Myosin VI regulates ciliogenesis by promoting the turnover of the centrosomal/satellite protein OFD1. EMBO Rep. *23*, e54160. https://doi.org/10.15252/embr.202154160.

59. Innocenti, M., Tenca, P., Frittoli, E., Faretta, M., Tocchetti, A., Di Fiore, P.P., and Scita, G. (2002). Mechanisms through which Sos-1 coordinates the activation of Ras and Rac. J. Cell Biol. *156*, 125–136. https://doi.org/10.1083/jcb.200108035.

60. Thielicke, W., and Sonntag, R. (2021). Particle Image Velocimetry for MATLAB: Accuracy and enhanced algorithms in PIVlab. J. Open Res. Software *9*, 12. https://doi.org/10.5334/jors.334.

61. Sun, W., Duan, T., Ye, P., Chen, K., Zhang, G., Lai, M., and Zhang, H. (2018). TSVdb: a web-tool for TCGA splicing variants analysis. BMC Genom. *19*, 405. https://doi.org/10.1186/s12864-018-4775-x.

62. Nakamuta, S., Yang, Y.-T., Wang, C.-L., Gallo, N.B., Yu, J.-R., Tai, Y., and Van Aelst, L. (2017). Dual role for DOCK7 in tangential migration of interneuron precursors in the postnatal forebrain. J. Cell Biol. *216*, 4313–4330. https://doi.org/10.1083/jcb.201704157.

63. Palamidessi, A., Frittoli, E., Garré, M., Faretta, M., Mione, M., Testa, I., Diaspro, A., Lanzetti, L., Scita, G., and Di Fiore, P.P. (2008). Endocytic Trafficking of Rac Is Required for the Spatial Restriction of Signaling in Cell Migration. Cell *134*, 135–147. https://doi.org/10.1016/j.cell.2008.05.034.

64. Nastały, P., Purushothaman, D., Marchesi, S., Poli, A., Lendenmann, T., Kidiyoor, G.R., Beznoussenko, G.V., Lavore, S., Romano, O.M., Poulikakos, D., et al. (2020). Role of the nuclear membrane protein Emerin in front-rear polarity of the nucleus. Nat. Commun. *11*, 2122. https://doi.org/10.1038/s41467-020-15910-9.

65. Villa, S., Palamidessi, A., Frittoli, E., Scita, G., Cerbino, R., and Giavazzi, F. (2022). Non-invasive measurement of nuclear relative stiffness from quantitative analysis of microscopy data. Eur. Phys. J. E Soft Matter *45*, 50. https://doi.org/10.1140/epje/s10189-022-00189-z.

66. Crocker, J.C., and Grier, D.G. (1996). Methods of Digital Video Microscopy for Colloidal Studies. J. Colloid Interface Sci. *179*, 298–310. https://doi.org/10.1006/jcis.1996.0217.

67. Bolte, S., and Cordelières, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. J. Microsc. *224*, 213–232. https://doi.org/10.1111/j.1365-2818.2006.01706.x.

