, G and K).

The effect of a 48-h drug treatment is expected to be biologically more similar to the effect of protein depletion by shRNA than the effects of more acute treatments; thus, we next compared the effects of blebbistatin and Y27632 treatment over 48 h with the effects of RHOA/RHOA-GEF depletion (Fig. 5 A). We did not consider formin inhibition in this analysis, because of the absence of a communication phenotype that could match the RHOA-GEF depletion phenotypes. Similarity analysis was performed by representing each experiment by a nine-dimensional feature vector composed of the normalized PCs 1–3 for speed, directionality, and coordination and calculating the similarity between every pair of conditions with the L1 norm. The pairwise (symmetric) similarity matrix confirmed distinct but related functionalities for RHOA/RHOA-GEFs/contractility in control of intercellular communication (Fig. 5 B). ARHGEF11, ARHGEF28, and ARHGEF3 fall in a first cluster and ARHGEF18, RHOA, and blebbistatin treatment in a second cluster. Treatments with 15 or 20 μ M Y27632, although similar among each other, differ in their effect from any

of the two other phenotypes. This demonstrates the sensitivity of our analysis to distinguish perturbations of the RhoA-myosin pathway from perturbation of Rock, which in part affects the RhoA-myosin axis but appears to be implicated in additional pathways driving collective migration. More critically, the sensitivity of our assay predicts that ARHGEF18 (and likely also ARHGEF2) regulate specifically RhoA-mediated activation of contractility, while ARHGEF11, ARHGEF28, and ARHGEF3 coregulate RhoA-independent pathways that do not converge on myosin II promoted processes. These may include RHOB/C (but not CDC42 or RAC1) for ARHGEF11 (Rümenapp et al., 1999; Jaiswal et al., 2011), RHOC (but not CDC42 or RAC1) for ARHGEF28 (van Horck et al., 2001; Bravo-Cordero et al., 2011, 2013), and RHOB (but not RHOC) for ARHGEF3 (Arthur et al., 2002), although the abundance of RhoC expression in 16HBE is fourfold less than that of RhoA (Wallace et al., 2011). Acute inhibition of all Rho isoforms (A, B, and C) with the small-molecule inhibitor Rhosin (Shang et al., 2012) showed a general motility reduction in a dose-dependent manner, indicating that Rho isoforms are required for collective migration (Fig. 5 C). We also excluded the possibility of crosstalk between RhoA and RhoC at the expression level. Western blots verified that knockdown of RhoA did not reduce RhoC (Fig. 5 D).

To test the prediction of differential regulation of RhoA and RhoC, we performed a set of RhoC knockdown experiments. Depletion of RhoC increased cell speed and coordination (PC1), long-range communication in directionality, and coordination (PC3; Fig. 5 E). A pairwise similarity matrix indicated that RhoC had an intermediate phenotype between the RhoA-ARHGEF18 and the cluster of the other RhoA-GEFs (Fig. 5, F and G).

Together, these analyses identified ARHGEF18 as the only RHOA-GEF exclusively activating the RHOA isoform (Blomquist et al., 2000; Herder et al., 2013), although biochemically it has also been described as a GEF for RAC1 (but not CDC42; Niu et al., 2003). Arhgef18 activates RhoA at tight junctions, directly interacting with myosin IIA and regulating tight-junction assembly (Terry et al., 2011; Xu et al., 2013; Durgan et al., 2015; Kim et al., 2015). Thus, we interpret the similarity of defects in long-range communication induced by Arhgef18 and RhoA depletion and direct inhibition of myosin II contraction as an indication that ARHGEF18 locates specifically at the top of an ARHGEF18/RHOA/MYO-II pathway (Fig. 5 H). ARHGEF11, ARHGEF28, and ARHGEF3 on one hand target the RHOA/MYO-II pathway. On the other hand, these three GEFs also target migration directionality, which is unaffected by RhoA-mediated signals. The intermediate phenotype of RHOC could be explained by integration of multiple pathways with similar but differential function (Kafri et al., 2009) and/or dynamic interactions between the molecular components that can be regulated in time and space (Guilluy et al., 2011a). The functional similarity between the ARHGEF18 and RHOC phenotypes, notably in cell speed, could be explained by a similar coil domain structure (Cook et al., 2014), predictive of competition, binding, or indirect interactions through other proteins of ARHGEF18 with RHOC. Determining how the ARHGEF18/RhoA-, RhoC-, and ARHGEF3,11,28-mediated pathways mechanistically differentiate between long-range communication, speed, and directionality will require an analysis of the spatiotemporal activation patterns during migration of these upstream GEFs in conjunction with the targeted GTPases, for example by construction of GEF activation biosensors.

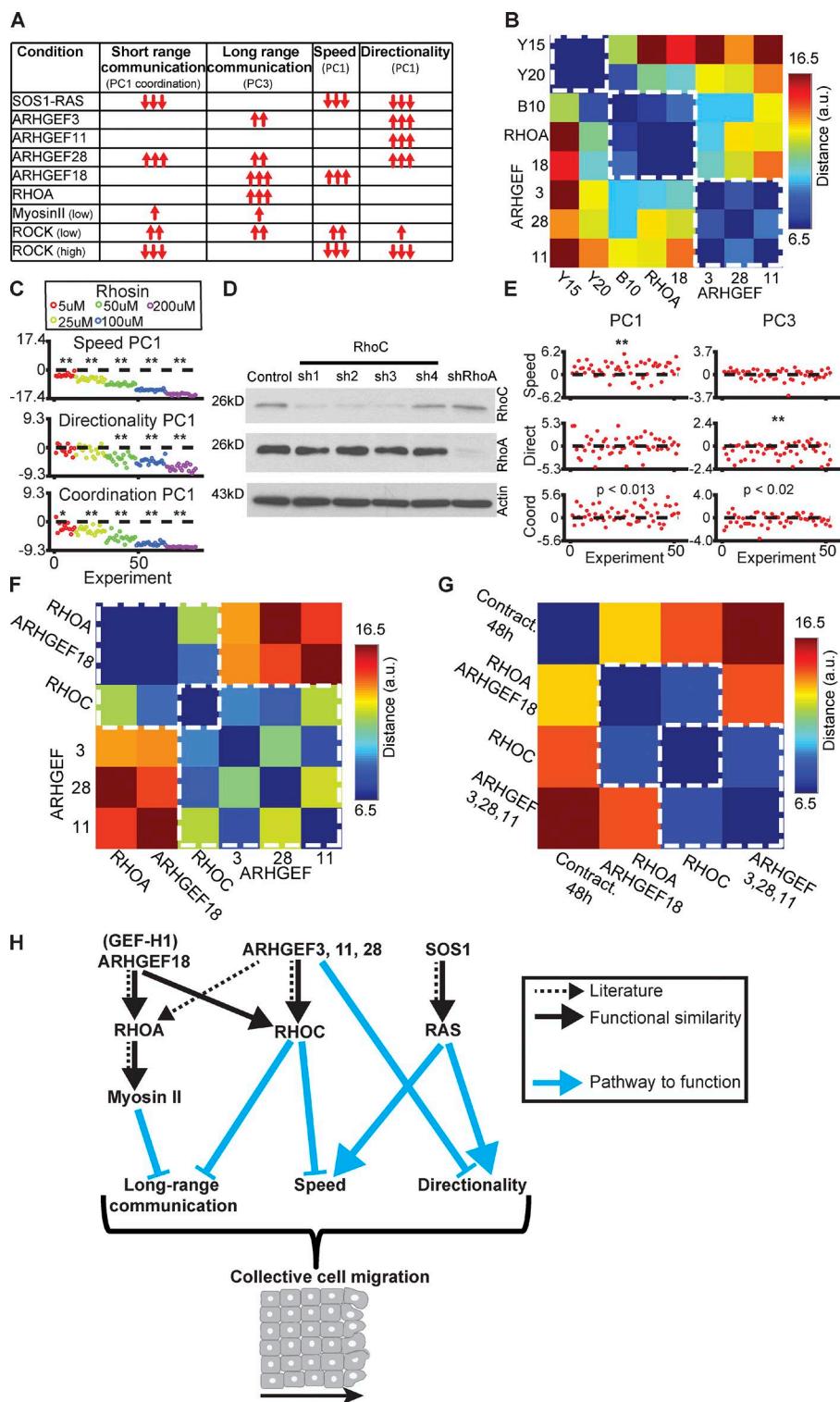


Figure 5. Distinct functional clusters along the RHOA-GEFs/RHOA/actomyosin pathway. (A) Summary of all phenotypic alterations by validated hits of our GEF-screen and for experiments using inhibition of contractility. Number of arrows corresponds to most significant p-value for a given attribute ($P \leq 0.01$, $P \leq 0.001$, or $P \leq 0.0001$). Empty table bins indicate no phenotype. (B) Pairwise distance matrix for knockdown of RhoA, RhoA-GEFs, and contractility experiments. Y, Y27632; B, blebbistatin; number indicates concentration applied over 48 h before wounding; a.u., arbitrary units. Dashed white boxes show functional clusters as defined by close phenotypic similarity. (C) Treatment with Rhosin (5, 25, 50, 100, or 200 μ M) induced motility in a dose-dependent fashion. Each point represents the difference in PC of the indicated variable between drug treatment in one well location and the mean of the daily control; $N = 2$ and $n = 12$ for 5 μ M and $N = 3$ and $n = 18$ for the other concentrations, one control well per daily experiment. Statistics via Wilcoxon signed rank test: *, $P \leq 0.01$; **, $P \leq 0.001$. (D) Western blots to identify expression crosstalk between RhoC and RhoA depletion. (E) Effect on PCs by depletion of RhoC. $N = 9$ and $n = 52$. Statistics via Wilcoxon signed rank test. Coord, coordination; Direct, directionality. (F) Pairwise distance matrix for knockdown of RhoA, RhoC, and RhoA-GEFs. (G) Pairwise distance matrix for the mean cluster phenotype: contractility (Contract.) applied for 48 h, RhoA and Arhgef18, RhoC, Arhgef3, Arhgef28, and Arhgef11. (H) Working model emerging from the phenotypic similarity established in this study and existing literature. Dotted black arrows are biochemical associations described in the literature (see citations in text), continuous black arrows indicate functional similarities, and cyan arrows link protein knockdown to a specific phenotype.

