CYFIP2-containing WAVE complexes inhibit cell migration by a competition mechanism

Anna Polesskaya[1], Arthur Boutillon[2, #], Sheng Yang[3, #], Yanan Wang[1], Stéphane Romero[1], Yijun Liu[3], Marc Lavielle[4], Sophie Vacher[5], Anne Schnitzler[5], Nicolas Molinie[1], Nathalie Rocques[1], Artem Fokin[1], Raphaël Guérois[6], Ivan Bièche[5], Baoyu Chen[3], Nicolas B. David[2], Alexis M. Gautreau[1]

co-second author

[1] CNRS UMR7654, Institut Polytechnique de Paris, 91120 Palaiseau, France.

[2] INSERM U1182, CNRS UMR7645, Institut Polytechnique de Paris, 91120 Palaiseau, France.

[3]Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA 50011, USA.

[4] INRIA Saclay & Center for Applied Mathematics (CMAP), Institut Polytechnique de Paris, 91120 Palaiseau, France.

[5] Pharmacogenomics Unit, Department of Genetics, Institut Curie, 26 rue d'Ulm, 75005 Paris, France.

[6] Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, University Paris-Saclay, CEA-Saclay, 91190 Gif-sur-Yvette, France.

Correspondence concerning mathematical modeling should be addressed to ML (Marc.Lavielle@inria.fr)

Correspondence concerning biomedical aspects and material requests should be addressed to the lead contact, AMG (alexis.gautreau@polytechnique.edu)

Running title: CYFIP2 inhibits cell migration

Keywords: Patient cohort, multiprotein complexes, Zebrafish embryo, Scar/WAVE, WAVE Regulatory Complex, Arp2/3.

Abbreviations: BIC, Bayesian Information Criteria; KD, knock-down; KO, knock-out; MFS, metastasis-free survival; RFS, Relapse-free survival.

ABSTRACT

Branched actin networks polymerized by the Arp2/3 complex are critical for cell migration. The WAVE complex is the major Arp2/3 activator at the leading edge of migrating cells. However, multiple distinct WAVE complexes can be assembled in a cell, due to the combinatorial complexity of paralogous subunits. When systematically analyzing the contribution of each WAVE complex subunit to the metastasis-free survival of breast cancer patients, we found that overexpression of the CYFIP2 subunit was surprisingly associated with good prognosis. Gain and loss of function experiments in transformed and untransformed mammary epithelial cells, as well as in prechordal plate cells in gastrulating zebrafish embryos, revealed that lamellipodium protrusion and cell migration were always inversely related to CYFIP2 levels. The role of CYFIP2 was systematically opposite to the role of the paralogous subunit CYFIP1 or of the NCKAP1 subunit, which determines levels of WAVE complexes. CYFIP2 showed no difference from CYFIP1 in assembling WAVE complexes or binding to active RAC1. CYFIP2-containing WAVE complexes, however, were less able to activate the Arp2/3 complex in response to RAC1 binding. CYFIP1- and CYFIP2-containing WAVE complexes thus compete for active RAC1 and produce different outcomes. Therefore, cell migration, lamellipodium protrusion and Arp2/3 activity are controlled by relative levels of CYFIP1 and CYFIP2.

INTRODUCTION

Vertebrate genomes are the result of two genome-wide duplications [1]. This explains why many protein families are encoded by up to four paralogous genes in the human genome, but by a single gene in invertebrates such as *Drosophila* or *C. elegans*. The availability of several paralogous genes in the human genome has permitted the emergence of new regulations or specialized functions of specific paralogs. In cancers, alteration of gene expression or mutation usually concerns a single specific member of the family, which has to be identified.

Ten to twenty percent of human proteins form stable multiprotein complexes [2]. These complexes are often referred to as molecular machines to emphasize that they perform elaborate functions through the coordination of their subunits [3]. When several subunits are encoded by paralogous genes, a combinatorial complexity arises. Different complexes, potentially displaying different regulations and functions, stem from the different assemblies of paralogous subunits. If a specific molecular machine is responsible for cancer progression, it is also critical to be able to identify it.