68. Manders, E.M.M., Verbeek, F.J., and Aten, J.A. (1993). Measurement of co-localization of objects in dual-colour confocal images. J. Microsc. *169*, 375–382. https://doi.org/10.1111/j.1365-2818.1993.tb03313.x.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-Desmoplakin	provided by Kathleen Green	N/A
anti-RAC1	Becton Dickinson	Cat#610651
anti-MYO6	Wollscheid et al.[31]	Eurogentec-1296
anti-GAPDH	Santa Cruz	Cat#32233
anti-DOCK7	Santa Cruz	Cat#398888
anti-RAB5A	Santa Cruz	Cat#166600
anti-CDC42	Santa Cruz	Cat#390210
anti-GFP	Sigma-Aldrich	Cat#G1544
anti-His	Cell Signaling	Cat#2366
Chemicals, peptides, and recombinant proteins		
Human EGF	Vinci Biochem	BPS-90201-1
GST-beads	GE Healthcare	GE17-0756-01
GFP-TRAP	Cromotek	RRID: AB_2631357
BODIPY^{TM} FL-GTP	Thermo Fisher Scientific	Cat# G12411
Carbachol	Calbiochem, Sigma	CAS 51-83-2
Deposited data		
Raw and analyzed data	This study, Mendeley Data	https://doi.org/10.17632/zsxcs45pt3.1
TCGA-BRCA	https://portal.gdc.cancer.gov/projects/TCGA-BRCA	phs000178
Experimental models: Cell lines		
MCF10.DCIS.com RAB5A	Malinverno et al.[14]	N/A
MCF10.DCIS.com RAB5A GFP-LifeAct mCherry-H2B	Malinverno et al.[14]	N/A
MCF10.DCIS.com RAB5A mCherry-H2B	Malinverno et al.[14]	N/A
MCF10.DCIS.com RAB5A GFP-DOCK7	This study	N/A
MCF10.DCIS.com RAB5A GFP-LifeAct mCherry-H2B RAC1 WT	This study	N/A
MCF10.DCIS.com RAB5A GFP-LifeAct mCherry-H2B RAC1 P62S	This study	N/A
MCF10.DCIS.com RAB5A RAC1-2G	This study	N/A
MCF10.DCIS.com RAB5A CDC42-2G	This study	N/A
HEK293T	ICLC	N/A
Oligonucleotides		
siRNA targeting sequence: MYO6 #1: 5′-GAGGCUGCACU AGAUACUUUGCUAA-3′	Thermo Fisher Scientific	NM_004999_stealth_1106
siRNA targeting sequence: MYO6 #2: 5′-GAGCCTTTGCCA TGGTACTTAGGTA-3′	Thermo Fisher Scientific	NM_004999_stealth_4
siRNA targeting sequence: DOCK7 #1: 5′- UUUAAGGUCA UCUUGAUCAUCCUGG-3′	Thermo Fisher Scientific	Cat # HSS131697
siRNA targeting sequence: DOCK7 #2: 5′- AUUAGGGUAA GUAUUUGGUAGGCGG-3′	Thermo Fisher Scientific	Cat # HSS131695
MYO6 isoform detection by PCR For. Sequence: CCGAG CTCATCAGTGATGAGGC	Wollscheid et al.[31]	N/A
MYO6 isoform detection by PCR Rev. Sequence: CCAAGC ATGATACACTTTTAGTCTCC	Wollscheid et al.[31]	N/A
Primer: DOCK7-FL. Forward: GCGGCCCCGAACTAGTGCCACCATGGTGAGCAAGG	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Primer: DHR2-LobeBC (aa1795-2072) of DOCK7. EcoRI-Forward: AAAGAATTCCGGATGTTTGGCACCTATTTTC	This study	N/A
Primer: DHR2-LobeBC (aa1795-2072) of DOCK7. XhoI-Reverse: AAACTCGAGTTAAGGGATCTTTCTGTTGATC	This study	N/A
Primer: DHR2-LobeAC (aa1597-2072) of DOCK7. EcoRI-Forward: AAAGAATTCGATCTGGTTTTCAATCTCC	This study	N/A
Primer: DHR2-LobeAC (aa1597-2072) of DOCK7. XhoI-Reverse: AAAGAATTCAAGGGTTACCAGACCTCTCC	This study	N/A
Primer: RAC1 (aa1-177). EcoRI-Forward: AAAGAATTCA TGCAGGCCATCAAGTG	This study	N/A
Primer: RAC1 (aa1-177). XhoI-Reverse: AAACTCGAGTT AGAGGACTGCTCG	This study	N/A
Recombinant DNA		
pLenti-RAC1-2G FRET Biosensor	Fritz et al.[37]	Addgene Plasmid #66111
pLenti-Cdc42-2G FRET Biosensor	Martin et al.[40]	Addgene Plasmid #68813
pGEX-GST-CRIB	Innocenti et al.[59]	N/A
Software and algorithms		
ImageJ	Schneider et al.	https://imagej.nih.gov/ij/
Fiji	N/A	https://imagej.net/software/fiji/downloads
PRISM	N/A	https://www.graphpad.com/features
MATLAB software PIVlab	Thielicke et al.[60]	https://www.mathworks.com/
TSVdb	Sun et al.[61]	http://www.tsvdb.com
Biorender	N/A	Biorender.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Simona Polo (simona.polo@ifom.eu).

Materials availability

All unique reagents generated in this study are available from the lead contact upon request and with a completed Materials Transfer Agreement.

Data and code availability

- This paper analyzed existing, publicly available data. The accession numbers for the datasets are listed in the key resources table (https://portal.gdc.cancer.gov/projects/TCGA-BRCA).
- This paper does not report original code.
- Raw data from main figures are deposited on Mendeley at https://doi.org/10.17632/zsxcs45pt3.1.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Constructs and reagents

GTP-CRIB plasmid[59] and GST-myosin VI (human isoform2, NP_001287828.1) constructs were previously described.[31] pLenti-RAC1-2G FRET Biosensor and pLenti-Cdc42-2G FRET Biosensor were obtained from Addgene (#66111 and #68813, respectively). DHR2 domain of DOCK2, DOCK6 and DOCK9 GEF proteins were amplified by PCR (primers listed in the "key resources table") using cDNA from Caco2 cells, cloned into the expression vector pEGFP-C1 and sequence-verified. pEGFP-DOCK7 was amplified by PCR from pCA-FLAG-DOCK7-FL[62] and cloned into the pEGFP-C1 expression vector. pSLIK-EGFP-DOCK7 and pSLIK-RAC1 wild-type (WT) and P29S constructs were generated starting from previous constructs (pCDNA3-RAC1[63] and pEGFP-DOCK7) with Gateway LR Clonase II Enzyme mix (Thermo Fisher Scientific) by subcloning their respective PCR products into a pENTR vector, followed by recombination into the pSLIK-HYGRO empty vector.