A second limitation of our approach is its inability to identify intercellular communication phenotypes upon overall reduction of motility because of inherent reduction of the spatial and temporal gradients in relation to controls (Fig. 2 E, PC3). For example, TRIO's Rho-targeting GEF domain is activated by a noncanonical Notch signaling pathway (Le Gall et al., 2008; Song and Giniger, 2011), known to be important in intercellular communication (Lai, 2004; Grego-Bessa et al., 2007). On the basis of our findings of a RhoA axis in the regulation of inter-

cellular communication, it would be obvious to hypothesize that this role of TRIO is modulated by its RhoA-targeting DPH2 domains (Bellanger et al., 2000). Unfortunately, this hypothesis could not be tested by knockdown experiments as performed here. Trio knockdown primarily caused an overall reduction of motility, possibly through its activating interaction with Rac1/RhoG mediated by the DPH1 domain. A differential analysis of domain-specific effects in GEFs with multiple GTPase interactions is beyond the scope of this work.

Consistent with several other studies (Vicente-Manzanares et al., 2009; Lim et al., 2010), our data suggest again that although necessary for the generation of motility forces, myosin II-mediated contractility acts as an inhibitor for cell-to-cell communication, likely because contractility intercepts passive mechanical force transduction through the cellular cortex (Ng et al., 2015). Hence, during collective migration myosin II contractility must be adjusted to balance two opposing objectives: generation of robust motile forces versus transmission of mechanical guidance cues. Here we show that signaling pathways upstream of myosin II, such as the RHOA/ROCK pathway, need to be balanced in the same manner.

Materials and methods

Cells and culture conditions

The human bronchial epithelial cell line, 16HBE14o (16HBE), was provided by the laboratory of D.C. Gruenert (University of California, San Francisco, San Francisco, CA). 16HBE cells were cultured in MEM (Memorial Sloan-Kettering Cancer Center [MSKCC] core facility), supplemented with 10% FBS (lot number 169905; Omega Scientific), GlutaMAX (35050; Gibco), and a mixture of penicillin-streptomycin (100×, 10,000 U/ml; 15140; Gibco). Stable cell lines were selected with 1.5 µg/ml puromycin (P7255; Sigma-Aldrich). HEK293T cells were purchased from ATCC and grown in DME high-glucose + sodium pyruvate (MSKCC core facility) supplemented with 10% FBS (lot number 169905; Omega Scientific) and a mixture of penicillin-streptomycin (100×, 10,000 U/ml; 15140; Gibco). All cells were cultured in a 37°C incubator with 5% CO₂.

Antibodies and chemical reagents

Primary antibodies used for Western blotting include ARHGEF18 (EB06163; Everest) at 1:500, BCR (N-20, sc-885; Santa Cruz Biotechnology, Inc.) at 1:1,000, Cdc42 (610929; BD) at 1:1,000, GAPDH (FL-335, sc-25778; Santa Cruz Biotechnology, Inc.) at 1:1,000, Intersectin 2 (H00050618-A01; Abnova) at 1:1,000, LARG (N-14, sc-15439; Santa Cruz Biotechnology, Inc.) at 1:1,000, α-PIX (4573S; Cell Signaling Technology) at 1:1,000, β-PIX (07-1450; EMD Millipore) at 1:2,000, Rac1 (23A8; Abcam) at 1:2,000, RhoA (sc-418; Santa Cruz Biotechnology, Inc.) at 1:500, SOS1 (C-23, sc-256; Santa Cruz Biotechnology, Inc.) at 1:1,000, SOS2 (C-19, sc-258; Santa Cruz Biotechnology, Inc.), Tiam1 (C-16, sc-872; Santa Cruz Biotechnology, Inc.), and α-tubulin (MCA77S; AbD Serotec) at 1:2,000. Secondary polyclonal antibodies conjugated with HRP for Western blot were from Dako and used at 1:5,000. Primary antibodies used for immunofluorescence include E-cadherin (13-1900; Invitrogen) at 1:100 and ZO-1 (61-7300; Invitrogen) at 1:100. Secondary antibodies conjugated with Alexa 488 or Alexa 568 (Invitrogen) were used at 1:400. Other reagents used include Hoechst 33342 (Sigma-Aldrich) at 1 µg/ml, Y27632 (stock in H₂O; HA139; Sigma-Aldrich) at indicated concentrations, blebbistatin (stock in DMSO; 203391; EMD Millipore) at indicated concentrations, ERKi (stock in DMSO; SCH772984; provided by N. Rosen laboratory, MSKCC, New York, NY) at 1 µM, GSK1120212 (stock in DMSO; S2673; Selleckchem) at 500 nM, PD0325901 (stock in DMSO; S1036; Selleckchem) at 500 nM, Rhosin (stock in DMSO; EMD Millipore) at indicated concentrations, and SMIFH2 (stock in DMSO; EMD Millipore) at indicated concentrations.

shRNAs

An shRNA library was constructed in pSUPERpuro, containing at least three hairpins per gene for 80 predicted human Rho GEFs. An

empty pSUPERpuro vector was used as a negative control (termed “control” in the figures and “pSuper” in Table S5). ARHGEF3, Rac1, RhoA, RhoC, and TRIO shRNAs were obtained from The RNAi Consortium library collection, in the pLKO.1 vector (MSKCC RNAi core facility) and a pLKO.1 vector with a nontargeting sequence was used as a negative control (catalog number SHC002; Sigma-Aldrich) for these experiments. The RhoA, Rac1, Cdc42, and RhoC hairpins reported in the study were Cdc42 sh1 (5'-GGAGAACCATATACT CTTG-3'), Cdc42 sh2 (5'-GATGACCCCTACTATTG-3'), Rac1 sh1 (TRCN0000004871; 5'-TTAAGAACACATCTGTTGCG-3'), Rac1 sh2 (TRCN0000004873; 5'-TAATTGTCAAAGACAGTAGGG-3'), RhoA sh1 (TRCN0000047710; 5'-GTACATGGAGTGTTTCAGCA AA-3'), RhoA sh2 (TRCN0000047711; 5'-CGATGTTACTGTAT GTGTT-3'), RhoC sh1 (5'-CTACTGTCTTGAGAACTATA-3'), RhoC sh2 (5'-GCGAACCGGATCAGTGCCTT-3'), RhoC sh3 (5'-TG ATGTCATCCTCATGTGCTT-3'), and RhoC sh4 (5'-GAATAAGAA GGACCTGAGGCA-3'). Hairpin sequences for the Rho GEFs are documented in Tables S2 and S3. Hairpins with unmeasurable knockdown were defined as “off-target” controls and were used to assess the extent of off-target effects and to define thresholds for hit identification in the screen that minimize false detection rates (Fig. 3 D).

Virus production and infection

For virus production, 90% confluent human embryonic kidney 293T cells cultured in a six-well plate were transfected with lentiviral or retroviral constructs using Lipofectamine 2000 (11668; Invitrogen) and Opti-MEM (31985; Invitrogen), according to the manufacturer's instructions. The culture media were removed the day after transfection, and media were collected three times at 24-h intervals. To infect cells, 2 × 10⁵ 16HBE cells were seeded in each well of a six-well plate and incubated on the following day with 1.5 ml virus-containing media supplemented with 1.5 µl polybrene (8 µg/ml stock; Sigma-Aldrich). Spin infection was performed at 2,250 rpm for 30 min. Cells were selected starting 2 d after infection with 1.5 µg/ml puromycin (Sigma-Aldrich). Pooled selected cell lines were used for all experiments.

Primers for PCR and qRT-PCR

Primers were designed using National Center for Biotechnology Information Primer-Blast to target all transcription variants of the gene and to span exon-exon boundaries to avoid amplifying genomic DNA. For qRT-PCR, primers were selected to amplify a PCR product between 70 and 150 bp in length with melting temperatures between 57°C and 63°C. Two sets of primers were examined for each gene, and the primer pair with the highest efficiency by qRT-PCR was selected for quantifying GEF expression. For some genes, QuantiTect primer assays (QIAGEN) were used. All primers used to quantify GEF expression are shown in Table S1.

RNA extraction and PCR

Total RNA was extracted using the RNeasy Plus Mini kit (74134; QIA GEN). cDNAs were prepared using Oligo dT or Random hexamer primers (IDT Technology). In brief, RNA mixtures were heated at 65°C for 5 min, then chilled on ice and mixed with 5× Reaction buffer (Thermo Fisher Scientific), RiboLock RNase inhibitor (EO0381; Thermo Fisher Scientific), dNTPs (Sigma-Aldrich), and RevertAid reverse transcription (EO0441; Thermo Fisher Scientific). Reverse transcription reactions were performed using the following PCR program: 25°C for 10 min, 42°C for 60 min, and 72°C for 10 min, followed by cooling at 4°C.

To examine gene expression, each PCR was performed in a total 20-µl reaction, with 100 ng cDNA as template, 0.5 µM forward and reverse primer, 2 µl 10× PCR buffer (without magnesium; Invitrogen), 1 µM deoxynucleotide triphosphates (Sigma-Aldrich), 1.5 mM MgCl₂,

0.2 μ l Taq polymerase (10342; Invitrogen), and 9.2 μ l H₂O. The PCR program used was 94°C for 3 min; followed by 30 cycles of 94°C for 45 s, 60°C for 30 s, and 72°C for 1 min; followed by 72°C for 10 min and 4°C. PCR products were electrophoresed on a 2% TAE gel and stained with ethidium bromide. For qRT-PCR, 1 μ g cDNA was used per reaction, in 25- μ l reactions containing 1.25 μ l 5 μ M forward and reverse primers and 12.5 μ l Maxima SYBR Green/ROX qPCR Master Mix (2x; K0221; Thermo Fisher Scientific). qRT-PCRs were performed on a Bio-Rad Laboratories iQ5 Multicolor RT-PCR Detection System with the following conditions: 95°C for 10 min; 40 cycles of 95°C for 15 s and 60°C for 60 s for gene expression detection; followed by 71 cycles of 60°C for 30 s, with increase of 0.5°C per cycle for melting curve detection. Gene expressions were normalized by the expression of GAP DH and HPRT and triplicate measurements were used for each sample.