Cell migration is controlled by several multiprotein complexes [4]. The Arp2/3 complex generates branched actin networks, which power membrane protrusions. At the protrusive edge, WAVE complexes activate the Arp2/3 complex [5,6]. The WAVE-Arp2/3 pathway depends on the activity of the small GTPase RAC1, which is necessary and sufficient to generate lamellipodia [7]. The RAC1-WAVE-Arp2/3 pathway controls protrusion lifetime and migration persistence through numerous feedback and feedforward loops [8]. This pathway has been implicated in the migration and invasion of tumor cells in various model systems [4].

The combinatorial complexity of WAVE complexes is daunting. A WAVE complex is composed of 5 generic subunits, hereafter referred to as WAVE, ABI, BRK, NAP and CYFIP. Except BRK, all human subunits are encoded by paralogous genes, 3 for WAVE and ABI, and 2 for NAP and CYFIP [9]. There are as many as 3x3x2x2, *i.e.* 36, possible WAVE complexes, just by combining the different paralogous subunits. Furthermore, the *ABI1* gene has been shown to be alternatively spliced and the resulting isoforms do not possess the same ability to mediate macropinocytosis, which, like lamellipodium formation, depends on the ability of branched actin to drive protrusions [10]. In mouse embryonic fibroblasts, WAVE2 is critical for the formation of peripheral ruffles, whereas WAVE1 is critical for dorsal ruffles [11]. Thus, evidence already exists for functional specialization among WAVE complexes.

WAVE complex subunits have been mostly reported to be overexpressed in tumors [4]. In line with their function in promoting cell migration and invasion, their overexpression is generally associated with high grades and poor prognosis. High levels of WAVE subunits is of poor prognosis for patients in breast, ovary, lung and liver cancers [12–16]. The overexpression of WAVE3 in colorectal cancers, however, is associated with good prognosis [17]. Similar to the general trend, high expression of the NAP paralogs, NCKAP1 and NCKAP1L, has been associated with poor prognosis in breast cancer and leukemia, respectively [18,19]. High expression of ABI1 has also been associated with poor prognosis in breast and ovary cancers [20,21].

Whereas most studies, including cancer studies, focused on one subunit, we measured the expression levels of all the paralogous genes encoding subunits in a large cohort of breast cancer patients, in an attempt to tackle the complexity of the WAVE complex. This systematic endeavor allowed us to examine each of the 36 possible WAVE complexes for their possible association with metastasis-free survival (MFS). We found no evidence for the involvement of a specific WAVE complex assembly. The first order determinant of MFS was whether WAVE complexes contained the NCKAP1 subunit. The second order determinant was whether WAVE complexes contained the CYFIP2 subunit. Surprisingly, however, we found that high levels of *CYFIP2* were associated with good prognosis. This unexpected effect on MFS could be accounted for by the fact that CYFIP2-containing complexes specifically impair cell migration in a variety of cell systems. CYFIP2-containing WAVE complexes are less activated by RAC1 than CYFIP1-containing WAVE complexes, suggesting that they titrate out the major activator of the WAVE complex.

RESULTS

Systematic analysis of WAVE complex subunits in breast cancer

In a cohort of 527 breast cancer patients (Table S1), we measured by qRT-PCR the mRNA levels of the 11 genes encoding WAVE complex subunits. Expression values in tumors were normalized to the expression in healthy breast tissue. We found that the expression of several subunits is profoundly deregulated in breast cancer (Table 1). *CYFIP2*, *NCKAP1L* and *ABI3* were up-regulated in 37%, 22% and 12% of tumors, respectively. Cases of overexpression were in different subgroups of breast cancer patients. *NCKAP1L* is mostly overexpressed in the Hormone Receptor (HR)- ERBB2+ subgroup. *ABI3* is mostly overexpressed in the HR- ERBB2-, triple negative subgroup. *CYFIP2* is mostly overexpressed in the HR+ ERBB2- subgroup and in low-grade tumors of good prognosis (Table S1). *WASF3* and *WASF1* are down-regulated in 46% and 27% of the cohort. Underexpression of these WAVE subunits is also mostly displayed in the good prognosis HR+ ERBB2- subgroup. We then examined if fluctuations in subunit expression were associated with prognosis.