All the other truncated constructs were engineered by site-directed mutagenesis or recombinant PCR and sequence-verified. Briefly, the DHR2 domain of DOCK7 (residues 1597–2072) was amplified and cloned into the pET28a, pET43 and pEGFP-C1 expression vectors. RAC1 (residues 1–177) was amplified and cloned into the pGEX-6P-1 expression vector. DHR2-ΔLobeA of DOCK7 (residues 1795–2072) was amplified and cloned into the pGEX-6P-1 and pEGFP-C1 expression vectors.

Cell lines and transfection procedures

The DCIS-RAB5A cell line and derivatives stably expressing GFP-LifeAct or mCherry-H2B were previously described.[17] Cells were grown at 37°C in humidified atmosphere with 5% CO_2 in Dulbecco's Modified Eagle Medium (DMEM): Nutrient Mixture F-12 (DMEM/F12) medium (Invitrogen) supplemented with 5% horse serum, 0.5 mg/mL hydrocortisone, 10 μg/mL insulin, and 20 ng/mL EGF. Phoenix-AMPHO cells (American Type Culture Collection, CRL-3213) were used as the packaging cell line for the generation of retroviral particles and cultured as recommended by the supplier. HEK293T cells were obtained from the BBCF-Biological Bank and Cell factory, INT, Milan, grown in DMEM supplemented with 10% fetal bovine serum and 2 mM L-glutamine and used as the packaging line for lentiviral vectors.

DCIS-RAB5A cells and derivatives were infected with pSLIK-EV (empty vector, CTR), pSLIK-EGFP-DOCK7-FL or pSLIK-RAC1 WT and P29S and selected with hygromycin to obtain stable inducible cell lines. In the case of DCIS-RAB5A expressing pSLIK-EGFP-DOCK7-FL, cells were FACS sorted with Beckman Coulter MoFlo Astrios to obtain a homogeneous population with low expression of GFP-DOCK7 in order to determine its localization. In these cell lines, doxycycline induction promotes the expression of RAB5A and DOCK7 or RAB5A and RAC1 WT/P29S. All cell lines were authenticated by cell fingerprinting and tested for mycoplasma contamination.

FRET-based analysis of RAC1 and CDC42 activation was performed infecting DCIS-RAB5A cells with pLent-RAC1-2G[37] or pLenti-CDC42-2G[40] second generation FRET biosensors, followed by selection with puromycin to obtain stable inducible cell lines.

Transient knock-down of myosin VI or DOCK7 was performed using Stealth siRNA oligonucleotides (listed in the "key resources table") from Thermo Fischer Scientific (Waltham, MA, USA). Cells were transfected twice using RNAiMax (Invitrogen), first in suspension and the following day in adhesion. Based on immunoblot analyses, cells were considered myosin VI or DOCK7-depleted three days after the first transfection. Two siRNA oligonucleotides/gene were used with comparable results.

Expression of myosin VI isoforms in DCIS-RAB5A cells was assessed by RT-PCR. RNA was isolated from cells grown at different confluencies using Maxwell RSC Instrument and Maxwell RSC simplyRNA Cells Kit. Retro-transcription was performed with QuantiTect Reverse Transcription Kit (Qiagen). The cDNA obtained was amplified by PCR using primers flanking the spliced regions (listed in the "key resources table") as previously described.[31]

Antibodies

The following antibodies were used at the indicated dilutions: anti-His (mouse, 1:1000, Cell Signaling), anti-RAC1 (mouse, 1:1000, BD), anti-E-cad (mouse, 1:200, BD-610181), anti-Desmoplakin (rabbit, 1:1000, NW6; kindly provided by Kathleen Green and Lisa Godsel), anti-myosin VI (rabbit, homemade,[31] 1:5000, Eurogentec-1296), anti-GAPDH (mouse, 1:2000, Santa Cruz-32233), anti-DOCK7 (mouse, 1:1000, Santa Cruz-398888), anti-RAB5A (mouse, 1:1000, Santa Cruz-166600), anti-CDC42 (rabbit, 1:1000, Santa Cruz-2462), and anti-GFP (rabbit, 1:5000, Sigma-G1544).