Western blotting

Cells were washed in ice-cold PBS and lysed in ice-cold RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% NP-40, and 0.1% SDS) supplemented with 5 mM Na₃VO₄, 10 mM NaF, 25 mM β -glycerophosphate, and 1 mM PMSF. Lysates were collected by cell scraping and cleared by centrifugation at 13,000 rpm for 1 min at 4°C and boiled in sample buffer (final concentration 50 mM Tris HCl, pH 6.8, 2% SDS, 10% glycerol, 0.1% bromophenol blue, and 100 mM DTT) for 5 min. Protein concentrations were determined by bicinchoninic acid assay (23225, Pierce BCA protein assay kit; Thermo Fisher Scientific). SDS-PAGE on 3%–8% Tris-acetate gels, 4%–12% Bis-Tris gels, or 12% Bis-Tris gels (NuPAGE; Thermo Fisher Scientific); transfer to polyvinylidene fluoride membranes (0.45 mm pore size; EMD Millipore); blocking; antibody binding; and ECL-mediated detection were performed as previously described (Durgan et al., 2015).

Validation of expression levels and knockdown efficiencies

Western blots were used to assess protein expression levels for 9 GEFs that had validated antibodies available. qRT-PCR was applied for the remaining 71 GEFs to assess gene expression levels. The knockdown of 11 GEFs could not be assessed by qRT-PCR because of failure in the production of efficient primers.

Immunofluorescence cell-cell junctions

16HBE cells grown on glass coverslips were washed with PBS and fixed in 3.7% (vol/vol) formaldehyde (Sigma-Aldrich) in PBS for 10 min at room temperature. For immunostaining, coverslips were washed three times in PBS and blocked with BTPA buffer (0.5% BSA, 0.02% sodium azide, and 0.25% Triton X-100 in PBS) for 30 min. After blocking, coverslips were incubated with primary antibodies against ZO-1 and E-cadherin (diluted in BTPA buffer) for 1 h at room temperature and washed three times in PBS for 5 min. Coverslips were then incubated with secondary antibodies and Hoechst (diluted in BTPA buffer) for 1 h at room temperature, washed three times in PBS for 5 min, and mounted onto microscope slides (Thermo Fisher Scientific) using fluorescent mounting media (Dako). Epifluorescence images were acquired with an upright Imager.A1 microscope (ZEISS), equipped with an EC-Plan-NEOFLUAR 40 \times /0.75 objective and a Hamamatsu Photonics Orca-ER 1394 C4742-80 camera, controlled by Axiovision software (ZEISS). Scale bars were added using ImageJ (National Institutes of Health).

Wound-healing assay

For wound-healing assays, 3 \times 10⁶ 16HBE cells were seeded in each well of a six-well tissue culture plate and incubated for 2 d before wounding. Wounding was performed by scratching with a P1000 pipette tip on the confluent monolayers, in the middle of each well, and a cell scraper was used to remove half of the cells from the plate.

After washing with PBS several times to remove cell debris and adding fresh 16HBE media, plates were imaged on an inverted Axiovert 200M microscope (ZEISS) equipped with an EC Plan-NEOFLUAR 10 \times /0.3 Ph1 objective, a Hamamatsu Orca-ER 1394 C4742-80 camera, a 37°C incubator and a CO₂ controller, controlled by Axiovision software (ZEISS). The time-lapse image sequences were recorded for 16 h at 5-min intervals.

Velocity measurements

Velocity fields were computed using custom cross-correlation-based particle image velocimetry using nonoverlapping image patches of size 15 \times 15 μ m (Zaritsky et al., 2012). The frame-to-frame displacement of each patch was defined by the maximal cross-correlation of a given patch with the subsequent image in the time-lapse image sequence. The search radius was constrained to an instantaneous speed of 90 μ m h⁻¹. At a frame rate of 5 min, this search radius corresponds to half the side length of a 15 \times 15 μ m patch.

Segmentation of monolayer contours

Monolayer contours were calculated with a segmentation algorithm that classified the image regions as either “cellular foreground” or “background.” Our goal was to optimize the robustness of the segmentation to enable unsupervised analysis of thousands of movies, each containing tens of time points. Small inaccuracies in the segmentation have little effect on the resulting kymograph, every movie was manually validated, and fewer than 1% of the well locations were discarded on the grounds of failed segmentation. To achieve robustness, we introduced several priors to the algorithm: (1) the image contains one continuous region of “cellular foreground” and one continuous region of “background,” and (2) the contour advances monotonically over time. These priors allowed us to estimate the initial contour at time 0 and then use the segmentation at time t as a seed to expand the “cellular foreground” to time $t+1$. The only pixels in question are those labeled “background” at time t and close enough to the “cellular foreground” region. The proximity threshold is calculated on the basis of the maximal cell velocity of 90 μ m h⁻¹.

The segmentation was performed in super-pixels with a size equivalent to a 15 \times 15 μ m patch. After cross-correlation-based particle image velocimetry, which assigns to each super-pixel a displacement vector and a cross-correlation score defining the match of the corresponding patch signals between consecutive frames, we took advantage of the observation that super-pixels associated with “cellular foreground,” especially those at the front of the monolayer, had lower cross-correlation scores because of their textured and dynamic nature. We thus assigned to each super-pixel a pseudo-intensity value (1 – cross-correlation score) and applied the Rosin thresholding method (Rosin, 2001) to label the super-pixels as “cellular foreground” versus “background.” Background super-pixels enclosed by foreground regions were relabeled as “cellular foreground,” and isolated “cellular foreground” super-pixels, usually attributed to debris or textured plate patterns, were relabeled as “background.” We next calculated the union of the “cellular foreground” region in the previous frame with the new “cellular foreground” region. Morphological closing and hole filling defined the final segmentation. For the first image in the time-lapse sequence, we defined the “cellular foreground” as the union of the cross-correlation-based segmentation and a texture-based segmentation. The latter was implemented by first representing each super-pixel by the distribution of its local binary patterns, a widely used representation of local texture in images (Ojala et al., 2002), followed by unsupervised K-means clustering with K = 2. This heuristic was found to robustly complement the cross-correlation-based segmentation by identifying super-pixels mislabeled as “background” at larger distances

from the wound edge, where cells do not sufficiently migrate and change appearance to generate a low enough cross-correlation score.

Correction for microscope repositioning error

During the multilocation filming, the microscope stage exhibited repositioning errors that caused systematic shifts in all vector fields. We estimated the stage shift for every image in a time-lapse sequence and corrected the vector fields accordingly. To accomplish this, we exploited patches in the background region that are expected to stay in the same position and subtracted their robust mean velocity from the vector fields in the “cellular foreground.”

Kymographs

Speed kymographs were constructed by calculating the mean speed of all patches in spatial bands of 15 μm from the monolayer’s front through time. Cell directionality is defined as the absolute ratio between the velocity component perpendicular to the monolayer edge and the velocity component parallel to the monolayer edge. Each bin in the directionality kymograph was calculated as the ratio obtained by the two-component decomposition of the speed kymograph to a component normal and parallel to the monolayer front. These components were calculated by considering the orientation of the wound edge. Coordination (or short-range communication) was measured by detecting clusters of cells moving in coordination. Coordination kymographs were constructed by recording the fraction of patches that participated in these coordinated clusters for every spatial band and time frame. Explicit detection of coordinated clusters was performed for every time frame by applying region-growing spatial clustering of the image-patch grid on the basis of the correlation of their velocity fields, as previously described (Zaritsky et al., 2014). In brief, region-growing segmentation (Nock and Nielsen, 2004) started with regions containing a single image patch and iteratively merging spatially adjacent patches on the basis of their velocity vector similarity. Two regions are merged if their similarity is lower than a given threshold, and the merged region vector is updated to be the mean of all contributing patch vectors. Merging is performed in ascending order of the similarity between the adjacent patches. Kymographs were calculated for cells located up to 180 μm from the monolayer and for the first 200 min after wound infliction. We chose 180 μm because the cell monolayers captured in the field of view of the vast majority of experiments were equal or wider than this value. We chose 200 min to focus our analysis on the transient phase between wound infliction and steady-state collective migration. Extending the time window to 400 min did not change the conclusions drawn from kymograph analysis.

Calculation of PCs

To obtain a compact representation of the kymographs, we averaged the kymographs in 3×4 bins at a resolution of 60 μm and 50 min (Zaritsky et al., 2012). To further reduce this 12-dimensional feature vector, we applied PCA (Jolliffe, 2002). The PCs are ranked by the spread of the data they capture. Mathematically, this is equivalent to the eigenvalues of the data covariance matrix. We used the 402 control experiments of the screen to calculate the PCA transformation for speed, directionality, and coordination and then applied the same transformation on control and knockdown experiments. First, the features were normalized (Fig. 2 B) to $x' = (x - \mu)/\sigma$, with feature x , mean μ , and standard deviation σ of the set of control experiments. The PCA transformation was calculated and applied to normalized features. Nearly identical PCA transformations were found when including the knockdown experiments (and for non-normalized features), indicating that the functional fluctuations between control and knockdown experiments are small compared with the day-to-day variation of control experiments. The coefficients of each

PC defined the projection of the 12-dimensional feature vector into the direction of the PC. The coefficients of PC1 were similar for all 12 features, implying that PC1 encodes the mean of speed, directionality, or coordination (dependent on the considered kymograph) across time and space. The coefficients of PC2 and PC3 reflected the pattern of a temporal (PC2) or spatial (PC3) gradient (Figs. 2 D, S1 F, and S3 D; see text). It should be noted that conditions with overall reduced motility, such as with CDC42 and RAC1 knockdown, not only reduce PC1 but inherently flatten the spatial and temporal gradients in speed, thus reducing PC3 but increasing PC2 of this measure. This increase of PC2 is due to the positive coefficients for the (slower) speed at the onset of migration, which is minimally reduced under these perturbations, while the negative coefficients for reduced speed of perturbed experiment at later times increase overall values. In Figs. 2 E; 4, F–G; 5, C and E; S1 H; S2 H; and S3, F, G, and K, we depict the shifts induced by knockdown by subtraction of the PC of control experiments from the PC of the knockdown experiment. For PC1, negative values indicate reductions in speed, directionality, or coordination. For PC3, which reflects the spatial gradient, negative difference values imply more immediate front-to-back propagation and faster establishment of the steady state, while for PC2 this is reversed.