Single cell migration assay

Random single cell migration experiment was performed as previously described.[14] Briefly, DCIS-RAB5A mCherry-H2B cells were transfected twice with siRNA oligos and seeded in sparse cell growth conditions in six-well plates (5×10^4 cells/well) in complete medium. RAB5A expression was induced 16 h before the experiment was initiated by adding fresh complete media supplemented with 2.5 μg/mL doxycycline to the cells. Cell migration was monitored by time-lapse microscopy. At the time of recording, fresh media containing EGF was added to the cells. The assay was performed using an environmental microscope incubator set to 37°C and 5% CO2 perfusion. An Olympus ScanR inverted microscope with 10× objective was used to acquire images every 5 min over a 24-h period. Tracking of cell nuclei and motility analysis were performed as described below.

For confined migration experiment, we employed fibronectin-coated micro-patterned lines of 10 μm diameter created through photolithography.[64] Briefly, glass coverslips were activated with plasma cleaner (Harrick Plasma), followed by coating with PLL-g-PEG (Surface Solutions GmbH, 0.1 mg/mL). After washing with phosphate-buffered saline (PBS), the surface was illuminated with UV light (UVO Cleaner, Jelight) using chromium photo-masks (JD-Photodata). The coverslips were then incubated with fibronectin (25 μg/mL), and 10^4 DCIS-RAB5A mCherry-H2B transfected cells were seeded over-night with 2.5 μg/mL doxycycline prior to analysis. Images were acquired every 5 min for 24 h using a humidity- and temperature-controlled inverted wide-field ScanR microscope. Tracking of cell nuclei and motility analysis were performed as previously described using a developed C++ software with the OpenCV [http://opencv.willowgarage.com/wiki/] and the GSL [http://www.gnu.org/software/gsl/] libraries. The migration analysis was performed by the C++ software coupled with R [www.R-project.org].

Wound healing assays

Assays were performed as previously described.[14] Briefly, cells transfected in suspension were seeded at confluency in six-well plates (1.5×10^6 cells/well) in complete medium and transfected again the following day. Three days after seeding, a uniform

monolayer is formed. RAB5A, DOCK7 and RAC1 WT/P29S expression was induced 16 h before the experiment was initiated by adding fresh complete media supplemented with 2.5 μg/mL doxycycline to the cells. The cell monolayer was scratched with a pipette tip and carefully washed with PBS to remove floating cells and create a cell-free wound area. The closure of the wound was monitored by time-lapse microscopy. At the time of recording, fresh media containing EGF was added to the cells. The assay was performed using an environmental microscope incubator set to 37°C and 5% CO_2 perfusion. An Olympus ScanR inverted microscope with 10× objective was used to acquire images every 5 min over a 24-h period. The percentage of area covered by cells (area coverage %) over time and wound-front speed were calculated using a custom Fiji and MATLAB code. The area covered over time was fitted with a straight line whose slope was used to estimate the velocity of wound closure.

Measurement of the cellular velocities and trajectories

To measure velocity and trajectories of the cells at various distance from the edge of the wound we used Fiji Trackmate plugin [10]. The obtained tracks were separated in three different area: FRONT (<50μm from the wound edge), MIDDLE (50μm < wound edge <150μm) and BACK (150μm < wound edge <300μm) according to the cell distance from the wound edge. The FRONT area corresponds to $n \leq 3$ cells in a row.

The distance between each nuclei centroid (identified by Trackmate analysis) and the wound edge (identified by the previously described wound healing analysis) was calculated using the MATLAB "bwdist" function. For each area (FRONT, MIDDLE, BACK) and for each track the directional change rate (https://imagej.net/plugins/trackmate/algorithms#mean-directional-change) was measured by Fiji Trackmate plugin as the average angle difference between two subsequent displacements. Directionality of the cells belonging to the different area are expressed as the inverse of the directional change rate parameter obtained.

Kymograph analysis of cell protrusions at the wound edge

For cell protrusion analysis at the wound edge, the wound healing assay was performed using Culture Inserts (Ibidi) to avoid debrides affecting the quality of the kymograph analysis. Inserts were placed in a 12-well plate and DCIS-RAB5A transfected cells were plated in each chamber (6×10^4 cells/chamber). A cell-free wound area was created by removal of the insert. Cell migration was monitored by a Leica widefield Thunder imager equipped with a Leica sCMOS DFC9000GT, using a Leica HC PL Fluotar 20× objective, NA 0.5. Images were acquired every 30 s for 1 h (100 ms exposure time). Images from 10 positions/condition were recorded every 30 s over a 1-h period. To measure the dynamic of protrusive structures, each visible lamellipodia was analyzed by tracing a single pixel wide line orthogonally to the edge of the wound. The resulting kymograph was obtained and analyzed following the plugin of the ImageJ Kymograph macro (available at https://www.embl.de/eamnet/html/kymograph.html). Briefly, a segmented line was traced following the edges of the lamellipodia in the kymographs, and every protrusion or retraction in the video were averaged to obtain a single value for the plot. Persistence was calculated considering the total amount of frames in which the protrusion is visible until retraction begins.