Scoring a knockdown well in relation to its daily control well

Time-lapse image sequences were acquired in at least three different locations in every knockdown and in every control well. To identify hits in the screen by comparison of knockdown and control conditions, we used metrics that take into account the variability within a well and between wells. Specifically, we applied three metrics: (1) the Davies-Bouldin index (Davies and Bouldin, 1979):

$$\frac{d(c_1, c_2)}{\sigma_1 + \sigma_2},$$

which divides the difference $d(c_1, c_2)$ between the per-well mean values of a measurement extracted from multiple locations in the knockdown and control wells by the sum of the mean distances σ_i between the measurement from individual locations in the same well and the corresponding well mean value; (2) the Dunn index (Dunn, 1974):

$$\frac{d(c_1, c_2)}{\max(\rho_1, \rho_2)},$$

which divides the difference between per-well mean values by the maximal distance in any one well between the measurement from individual locations and the corresponding well mean value; (3) the silhouette coefficient (Rousseeuw, 1987):

$$\frac{\sum_{i=1}^n s_i + \sum_{j=1}^m s_j}{n + m},$$

where n and m are the numbers of locations in control and knockdown wells, and

$$\forall k, s_k = \frac{b - a}{\max(a, b)},$$

where a is the mean difference between the measurement extracted in location k and the measurements extracted in all other locations of the same well, and b is the mean difference between the measurement in location k and the measurements extracted in all locations of the other well.

Off-target control experiments (see text) were used to normalize the three metrics across different PC measurements and to define a z -score = $(x - \mu)/\sigma$, with μ and σ denoting the mean and standard deviation of a particular PC metric in the off-target control population and x denoting the metrics of that PC for a knockdown experiment. We defined a single combined z -score scalar for every PC measure by accumulating the three metric-specific z -scores.

Selecting threshold for identification of hits

The 24 off-target and 18 positive controls were used to calculate a threshold for hit identification in the screen, which has the desired property of minimizing the number of false alarms to allow us to focus on real hits for follow-up experiments. Because the positive controls (CDC42, RAC1, and β -PIX) were selected on the basis of their known effect in reducing motility, the threshold was calculated to minimize false alarms in wound-healing rate. Calculating z -scores by standardization relative to the off-target controls made the threshold transferable to the other PC measures. Specifically, we calculated the z -score threshold that maximizes the F -measure, defined as

$$F\text{-measure} = 2 \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}},$$

with recall

$$\left(\frac{\text{TP}}{\text{TP} + \text{FN}} \right)$$

and precision

$$\left(\frac{\text{TP}}{\text{TP} + \text{FP}} \right)$$

derived from the true-positive (TP), false-negative (FN), and false-positive (FP) rates. The true-positive and false-positive rates were estimated on the basis of the positive controls, whereas the false-negative rate was estimated on the basis of the off-target controls. Because of the small number of control experiments, the threshold had low confidence. To resolve this, we applied bootstrapping to estimate the distribution of the maximal F -measure thresholds. Off-target and positive controls were randomly selected with repetition from the original groups. The process was repeated 10,000 times to define the distribution of z -score values maximizing the F -measure, and the final threshold of z -score = 9.8 was selected with 99.9% confidence from this distribution. There were no false-positive occurrences in five of six PC measures, and one false positive occurred for the sixth PC measure (directionality PC2); the false-negative rate was 44.4%.

Analysis of follow-up experiments

To validate the hits that were identified in the screen, we replicated experiments and statistically confirmed the phenotypes. For every experiment, we subtracted the mean PC measure of the daily control locations from the corresponding PC measure in every knockdown location. Thus, each data point represents the deviation of a location to the mean control. The null hypothesis was that the data come from a distribution whose median is zero. Statistical significance was inferred using the nonparametric Wilcoxon signed rank test, and a p-value of 0.01 was selected as the significance threshold. At least two different hairpins were required to validate a hit. SOS1 experiments were replicated with N (number of independent experiments) = 3–6 and n (total number of locations in knockdown wells) = 18–26; RHOA-GEFs were validated with n = 6–9 and n = 23–53 and contractility experiments (treated for 48 h) with n = 3 and n = 18. RhoGTPases were validated with n = 6–9 and n = 24–47 and β -PIX with n = 3 and n = 11.

Calculating similarities between different conditions

To calculate overall similarities between different experimental conditions, we included the three first PCs for speed, directionality, and coordination in a nine-dimensional vector for each well location of a knockdown experiment. As in the follow-up analysis, PC measures in the locations of a knockdown well were related to the mean of the corresponding PC measures in all locations of the daily control well. An

experimental condition was represented as the mean nine-dimensional vector of all locations in all wells under the same treatment regimen. This defined a matrix with nine rows (one per measure) and k columns, where k is the number of different conditions examined for the analysis (e.g., eight for the analysis described in Fig. 5 B). Each of the nine measures was standardized across conditions by subtracting the mean and dividing by the standard deviation (z -scores) for each row in the matrix. This encoded the divergence of every specific condition from the mean across the examined experimental conditions. Similarity was calculated between every pair of conditions (columns) in nine dimensions of the standardized measures using the $\|L1\|$ distance metric. High values reflect dissimilarity, and low values represent more similar phenotypes in this nine-dimensional space.

Data and source-code availability

All following metadata are made available: (1) Primers used to quantify GEF expression are listed in Table S1. (2) shRNA hairpin sequences and knockdown efficiency (Western blots and qRT-PCR) are listed in Table S2 and summarized in Table S3. (3) Details of all screen experiments are provided in Table S4 and of all screen plus follow-up studies in Table S5: molecular perturbation, raw-data file name, file format (.zvi or .tif), physical pixel size in micrometers, measured knockdown efficiency (%), and experiment date. (4) Screen z -scores are provided in Table S6 and coordination z -scores from follow-up experiments in Table S7.

All raw and processed data are available upon request. These data include raw imaging data, processed kymographs (per location and mean per day), and quantification of phenotypes for genes that were followed up. The full data set size is 3 TB. Potential requesters are asked to send a drive of sufficient size.

The Matlab source code is available to the public at <https://github.com/DanuserLab/MonolayerKymographs>. It enables calculation of the velocity fields and segmentation of the monolayer front and the wound-healing rate for every time frame in the raw time-lapse images, and it generates kymographs for speed, directionality, and coordination for a full video. It will also produce the 12-dimensional feature vector representation and PC projection on the basis of the transformation that was calculated for this study. We also provide a script to perform PCA and compute the transformation given a set of high-dimensional experiments. Input is as follows: $M = d \times n$ matrix, where d is the number of dimensions (in our case, d = 12), and n is the number of experiments. Output is a PCA transformation (a wrapper for Matlab's PCA code). This code includes means of visualization of the PCA weights (such as in Fig. 2 D) to enable interpretation. Documentation and a test example are included with the source code.

Online supplemental material

Fig. S1 depicts alterations of collective motility induced by depletion of Rho GTPases or the GEF β -PIX. Fig. S2 shows screen hits that were not followed up (A–E) and follows the SOS1-RAS pathway regulation of collective cell migration. Fig. S3 shows the effects of RhoA GEF depletion on short-range communication as quantified by the coordination parameter (A–F) and the effects of perturbation of actomyosin contractility on intercellular communication (G–I). Video 1 shows phase-contrast live imaging of the wound-healing response of 16HBE cells. Tables S1–S7 are available as Excel files. Primers used to quantify GEF expression are listed in Table S1. shRNA hairpin sequences and KD efficiency (Western blots and qRT-PCR) are listed in Table S2 and summarized in Table S3. Details of all screen experiments are provided in Table S4 and of all screen plus follow-up studies in Table S5. Screen z scores are provided in Table S6 and coordination z -scores from follow-up experiments in Table S7.

Acknowledgments

We thank A. Jamieson for helping package the source code and creating the repository. We thank C. Schaefer and N. Bendris for fruitful discussions and advice. We dedicate this work to our mentor and colleague A. Hall, who initiated this project.

This work was supported by the National Institutes of Health (P01 GM103723 to G. Danuser, A. Hall, and M. Overholtzer; P30 CA008748 to M. Overholtzer).

The authors declare no competing financial interests.

Author contributions: A. Hall and G. Danuser conceived the study. M. Overholtzer, A. Hall, and G. Danuser guided A. Zaritsky, Y.-Y. Tseng, MA. Rabadán, and S. Krishna. A. Zaritsky, Y.-Y. Tseng, and MA. Rabadán designed the experiments. Y.-Y. Tseng, MA. Rabadán, and S. Krishna performed all experiments. A. Zaritsky developed analytic tools and analyzed the data. A. Zaritsky and Y.-Y. Tseng interpreted the data. A. Zaritsky drafted the manuscript. All authors wrote and edited the manuscript and approve of its content.