In order to extract the lamellipodia region area from phase contrast time lapse images acquired as described above, we used a fully convolutional neural network.[36] The network was trained with 65 images of leading-edge cell protrusions at the wound.

Cell sheet streaming and kinetic parameter measurements

DCIS-RAB5A and derivative cell lines were transfected and seeded as described for the wound healing experiment. RAB5A, DOCK7 and RAC1 WT/P29S expression was induced 16 h before the experiment was initiated with 2.5 μg/mL doxycycline. The assay was performed using an environmental microscope incubator set to 37°C and 5% CO2 perfusion. An Olympus ScanR inverted microscope with 10× objective was used to acquire images every 5 min over a 24-h period.

Quantification of monolayer dynamics in the streaming assays was performed using both Particle Imaging Velocimetry (PIV) and Particle Tracking (PT). PIV analysis of phase-contrast image sequences was performed using the MATLAB software PIVlab.[60] In all cases, we adopted a final interrogation area of $20.8x20.8\ \mu m^2$, close to the typical cell projected area. Spurious contributions to the velocity field corresponding to instantaneous global translations between consecutive frames (due to stage positioning errors) were removed through the smoothing procedure described in.[65]

From the resulting velocity field $\boldsymbol{v}(\boldsymbol{x},t)$, where $\boldsymbol{x}$ is the position in the monolayer plane, the root mean squared velocity v_{RMS} and the polar order parameter ψ were computed as:

$$v_{RMS} = < \sqrt{\langle |\boldsymbol{v}|^2 \rangle_{\boldsymbol{x}}} >_{j,t}$$

$$\psi = \left\langle \frac{|\langle \boldsymbol{v} \rangle_{\boldsymbol{x}}|^2}{\langle |\boldsymbol{v}|^2 \rangle_{\boldsymbol{x}}} \right\rangle_{j,t}$$

Where $< \ldots >_{\boldsymbol{x}}$ denotes the space average while $< \ldots >_{j,t}$ denotes the average over different field of views (FOVs) j and over different time points t. For each experiment, at least 5 different FOV were considered, while the time average was made over a time window from 4 to 20 h after time-lapse recording starts. This time window roughly corresponded to the peak of RAB5A-induced motility.[14]

The velocity spatial correlation function was evaluated from the PIV velocity field as

$$C_{vv}(r) = \left\langle \left\langle \frac{\langle \boldsymbol{v}(\boldsymbol{x},t)\cdot\boldsymbol{v}(\boldsymbol{x}+\boldsymbol{r},t)\rangle_{\boldsymbol{x}}}{\left\langle |\boldsymbol{v}(\boldsymbol{x},t)|^2 \right\rangle_{\boldsymbol{x}}} \right\rangle_{\theta} \right\rangle_{j,t}$$

Where $\langle \ldots \rangle_{\theta}$ denotes the azimuthal average over different orientations of the relative position vector $\boldsymbol{r}$ and $r \equiv |\boldsymbol{r}|$. The obtained velocity space correlation was fitted with a stretched exponential function $C_{vv}(r) = \exp\left[-\left(\frac{r}{l}\right)^{\beta} \right]$. An estimate for the correlation length L_{corr} was obtained by calculating the average decay length $L_{corr} \equiv \int_0^{\infty} dr \exp\left[-\left(\frac{r}{l}\right)^{\beta} \right] = \frac{1}{\beta}\Gamma\left(\frac{1}{\beta}\right) l$, where Γ is the gamma function.

In order to quantify the relative motion between neighboring cells, we performed particle tracking on fluorescent nuclei. The particle tracking algorithm is described in detail in.[65] Briefly, we performed for each frame I a seeded watershed transformation of the gradient image ∇I to segment each fluorescent nucleus k and compute its centers of mass $\boldsymbol{x}_k$, assumed to coincide with the geometrical center of mass of its projection on the plane. Single nuclei trajectories were then obtained by linking nuclei positions in subsequent frames using the available MATLAB implementation by D. Blair and E. Dufresne (http://physics.georgetown.edu/matlab) of the Grier Crocker tracking algorithm.[66] From nuclear trajectories, the instantaneous velocity $\boldsymbol{v}_{INS,k}$ was then evaluated as $\boldsymbol{v}_{INS,k}(t) = [\boldsymbol{x}_k(t+\Delta t) - \boldsymbol{x}_k(t)]/\Delta t$, where Δt is the time between two acquired frames. In order to reduce tracking noise, we computed a weighted moving average $\boldsymbol{v}_k(t)$ by convolving the instantaneous velocity with a Gaussian filter of width 10 min.

We calculated the root mean squared relative velocity Δv_{RMS} of two nuclei at distance r as

$$\Delta v_{RMS}(r) = \left\langle \sqrt{\frac{\int_{r-\delta r}^{r+\delta r} \sum_{k,k'} |\boldsymbol{v}_k - \boldsymbol{v}_{k'}|^2 \delta(r' - |\boldsymbol{x}_k - \boldsymbol{x}_{k'}|) dr'}{\int_{r-\delta r}^{r+\delta r} \sum_{k,k'} \delta(r' - |\boldsymbol{x}_k - \boldsymbol{x}_{k'}|) dr'}} \right\rangle_{j,t},$$

where k and k' run over all nuclei within the same FOV and the amplitude of the integration interval $[r-\delta r, r+\delta r]$ corresponds to 1.3 micrometer. To obtain an estimate of the relative motion of adjacent cells, we evaluated by linear interpolation Δv_{RMS} for $r = 14\ \mu m$, corresponding to the average distance between the center of mass of a nucleus and the ones of its first neighbors.

All of the kinetic parameters were evaluated separately for each independent experiment. Each data point in Figures 1E and 1F, in the central panel of Figure G, and in the left panel of Figure 1G corresponds to independent experiments.

Measurements of cryptic lamellipodia dynamics

Assays were performed as previously described.[14] Briefly, DCIS-RAB5A cells stably expressing EGFP-LifeAct were mixed in a 1:10 ratio with unlabeled DCIS-RAB5A cells and transfected and seeded as described for the wound healing experiment. Cell migration was monitored by time-lapse phase-contrast and fluorescence microscopy with an Olympus ScanR inverted microscope using a 20× objective and images were acquired every 90 s over a 6-h period. The quantification of cryptic lamellipodia protrusion velocity was performed using the ADAPT plug-in of Fiji. Cryptic lamellipodia directionality was measured as the angle Φ delimited by the direction of the single lamellipodium and the direction vector of the collective pack locomotion.

$0^{\circ} \leq \Phi \leq 45^{\circ}$ indicates that protrusion and collective migration have the same direction; $\Phi = 180^{\circ}$ indicates that protrusion and collective migration have opposite directions. The assay was repeated five times for each condition and at least 25 cells/condition were counted for each experiment.

Immunofluorescence (IF)

DCIS-RAB5A cells were grown on coverslips and fixed with 4% paraformaldehyde (PFA) for 10 min. Coverslips were incubated in PBS with 2% BSA for 1 h for blocking, followed by incubation with primary antibodies for 1 h at room temperature (RT) (overnight at 4°C in case of confluent cells) and secondary antibodies for 1 h at RT in PBS containing 1% BSA. Incubation with DAPI (Sigma-Aldrich, cat. D9542) for 10 min was performed to stain the nuclei. Coverslips were mounted on glass slides using Mowiol Mounting Medium (Calbiochem) and images were acquired using Leica TCS SP8 laser confocal scanner mounted on a Leica DMI 6000B inverted microscope equipped with HCX PL APO 63×/1.4 NA oil immersion objective.

For Figure 2B, quantification of the mean fluorescence signal of myosin VI across the cells in Z-stacks images was performed manually using Fiji. The assay was repeated four times and 15 cells/experiment were analyzed.

Colocalization analysis was carried out adapting the JaCOP FIJI4,5 plugin[67] and using a custom pipeline able to process multiple folders and multicolor images. Manders' coefficients[68] were calculated considering phalloidin or myosin VI signal as image A and

DOCK7 signal as image B. For myosin VI-DOCK7 colocalization (Figure 5C), the XZ resliced images of the z-stacks images were split into two parts to distinguish the contribution of the basal and the apical regions of the cells.

FRET based RAC1/CDC42 activation assay

DCIS-RAB5A cells stably expressing RAC1-2G[37] or CDC42-2G[40] FRET biosensors were mixed in a 1:10 ratio with unlabeled DCIS-RAB5A cells and transfected and seeded at confluency on coverslips. Three days after seeding, RAB5A expression is induced for 16 h and coverslips were fixed with 4% PFA for 10 min. Images were acquired with GE HealthCare Deltavision OMX system, equipped with 2 PCO Edge 5.5 sCMOS cameras, using a 60 ×1.42 NA Oil immersion objective.

A custom Fiji plugin was developed to calculate single cell FRET signal. Briefly, YPF channel was used to identify single cell edge allowing background removal (https://imagej.net/plugins/rolling-ball-background-subtraction) and Gaussian filter. The single cells were segmented using the ImageJ Li's threshold method (https://imagej.net/plugins/auto-threshold#li). Cell periphery (<1.5 μm from the edge) was used to quantify the FRET intensity signal. The ratio between the FRET channel and the CFP channel was calculated using a python script to obtain the distribution parameters.