Submitted: 21 September 2016

Revised: 14 March 2017

Accepted: 12 April 2017

References

- Arthur, W.T., S.M. Ellerbroek, C.J. Der, K. Burridge, and K. Wennerberg. 2002. XPLN, a guanine nucleotide exchange factor for RhoA and RhoB, but not RhoC. *J. Biol. Chem.* 277:42964–42972. <http://dx.doi.org/10.1074/jbc.M207401200>
- Bazellières, E., V. Conte, A. Elosegui-Artola, X. Serra-Picamal, M. Bintanel-Morillo, P. Roca-Cusachs, J.J. Muñoz, M. Sales-Pardo, R. Guimerà, and X. Trepat. 2015. Control of cell-cell forces and collective cell dynamics by the intercellular adhesome. *Nat. Cell Biol.* 17:409–420. <http://dx.doi.org/10.1038/ncb3135>
- Bellanger, J.-M., C. Astier, C. Sardet, Y. Ohta, T.P. Stossel, and A. Debant. 2000. The Rac1- and RhoG-specific GEF domain of Trio targets filamin to remodel cytoskeletal actin. *Nat. Cell Biol.* 2:888–892. <http://dx.doi.org/10.1038/35046533>
- Birkenfeld, J., P. Nalbant, B.P. Bohl, O. Pertz, K.M. Hahn, and G.M. Bokoch. 2007. GEF-H1 modulates localized RhoA activation during cytokinesis under the control of mitotic kinases. *Dev. Cell.* 12:699–712. <http://dx.doi.org/10.1016/j.devcel.2007.03.014>
- Blomquist, A., G. Schwörer, H. Schablowski, A. Psoma, M. Lehnen, K.H. Jakobs, and U. Rümenapp. 2000. Identification and characterization of a novel Rho-specific guanine nucleotide exchange factor. *Biochem. J.* 352:319–325. <http://dx.doi.org/10.1042/bj3520319>
- Bravo-Cordero, J.J., M. Oser, X. Chen, R. Eddy, L. Hodgson, and J. Condeelis. 2011. A novel spatiotemporal RhoC activation pathway locally regulates cofilin activity at invadopodia. *Curr. Biol.* 21:635–644. <http://dx.doi.org/10.1016/j.cub.2011.03.039>
- Bravo-Cordero, J.J., V.P. Sharma, M. Roh-Johnson, X. Chen, R. Eddy, J. Condeelis, and L. Hodgson. 2013. Spatial regulation of RhoC activity defines protrusion formation in migrating cells. *J. Cell Sci.* 126:3356–3369. <http://dx.doi.org/10.1242/jcs.123547>
- Cai, D., S.-C. Chen, M. Prasad, L. He, X. Wang, V. Choesmel-Cadamuro, J.K. Sawyer, G. Danuser, and D.J. Montell. 2014. Mechanical feedback through E-cadherin promotes direction sensing during collective cell migration. *Cell.* 157:1146–1159. <http://dx.doi.org/10.1016/j.cell.2014.03.045>
- Chang, Y.-C., P. Nalbant, J. Birkenfeld, Z.-F. Chang, and G.M. Bokoch. 2008. GEF-H1 couples nocodazole-induced microtubule disassembly to cell contractility via RhoA. *Mol. Biol. Cell.* 19:2147–2153. <http://dx.doi.org/10.1091/mbc.E07-12-1269>
- Cherfils, J., and M. Zeghouf. 2013. Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol. Rev.* 93:269–309. <http://dx.doi.org/10.1152/physrev.00003.2012>
- Cook, D.R., K.L. Rossman, and C.J. Der. 2014. Rho guanine nucleotide exchange factors: regulators of Rho GTPase activity in development and disease. *Oncogene.* 33:4021–4035. <http://dx.doi.org/10.1038/onc.2013.362>
- Das, T., K. Safferling, S. Rausch, N. Grabe, H. Boehm, and J.P. Spatz. 2015. A molecular mechanotransduction pathway regulates collective migration of epithelial cells. *Nat. Cell Biol.* 17:276–287. <http://dx.doi.org/10.1038/ncb3115>
- Davies, D.L., and D.W. Bouldin. 1979. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* 1:224–227.
- Dunn, J.C. 1974. Well-separated clusters and optimal fuzzy partitions. *J. Cybern.* 4:95–104.
- Durgan, J., G. Tao, M.S. Walters, O. Florey, A. Schmidt, V. Arbelaez, N. Rosen, R.G. Crystal, and A. Hall. 2015. SOS1 and Ras regulate epithelial tight junction formation in the human airway through EMP1. *EMBO Rep.* 16:87–96. <http://dx.doi.org/10.1525/embr.201439218>
- Goicoechea, S.M., S. Awadia, and R. Garcia-Mata. 2014. I'm coming to GEF you: regulation of RhoGEFs during cell migration. *Cell Adhes. Migr.* 8:535–549. <http://dx.doi.org/10.4161/cam.28721>
- Grego-Bessa, J., L. Luna-Zurita, G. del Monte, V. Bolós, P. Melgar, A. Arandilla, A.N. Garratt, H. Zang, Y.S. Mukouyama, H. Chen, et al. 2007. Notch signaling is essential for ventricular chamber development. *Dev. Cell.* 12:415–429. <http://dx.doi.org/10.1016/j.devcel.2006.12.011>
- Guilluy, C., R. Garcia-Mata, and K. Burridge. 2011a. Rho protein crosstalk: another social network? *Trends Cell Biol.* 21:718–726. <http://dx.doi.org/10.1016/j.tcb.2011.08.002>
- Guilluy, C., V. Swaminathan, R. Garcia-Mata, E.T. O'Brien, R. Superfine, and K. Burridge. 2011b. The Rho GEFs LARG and GEF-H1 regulate the mechanical response to force on integrins. *Nat. Cell Biol.* 13:722–727. <http://dx.doi.org/10.1038/ncb2254>
- Hayer, A., L. Shao, M. Chung, L.-M. Joubert, H.W. Yang, F.-C. Tsai, A. Bisaria, E. Betzig, and T. Meyer. 2016. Engulfed cadherin fingers are polarized junctional structures between collectively migrating endothelial cells. *Nat. Cell Biol.* 18:1311–1323. <http://dx.doi.org/10.1038/ncb3438>
- Herder, C., J.M. Swiercz, C. Müller, R. Perivali, R. Quiring, S. Offermanns, J. Wittbrodt, and F. Loosli. 2013. ArhGEF18 regulates RhoA-Rock2 signaling to maintain neuro-epithelial apico-basal polarity and proliferation. *Development.* 140:2787–2797. <http://dx.doi.org/10.1242/dev.096487>
- Hidalgo-Carcedo, C., S. Hooper, S.I. Chaudhry, P. Williamson, K. Harrington, B. Leitinger, and E. Sahai. 2011. Collective cell migration requires suppression of actomyosin at cell-cell contacts mediated by DDR1 and the cell polarity regulators Par3 and Par6. *Nat. Cell Biol.* 13:49–58. <http://dx.doi.org/10.1038/ncb2133>
- Higashida, C., T. Miyoshi, A. Fujita, F. Ocegueda-Yanez, J. Monypenny, Y. Andou, S. Narumiya, and N. Watanabe. 2004. Actin polymerization-driven molecular movement of mDia1 in living cells. *Science.* 303:2007–2010. <http://dx.doi.org/10.1126/science.1093923>
- Hodge, R.G., and A.J. Ridley. 2016. Regulating Rho GTPases and their regulators. *Nat. Rev. Mol. Cell Biol.* 17:496–510. <http://dx.doi.org/10.1038/nrm.2016.67>
- Jaffe, A.B., and A. Hall. 2005. Rho GTPases: biochemistry and biology. *Annu. Rev. Cell Dev. Biol.* 21:247–269. <http://dx.doi.org/10.1146/annurev.cellbio.21.020604.150721>
- Jaiswal, M., L. Gremer, R. Dvorsky, L.C. Haeusler, I.C. Cirstea, K. Uhlenbrock, and M.R. Ahmadian. 2011. Mechanistic insights into specificity, activity, and regulatory elements of the regulator of G-protein signaling (RGS)-containing Rho-specific guanine nucleotide exchange factors (GEFs) p115, PDZ-RhoGEF (PRG), and leukemia-associated RhoGEF (LARG). *J. Biol. Chem.* 286:18202–18212. <http://dx.doi.org/10.1074/jbc.M111.226431>
- Jolliffe, I. 2002. Principal Component Analysis. Springer, New York. 488 pp.
- Kafri, R., M. Springer, and Y. Pilpel. 2009. Genetic redundancy: new tricks for old genes. *Cell.* 136:389–392. <http://dx.doi.org/10.1016/j.cell.2009.01.027>
- Kim, J.H., X. Serra-Picamal, D.T. Tambe, E.H. Zhou, C.Y. Park, M. Sadati, J.A. Park, R. Krishnan, B. Gweon, E. Millet, et al. 2013. Propulsion and navigation within the advancing monolayer sheet. *Nat. Mater.* 12:856–863. <http://dx.doi.org/10.1038/nmat3689>
- Kim, M., A. M Shewan, A.J. Ewald, Z. Werb, and K.E. Mostov. 2015. p114RhoGEF governs cell motility and lumen formation during tubulogenesis through a ROCK-myosin-II pathway. *J. Cell Sci.* 128:4317–4327. <http://dx.doi.org/10.1242/jcs.172361>
- Ladoux, B., R.-M. Mège, and X. Trepat. 2016. Front-rear polarization by mechanical cues: from single cells to tissues. *Trends Cell Biol.* 26:420–433. <http://dx.doi.org/10.1016/j.tcb.2016.02.002>
- Lai, E.C. 2004. Notch signaling: control of cell communication and cell fate. *Development.* 131:965–973. <http://dx.doi.org/10.1242/dev.01074>
- Le Gall, M., C. De Mattei, and E. Giniger. 2008. Molecular separation of two signaling pathways for the receptor, Notch. *Dev. Biol.* 313:556–567. <http://dx.doi.org/10.1016/j.ydbio.2007.10.030>
- Lim, J.I., M. Sabouri-Ghomi, M. Machacek, C.M. Waterman, and G. Danuser. 2010. Protrusion and actin assembly are coupled to the organization of

- lamellar contractile structures. *Exp. Cell Res.* 316:2027–2041. <http://dx.doi.org/10.1016/j.yexcr.2010.04.011>
- Matsabayashi, Y., W. Razzell, and P. Martin. 2011. ‘White wave’ analysis of epithelial scratch wound healing reveals how cells mobilise back from the leading edge in a myosin-II-dependent fashion. *J. Cell Sci.* 124:1017–1021. <http://dx.doi.org/10.1242/jcs.080853>
- Mayor, R., and S. Etienne-Manneville. 2016. The front and rear of collective cell migration. *Nat. Rev. Mol. Cell Biol.* 17:97–109. <http://dx.doi.org/10.1038/nrm.2015.14>
- Ng, M.R., A. Besser, G. Danuser, and J.S. Brugge. 2012. Substrate stiffness regulates cadherin-dependent collective migration through myosin-II contractility. *J. Cell Biol.* 199:545–563. <http://dx.doi.org/10.1083/jcb.201207148>
- Ng, M.R., A. Besser, J.S. Brugge, and G. Danuser. 2015. Correction: mapping the dynamics of force transduction at cell-cell junctions of epithelial clusters. *eLife.* 4:e06656. <http://dx.doi.org/10.7554/eLife.06656>
- Nimnuan, A.S., B.A. Yatsula, and D. Bar-Sagi. 1998. Coupling of Ras and Rac guanosine triphosphatases through the Ras exchanger Sos. *Science.* 279:560–563. <http://dx.doi.org/10.1126/science.279.5350.560>
- Niu, J., J. Profirovic, H. Pan, R. Vaikunaite, and T. Voyno-Yasenetskaya. 2003. G protein $\beta\gamma$ subunits stimulate p114RhoGEF, a guanine nucleotide exchange factor for RhoA and Rac1: regulation of cell shape and reactive oxygen species production. *Circ. Res.* 93:848–856. <http://dx.doi.org/10.1161/01.RES.0000097607.14733.0C>
- Nock, R., and F. Nielsen. 2004. Statistical region merging. *IEEE Trans. Pattern Anal. Mach. Intell.* 26:1452–1458. <http://dx.doi.org/10.1109/TPAMI.2004.110>
- Notbohm, J., S. Banerjee, K.J. Utuaje, B. Gweon, H. Jang, Y. Park, J. Shin, J.P. Butler, J.J. Fredberg, and M.C. Marchetti. 2016. Cellular contraction and polarization drive collective cellular motion. *Biophys. J.* 110:2729–2738. <http://dx.doi.org/10.1016/j.bpj.2016.05.019>
- Ojala, T., M. Pietikainen, and T. Maenpaa. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 24:971–987. <http://dx.doi.org/10.1109/TPAMI.2002.1017623>
- Omelchenko, T., and A. Hall. 2012. Myosin-IXA regulates collective epithelial cell migration by targeting RhoGAP activity to cell-cell junctions. *Curr. Biol.* 22:278–288. <http://dx.doi.org/10.1016/j.cub.2012.01.014>
- Omelchenko, T., M.A. Rabadian, R. Hernández-Martínez, J. Grego-Bessa, K.V. Anderson, and A. Hall. 2014. β -Pix directs collective migration of anterior visceral endoderm cells in the early mouse embryo. *Genes Dev.* 28:2764–2777. <http://dx.doi.org/10.1101/gad.251371.114>
- Pascual-Vargas, P., S. Cooper, J. Sero, V. Bousgouni, M. Arias-Garcia, and C. Bakal. 2017. RNAi screens for Rho GTPase regulators of cell shape and YAP/TAZ localisation in triple negative breast cancer. *Sci. Data.* 4:170018. <http://dx.doi.org/10.1038/sdata.2017.18>
- Plutoni, C., E. Bazellieres, M. Le Borgne-Rochet, F. Comunale, A. Brugues, M. Séveno, D. Planchon, S. Thuault, N. Morin, S. Bodin, et al. 2016. P-cadherin promotes collective cell migration via a Cdc42-mediated increase in mechanical forces. *J. Cell Biol.* 212:199–217. <http://dx.doi.org/10.1083/jcb.201505105>
- Reffay, M., M.C. Parrini, O. Cochet-Escartin, B. Ladoux, A. Buguin, S. Coscoy, F. Amblard, J. Camonis, and P. Silberzan. 2014. Interplay of RhoA and mechanical forces in collective cell migration driven by leader cells. *Nat. Cell Biol.* 16:217–223. <http://dx.doi.org/10.1038/ncb2917>
- Rosin, P.L. 2001. Unimodal thresholding. *Pattern Recognit.* 34:2083–2096. [http://dx.doi.org/10.1016/S0031-3203\(00\)00136-9](http://dx.doi.org/10.1016/S0031-3203(00)00136-9)
- Rousseeuw, P.J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20:53–65. [http://dx.doi.org/10.1016/0377-0427\(87\)90125-7](http://dx.doi.org/10.1016/0377-0427(87)90125-7)
- Rümenapp, U., A. Blomquist, G. Schwörer, H. Schablowski, A. Psoma, and K.H. Jakobs. 1999. Rho-specific binding and guanine nucleotide exchange catalysis by KIAA0380, a dbl family member. *FEBS Lett.* 459:313–318. [http://dx.doi.org/10.1016/S0014-5793\(99\)01270-3](http://dx.doi.org/10.1016/S0014-5793(99)01270-3)
- Schmidt, A., and A. Hall. 2002. Guanine nucleotide exchange factors for Rho GTPases: turning on the switch. *Genes Dev.* 16:1587–1609. <http://dx.doi.org/10.1101/gad.1003302>
- Serra-Picamal, X., V. Conte, R. Vincent, E. Anon, D.T. Tambe, E. Bazellieres, J.P. Butler, J.J. Fredberg, and X. Trepat. 2012. Mechanical waves during tissue expansion. *Nat. Phys.* 8:628–634. <http://dx.doi.org/10.1038/nphys2355>
- Shang, X., F. Marchioni, N. Sipes, C.R. Evelyn, M. Jerabek-Willemsen, S. Duhr, W. Seibel, M. Wortman, and Y. Zheng. 2012. Rational design of small molecule inhibitors targeting RhoA subfamily Rho GTPases. *Chem. Biol.* 19:699–710. <http://dx.doi.org/10.1016/j.chembiol.2012.05.009>
- Simpson, K.J., L.M. Selfors, J. Bui, A. Reynolds, D. Leake, A. Khvorova, and J.S. Brugge. 2008. Identification of genes that regulate epithelial cell migration using an siRNA screening approach. *Nat. Cell Biol.* 10:1027–1038. <http://dx.doi.org/10.1038/ncb1762>
- Song, J.K., and E. Giniger. 2011. Noncanonical Notch function in motor axon guidance is mediated by Rac GTPase and the GEF1 domain of Trio. *Dev. Dyn.* 240:324–332. <http://dx.doi.org/10.1002/dvdy.22525>
- Tambe, D.T., C.C. Hardin, T.E. Angelini, K. Rajendran, C.Y. Park, X. Serra-Picamal, E.H.H. Zhou, M.H. Žaman, J.P. Butler, D.A. Weitz, et al. 2011. Collective cell guidance by cooperative intercellular forces. *Nat. Mater.* 10:469–475. <http://dx.doi.org/10.1038/nmat3025>
- Terry, S.J., C. Zihni, A. Elbediwy, E. Vitiello, I.V. Leefa Chong San, M.S. Balda, and K. Matter. 2011. Spatially restricted activation of RhoA signalling at epithelial junctions by p114RhoGEF drives junction formation and morphogenesis. *Nat. Cell Biol.* 13:159–166. <http://dx.doi.org/10.1038/ncb2156>
- Timpson, P., E.J. McGhee, J.P. Morton, A. von Kriegsheim, J.P. Schwarz, S.A. Karim, B. Doyle, J.A. Quinn, N.O. Carragher, M. Edward, et al. 2011. Spatial regulation of RhoA activity during pancreatic cancer cell invasion driven by mutant p53. *Cancer Res.* 71:747–757. <http://dx.doi.org/10.1158/0008-5472.CAN-10-2267>
- van Horck, F.P., M.R. Ahmadian, L.C. Haeusler, W.H. Moolenaar, and O. Kränenburg. 2001. Characterization of p190RhoGEF, a RhoA-specific guanine nucleotide exchange factor that interacts with microtubules. *J. Biol. Chem.* 276:4948–4956. <http://dx.doi.org/10.1074/jbc.M003839200>
- Vicente-Manzanares, M., X. Ma, R.S. Adelstein, and A.R. Horwitz. 2009. Non-muscle myosin II takes centre stage in cell adhesion and migration. *Nat. Rev. Mol. Cell Biol.* 10:778–790. <http://dx.doi.org/10.1038/nrm2786>
- Vitorino, P., and T. Meyer. 2008. Modular control of endothelial sheet migration. *Genes Dev.* 22:3268–3281. <http://dx.doi.org/10.1101/gad.1725808>
- Wallace, S.W., A. Magalhaes, and A. Hall. 2011. The Rho target PRK2 regulates apical junction formation in human bronchial epithelial cells. *Mol. Cell. Biol.* 31:81–91. <http://dx.doi.org/10.1128/MCB.01001-10>
- Wang, X., L. He, Y.I. Wu, K.M. Hahn, and D.J. Montell. 2010. Light-mediated activation reveals a key role for Rac in collective guidance of cell movement in vivo. *Nat. Cell Biol.* 12:591–597. <http://dx.doi.org/10.1038/ncb2061>
- Weber, G.F., M.A. Bjerke, and D.W. DeSimone. 2012. A mechanoresponsive cadherin-keratin complex directs polarized protrusive behavior and collective cell migration. *Dev. Cell.* 22:104–115. <http://dx.doi.org/10.1016/j.devcel.2011.10.013>
- Xu, X., D. Jin, J. Durgan, and A. Hall. 2013. LKB1 controls human bronchial epithelial morphogenesis through p114RhoGEF-dependent RhoA activation. *Mol. Cell. Biol.* 33:2671–2682. <http://dx.doi.org/10.1128/MCB.00154-13>
- Zaritsky, A., S. Natan, E. Ben-Jacob, and I. Tsarfaty. 2012. Emergence of HGF/SF-induced coordinated cellular motility. *PLoS One.* 7:e44671. <http://doi.org/10.1371/journal.pone.0044671>
- Zaritsky, A., D. Kaplan, I. Hecht, S. Natan, L. Wolf, N.S. Gov, E. Ben-Jacob, and I. Tsarfaty. 2014. Propagating waves of directionality and coordination orchestrate collective cell migration. *PLOS Comput. Biol.* 10:e1003747. <http://dx.doi.org/10.1371/journal.pcbi.1003747>
- Zaritsky, A., E.S. Welf, Y.-Y. Tseng, M.A. Rabadian, X. Serra-Picamal, X. Trepat, and G. Danuser. 2015. Seeds of locally aligned motion and stress coordinate a collective cell migration. *Biophys. J.* 109:2492–2500. <http://dx.doi.org/10.1016/j.bpj.2015.11.001>