Protein expression and purification

Recombinant proteins were expressed in E. coli BL21 (DE3) at 18°C for 16 h after induction with 0.5 mM IPTG at an OD600 of 0.6.

For GST-fusion proteins, cell pellets were resuspended in lysis buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 1 mM EDTA, 0.1% NP40, 5% glycerol, 0.1 mM PMSF, and 1:500 protease inhibitor cocktail (Calbiochem). Sonicated lysates were cleared by centrifugation at 16,000 rpm for 45 min at 4°C. Supernatants were incubated with 1 mL of glutathione Sepharose beads (Cytiva) per liter of bacterial culture for 4 h at 4°C. Beads were washed four times with lysis buffer followed by four washes with high salt buffer (20 mM Tris-HCl pH 8.0, 1 M NaCl, 1 mM EDTA, 1 mM DTT, and 5% glycerol) and finally equilibrated in cleavage buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM DTT, and 5% glycerol). PreScission protease was added at a 1:50 (w/w) ratio (protease:substrate) and incubated for 16 h at 4°C. Cleaved proteins were concentrated in Amicon Ultra Centrifugal Filters (MW cut-off 10 and 30 kDa, respectively) (Merck Millipore) and loaded onto a Superdex 75 10/300 GL column (Cytiva) equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 5% glycerol, and 1 mM DTT. Fractions containing purified proteins were collected and concentrated.

For HisMBP-fusion DHR2 (DOCK7), cell pellets were lysed in 50 mM Na-Phosphate buffer pH 7.5, 200 mM NaCl, 10 mM imidazole, and 5% glycerol. After sonication and clearance of the lysate by centrifugation, supernatants were incubated with 1 mL Ni-NTA agarose beads (Qiagen) per 1 L of bacterial culture. Beads were washed four times with lysis buffer followed by four washes with high salt/imidazole buffer (20 mM Na-Phosphate buffer pH 7.5, 1 M NaCl, 30 mM imidazole, and 5% glycerol). Proteins were then eluted from beads with buffer containing 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 300 mM imidazole, and 5% glycerol, and dialyzed over night at 4°C in the same buffer without imidazole. Dialyzed proteins were concentrated and purified by SEC as described for the GST-fusion proteins.

Co-immunoprecipitation and pull-down assays

For co-immunoprecipitation (co-IP) analysis, 1 mg of fresh lysates were incubated with specific antibodies for 2 h at 4°C. Cells were lysed in JS buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 1.5 mM MgCl2, 5 mM EGTA, 10% glycerol, and 1% Triton X-100) supplemented with 20 mM sodium pyrophosphate, pH 7.5, 250 mM sodium fluoride, 2 mM PMSF, 10 mM sodium orthovanadate, and protease inhibitors (Calbiochem), and lysates were cleared by centrifugation.

For pull-down experiments, 500 μg of transfected HEK293T cellular lysates were incubated with 1 μM of GST-fusion proteins immobilized onto GSH beads for 2 h at 4°C in JS buffer. After extensive washes with JS buffer, beads were re-suspended in Laemmli buffer and proteins were analyzed through sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE, 4–20% TGX precast gel, Bio-Rad). Detection was performed either by staining the gels with Coomassie or by immunoblotting using specific antibodies.

For the evaluation of direct binding, recombinant GST-fusion proteins (MYO6 tail, MyUb-CBD and RAC1, respectively) immobilized onto beads were incubated with purified HisMBP-tagged DHR2 domain of DOCK7 (15 μM final concentration) for 3 h at 4°C in low salt buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 5% glycerol, and 1 mM DTT). Beads were washed four times with the same buffer supplemented with 1% Triton X- and the samples were run on SDS-PAGE. Detection was performed by immunoblotting using anti-His antibody. Ponceau-stained membranes were used to show equal loading.

The GTP-CRIB assay was performed as described in.[63] Briefly, 500 μg of cell lysates from DCIS-RAB5A mock or myosin VI depleted cell monolayers were incubated with purified GST-CRIB (5 μM final concentration) for 1 h at 4°C. Beads were washed five times with lysis buffer and subjected to SDS-Page followed by immunoblotting with anti-RAC1 or CDC42 antibodies. Quantification was performed by normalizing the intensities of the bands to the total amount of RAC1 in the lysates. Data are reported as fold change with respect to RAC1-GTP levels in the corresponding mock sample for each experiment. Five independent experiments were performed.