Supplemental material

JCB

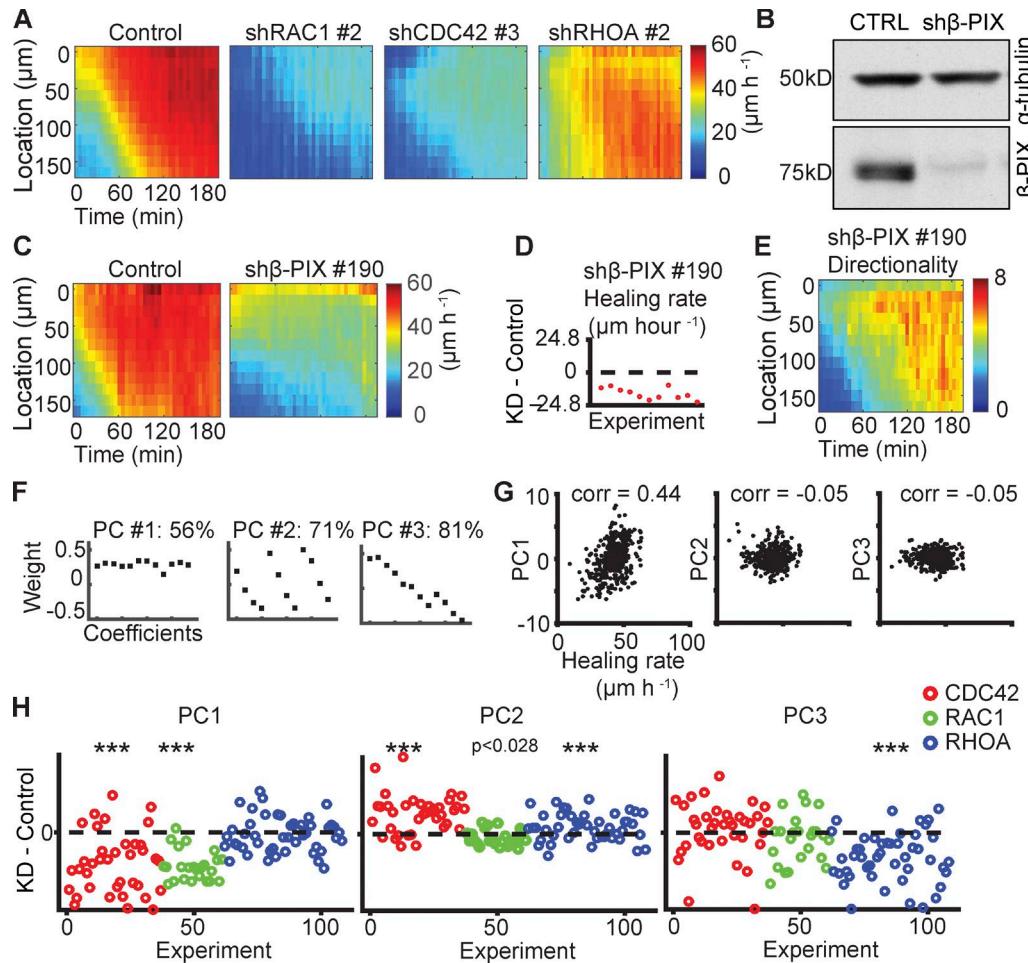
Zaritsky et al., <https://doi.org/10.1083/jcb.201609095>

Figure S1. Alterations of collective motility induced by depletion of Rho GTPases or the GEF β -PIX. (A) Validation of Rho GTPase knockdown effect on speed kymograph by the second hairpin. Kymographs are averages of three or four locations in a well. (B) Western blot showing effective depletion of β -PIX. (C) Speed kymographs. (D) Wound-healing rate. Each point was calculated as the difference between the migration rate in one location and the mean of the same day's control experiment. $N = 3$ d and $n = 11$ locations, $P = 0.001$ via Wilcoxon signed rank test. (E) Directionality kymograph, averaged over four locations in one well. (F) Coefficients of the first three principal components (PCs) in directionality variation, calculated from 402 control experiments. Together, three first PCs account for 81% of directionality variation. (G) Association of PCs in directionality with wound-healing rate. (H) Effect of Cdc42, Rac1, and RhoA depletion on PCs in directionality. For each experiment the difference between knockdown (KD) and control in the PC space is shown. Statistics via Wilcoxon signed rank test: ***, $P \leq 0.0001$.

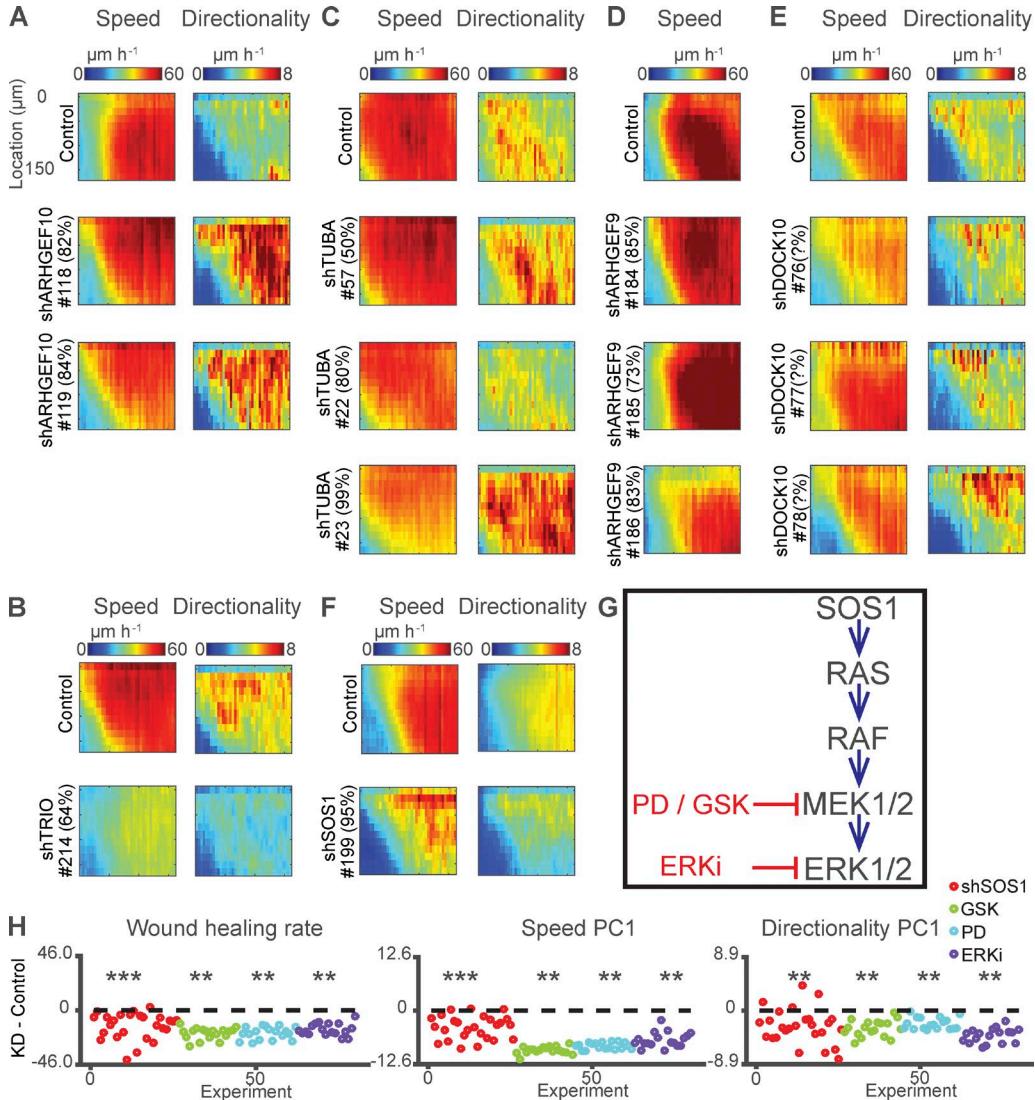


Figure S2. Screen hits that were not followed up and effects of SOS1-RAS pathway inhibition on collective cell migration. (A–E) Screen hits that were not followed up. (A) ARHGEF10 (also named Gef10), a RHOA-GEF (Aoki et al., 2009), showed increased directionality for two distinct validated hairpins. It was not followed up because it was not a paralog of the other four RhoA-GEFs hits. (B and C) TRIO (B) is a dual GEF activating RhoA and Rac1 (Bellanger et al., 1998), and TUBA (C) is a CDC42-specific GEF (Salazar et al., 2003). Both Trio and Tuba knockdowns showed reduced speed, and down-regulation of Trio reduced directionality while Tuba had an inconsistent directionality phenotype. Tuba showed a dose dependency alteration: reduced speed for a more effective depletion. Both TRIO and TUBA were not followed up, because we focused on intercellular communication alterations. (D) ARHGEF9 is a CDC42-specific GEF (Reid et al., 1999). Hairpin 184 was identified as a hit in speed PC3 (slower propagation than the control). The reason was that the control experiment on that day was an outlier (rapid front-to-back propagation in speed; PC3). Thus, the PC3 in the knockdown experiment was erroneously identified as a hit. (E) DOCK10 is a dual GEF for CDC42 and RAC1 (Ruiz-Lafuente et al., 2015). Hairpin 78 came as a hit in directionality PC2. None of the three primers for DOCK10 worked with qRT-PCR to assess KD efficiency; thus, they were labeled “unknown.” Together with the inability to observe a similar phenotype in the two other hairpins, and the fact that temporal gradient is not an intercellular phenotype, we decided not to follow up on DOCK10. (F–H) SOS1-RAS pathway regulates collective cell migration. (F) Speed kymographs. (G) SOS-RAS pathway. Inhibition of MEK with GSK or PD and of ERK by ERKi. (H) SOS1-RAS follow-up experiments. Each data point is the subtraction of the mean measured phenotype of daily controls from the certain location within a well (KD – control). Left to right: wound-healing rate, PC1 speed (encoding magnitude), and PC1 directionality. N , number of wells; n , number of locations ($N = 3\text{--}6$ and $n = 18\text{--}26$), one well per daily experiment, number of daily experiments (replicates) ≥ 3 . SOS1: $N = 6$ and $n = 26$; GSK, PD, and ERKi: $N = 3$ and $n = 18$. Statistics via Wilcoxon signed rank test: **, $P \leq 0.001$; ***, $P \leq 0.001$. Note that with more replicates, we were able to reveal a systematic reduction of speed and directionality that were not identified by the primary screen.