AlphaFold2-multimer prediction

The computational resources of the High-Performance Computing Biowulf cluster of the NIH (http://hpc.nih.gov) was used to run AlphaFold2-Multimer. The top-ranking structure from 50 predicted structures was selected for visualization and display. Structures were analyzed and figures generated by using PyMol (PyMOL Molecular Graphics System, http://www.pymol.org).

GEF activity assay

Before starting the assay, all recombinant purified proteins were dialyzed in buffer containing 20 mM Tris-HCl pH 7.0, 150 mM NaCl and 10 mM MgCl2.

Recombinant RAC1 (10 μM final concentration) was pre-incubated with GDP (15 μM final concentration) for 30 min on ice in buffer containing 20 mM Tris-HCl pH 7.0, 150 mM NaCl, 10 mM MgCl2, and 0.2 mg/mL BSA. Subsequently, fluorescent boron-dipyrromethene-fluor (BODIPY-FL)-GTP (2.4 mM final concentration) was added to the GDP-loaded RAC1. To detect intrinsic RAC1 activity, buffer was added to the reaction prior to measurement. As positive control we used EDTA (12 mM final concentration). In the testing conditions, recombinant HisMBP-tagged DHR2 DOCK7 domain was added at the indicated final concentration. To evaluate myosin VI activity, recombinant MYO6 tail construct at the indicated concentration was pre-incubated on ice for 30 min with HisMBP-tagged DHR2 DOCK7 and then added to the reaction mixture. Kinetic hydrolysis of BODIPY-FL-GTP was measured at 30°C by monitoring the increase in fluorescence at excitation/emission wavelengths of 485/535 nm in a black 384-well microplate (Corning) using EnVision (PerkinElmer) plate reader. A reaction containing buffer, GDP and BODIPY-FL-GTP at the same final concentration was set up in order to subtract background value. All reactions were performed in technical triplicates. At the end of the measurement, samples were run on SDS-PAGE and analyzed by Coomassie staining. The assay was repeated at least three times/condition.

IHC analysis and quantification

Formalin-fixed and paraffin embedded (FFPE) tissue samples of human breast cancer cases (8 ductal *in situ* carcinoma and 8 invasive ductal carcinoma cases, collected and handled according to the Helsinki Declaration) were selected for the quantitative *in situ* immunophenotypical analyses. Four-micrometers-thick tissue sections were deparaffinized, rehydrated and unmasked using Novocastra Epitope Retrieval Solutions pH6 in thermostatic bath at 98°C for 30 min. Subsequently, the sections were brought to room temperature and washed in PBS. After neutralization of the endogenous peroxidase with 3% H_2O_2 and Fc blocking by 0.4% casein in PBS (Novocastra, Leica Microsystems), the samples were incubated for 90 min at room temperature with anti-myosin VI primary antibody (code #1296, homemade, 1:250). IHC staining was revealed using Novolink Polymer Detection Systems (Novocastra, Leica Microsystems) and Romulin AEC Chromogen kit (BioOptica) as substrate chromogen. Slides were counterstained with Harris hematoxylin (Novocastra, Leica Microsystems) and analyzed under a Zeiss Axioscope A1 microscope. Microphotographs were collected using a Zeiss Axiocam 503 Color digital camera with the Zen 2.0 Software (Zeiss). Quantitative analyses of IHC staining were performed by calculating the average percentage of positive signals in five nonoverlapping fields at medium-power magnification (200X) using the Positive Pixel Count v9 (2+ moderate positivity and 3+ strong positivity) ImageScope software. Average percentage of positive signals (3+ or 2+) was calculated in five nonoverlapping fields of view (200x). This study was approved by the University of Palermo Ethical Review Board (approval number 09/2018).

Expression profile analysis

The transcriptional myosin VI isoforms expression was analyzed by using TSVdb (http://www.tsvdb.com),[61] using RNA-seq data level of The Cancer Genome Atlas (TCGA) breast cancer (BRCA) dataset which includes 981 invasive breast carcinoma samples (and 112 normal breast samples) with complete clinical and pathological information together with PAM50 and SigClust Subtype assignments. Briefly, we downloaded the normalized expression level (RPKM) of short (isoform_uc003pii) and long (isoform_uc003pih) myosin VI isoforms mapped by TSVdb on GRCh37/hg19 genome assembly. RPKM data were +1 trimmed to avoid ‘‘0’’ or ‘‘Null’’ values before processing these to obtain ratios of expression between short versus long myosin VI isoforms. Statistical analyses and relative plots were done using JMP 17 (SAS) software.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using JMP 17 (SAS) software or PRISM software package. Statistical tests for data analysis included the log-rank test, Student’s t test (two-tailed), and two-way ANOVA. Multivariate statistical analysis was performed using a Cox regression model. p values are reported in each figure legends. $p < 0.05$ was considered statistically significant. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$. ns, not significant.