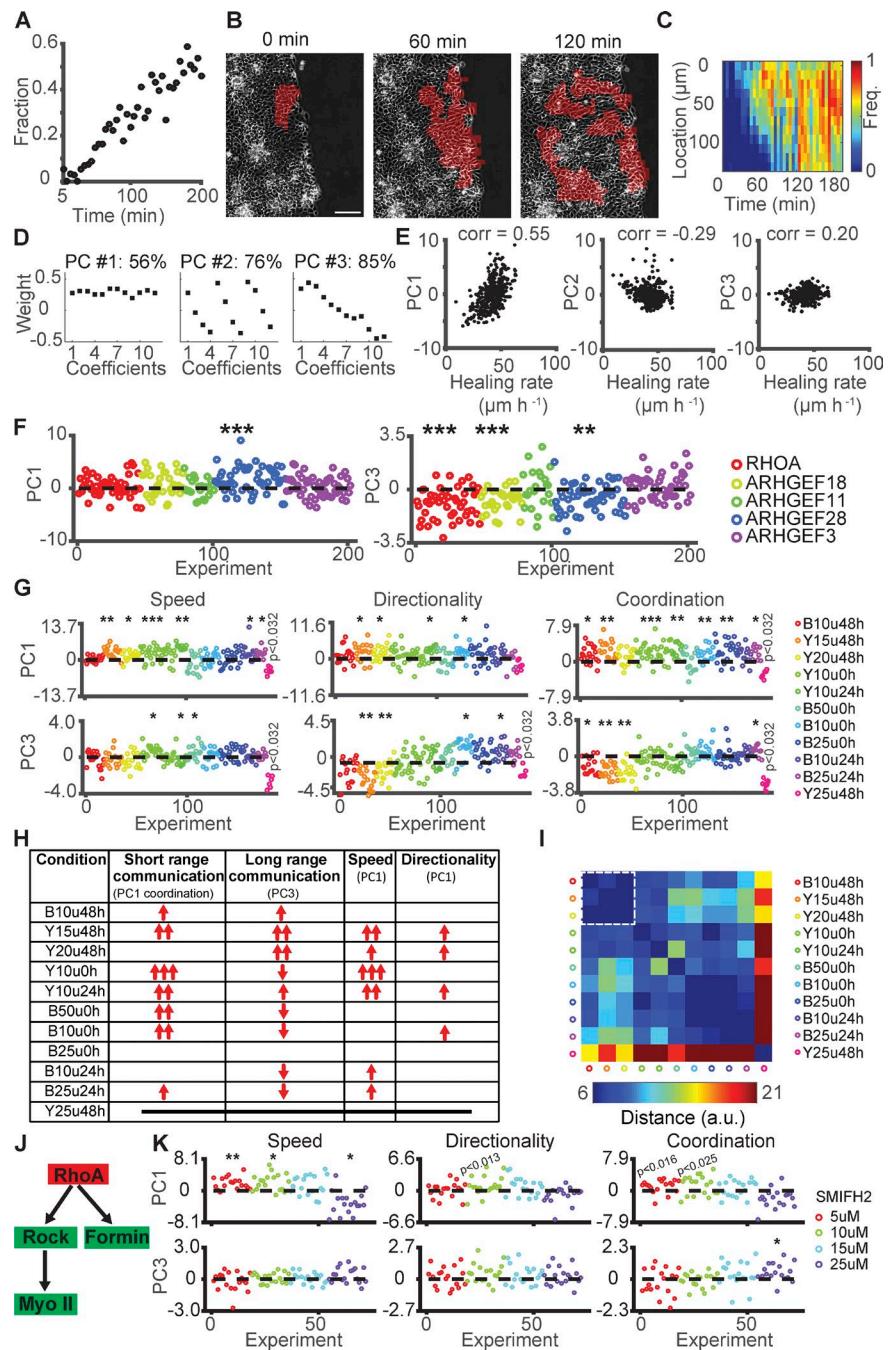
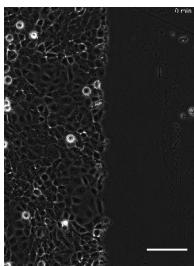


Figure S3. Effects of RhoA GEF depletion on short-range communication as quantified by the coordination parameter and effects of perturbation of actomyosin contractility on intercellular communication. (A) Fraction of cells migrating in coordinated clusters (coordination) over time. (B) Representative formation of clusters in a control experiment. Bar, 100 μm. (C) Kymograph of coordination. (D) Coefficients of the first three principal components (PCs) of coordination extracted from $n = 402$ control experiments. (E) Association of PCs with wound-healing rate. (F) Difference in PC1 and PC3 between GEF knockdown and control. Each data point is the subtraction of the mean PC value of the daily controls from the PC value extracted in one well location. Same experiments as in Fig. 4 G. (G) Shifts in spatiotemporal dynamics of speed, directionality, and coordination upon inhibition of myosin II by preincubation of cells with blebbistatin or Y27632 (denoted B or Y, respectively) at low doses (10, 15, 25, or 50 μM, denoted u) for 0, 24, or 48 h (denoted Oh, 24h or 48h) before wound-healing experiment. Each point represents the difference in PC of the indicated variable between GEF knockdown in one well location and the mean of the daily control; N , number of wells; n , number of locations ($N = 1\text{--}5$ and $n = 6\text{--}35$), one control well per daily experiment. B10u48h: $N = 3$ and $n = 18$; Y15u48h: $N = 3$ and $n = 18$; Y20u48h: $N = 3$ and $n = 18$; Y10u0h: $N = 5$ and $n = 35$; Y10u24h: $N = 2$ and $n = 11$; B50u0h: $N = 3$ and $n = 15$; B10u0h: $N = 3$ and $n = 17$; B25u0h: $N = 4$ and $n = 23$; B10u24h: $N = 2$ and $n = 13$; B25u24h: $N = 2$ and $n = 10$; Y20u48h: $N = 1$ and $n = 6$. Statistics via Wilcoxon signed rank test: *, $P \leq 0.01$; **, $P \leq 0.001$; ***, $P \leq 0.0001$. (H) Summary of alterations induced by contractility inhibition. Number of arrows corresponds to p-values in F and G. (I) Pairwise distance matrix measuring the similarity of perturbation effects (Materials and methods). Dashed white line marks the 48-h treatment functional cluster. (J) RHOA is upstream of the ROCK-myosin II and mDia pathways. (K) Increased speed and subtle increase in directionality and coordination upon pan-inhibition of formin-family members by acute treatment with low dosages of SMIFH2 (5, 10, 15 or 25 μM; Rizvi et al., 2009). Each point represents the difference in PC of the indicated variable between GEF knockdown in one well location and the mean of the daily control; $N = 3$ and $n = 18$ for each experimental conditions, one control well per daily experiment. Statistics via Wilcoxon signed rank test.



Video 1. Phase-contrast live imaging of the wound-healing response of 16HBE cells. The video was recorded at 5 min per frame.

Tables S1-S7 are available as Excel files.

Table S1 lists primers used to quantify GEF expression for 80 of 81 GEFs (some had two different primers).

Table S2 lists shRNA hairpin sequences and knockdown efficiency. Knockdown efficiency was assessed via Western blots and qRT-PCR. N/A indicates that the primer could not be validated by qRT-PCR because of failures in the production of efficient primers (11 GEFs). Five GEFs were not expressed in 16HBE cells. Red indicates that (16) GEFs were all hairpins with knockdown efficiency less than 50%.

Table S3 is a summary of Table S2. List of GEFs with zero, one, two, or three hairpins with knockdown efficiency >50%.

Table S4 provides details for all screen experiments. Details of every time-lapse experiment recorded in the screen. Every entry describes a location in a well. Each day contains a control well and multiple wells with a GEF knockdown. The information for every video includes the molecular perturbation in the format GENE_HAIRPIN, file name of raw data, file format (.zvi or .tif), physical pixel size in micrometers, measured knockdown efficiency (%), and experiment date.

Table S5 provides details for all screen and follow-up experiments. Same as Table S4, including all experiments performed in this study.

Table S6 lists screen z-scores. For every combination of GEF, hairpin, and experiment day, the table reports the z-score for the wound-healing rate and PCs 1-3 for speed and directionality. Z-scores were calculated as described in Materials and methods.

Table S7 lists coordination z-scores for every combination of GEF, hairpin, and experiment day.

References

- Aoki, T., S. Ueda, T. Kataoka, and T. Satoh. 2009. Regulation of mitotic spindle formation by the RhoA guanine nucleotide exchange factor ARHGEF10. *BMC Cell Biol.* 10:56. <http://dx.doi.org/10.1186/1471-2121-10-56>
- Bellanger, J.-M., J.-B. Lazaro, S. Diriong, A. Fernandez, N. Lamb, and A. Debant. 1998. The two guanine nucleotide exchange factor domains of Trio link the Rac1 and the RhoA pathways in vivo. *Oncogene*. 16:147–152. <http://dx.doi.org/10.1038/sj.onc.1201532>
- Reid, T., A. Bathoorn, M.R. Ahmadian, and J.G. Collard. 1999. Identification and characterization of hPEM-2, a guanine nucleotide exchange factor specific for Cdc42. *J. Biol. Chem.* 274:33587–33593. <http://dx.doi.org/10.1074/jbc.274.47.33587>
- Rizvi, S.A., E.M. Neidt, J. Cui, Z. Feiger, C.T. Skau, M.L. Gardel, S.A. Kozmin, and D.R. Kovar. 2009. Identification and characterization of a small molecule inhibitor of formin-mediated actin assembly. *Chem. Biol.* 16:1158–1168. <http://dx.doi.org/10.1016/j.chembiol.2009.10.006>
- Ruiz-Lafuente, N., M.-J. Alcaraz-García, A.-M. García-Serna, S. Sebastián-Ruiz, M.-R. Moya-Quiles, A.-M. García-Alonso, and A. Parrado. 2015. Dock10, a Cdc42 and Rac1 GEF, induces loss of elongation, filopodia, and ruffles in cervical cancer epithelial HeLa cells. *Biol. Open*. 4:627–635. <http://dx.doi.org/10.1242/bio.20149050>
- Salazar, M.A., A.V. Kwiatkowski, L. Pellegrini, G. Cestra, M.H. Butler, K.L. Rossman, D.M. Serna, J. Sondek, F.B. Gertler, and P. De Camilli. 2003. Tuba, a novel protein containing bin/amphiphysin/Rvs and Dbl homology domains, links dynamin to regulation of the actin cytoskeleton. *J. Biol. Chem.* 278:49031–49043. <http://dx.doi.org/10.1074/jbc.M308104200>