Since the outcome of patients is known in the cohort and given the role of the WAVE complex in tumor cell invasion, we were especially interested in the metastasis-free survival (MFS). MFS starts at the date of surgery and terminates at the date of the last news from the patient, of metastasis diagnostic, or of death. We applied to these right-censored data a classical Cox univariate model using the expression level of each subunit as the variable. We sorted the different subunit genes according to increasing p-values (Fig. S1). The first three genes were *NCKAP1*, *CYFIP2* and *NCKAP1L*. The levels of *NCKAP1* mRNA, within their natural fluctuations, were significantly associated with MFS (p=0.012, Fig. S1). Indeed, we previously reported that high levels of *NCKAP1* were associated with poor MFS [18]. Levels of *CYFIP2* and *NCKAP1L* also appeared significantly associated with MFS, but with a lower significance, p=0.138 and 0.288, respectively.

Our goal when measuring expression levels of all WAVE subunits in the cohort was to examine whether a particular combination of subunits would create a specific WAVE complex conferring invasive properties to tumor cells. This is why we chose to perform highly accurate measurements by qRT-PCR in our cohort of 527 patients, even if global analyses were already available in public databases containing a larger number of patients. To analyze the association of various WAVE assemblies with MFS, we needed to transform and normalize our variables, i.e. subunit levels. Using a monotonous function of the type $\log(x\text{-}c)$, levels of each subunit fitted a Gaussian distribution. Then we normalized transformed variables around 0 with a variance of 1, to allow a better comparison between different subunit levels. Transformation and normalization did not change the relative association of subunit levels with MFS, since, by univariate Cox analysis, the 3 most powerful subunits to predict MFS were still, first, *NCKAP1* with a p-value of 0.005, second, *CYFIP2* with a p-value of 0.059, just above the classical 5% significance level, but far above the third subunit, *NCKAP1L*, with a p-value of 0.397 (Fig. S1).

During these simple Cox analyses of the original subunit levels or of the transformed and normalized variables, we were struck by the fact that *NCKAP1* and *CYFIP2* had opposite coefficients for the association with MFS. Indeed, high levels of *NCKAP1* were associated with poor MFS, whereas high levels of *CYFIP2* were associated with good MFS (Fig.S1).

Using transformed and normalized variables, we were able to perform a multivariate Cox analysis to analyze the association of each of the 36 possible WAVE complexes with MFS (Fig.S1). We sorted the 36 WAVE complexes according to increasing p-values. The 18 best combinations all contained NCKAP1 as the NAP subunit, while the best 9 combinations also contained CYFIP2 as the CYFIP subunit, suggesting that NCKAP1 is the first order predictor, whereas CYFIP2 is the second order predictor in our cohort. The multivariate Cox analysis does not suggest a specific WAVE assembly that would be particularly associated with MFS, what seems to matter is whether the assembly contains NCKAP1 and/or CYFIP2. Computer simulations using random permutations of values confirmed that the prediction powers of NCKAP1 and CYFIP2 cannot be attributed to chance (Fig.S1).

We then evaluated further multivariate Cox models by adding up to 5 variables using NCKAP1, CYFIP2, WASF3 and ABI2 and BRK1 subunits in this order. The log-likelihood criterium increased when more subunits were introduced, but the log-likelihood always increases when further variables are added. Therefore, we compared the models using Bayesian Information Criteria (BIC). BIC introduces a penalty term for the number of variables used in the model to avoid overfitting. The model with 2 variables, NCKAP1 and CYFIP2, had the smallest BIC (Fig.S1) and thus appeared as the optimal model of MFS in our cohort. MFS over time can be accurately predicted from mRNA levels of *NCKAP1* and *CYFIP2*. In our optimal model, *NCKAP1* is a first order predictor with a p-value of 0.001, whereas *CYFIP2* is the second order predictor with a p-value of 0.012. Importantly, in this multivariate model, as in the initial univariate models, *NCKAP1* and *CYFIP2* have opposite coefficients, indicating that up-regulation of *NCKAP1*, but down-regulation of *CYFIP2*, are associated with poor prognosis. In the model, the higher the *CYFIP2* value, the better the MFS, for a given value of *NCKAP1*. To illustrate how the second order predictor *CYFIP2* modulates the MFS, we ran the model with expression levels found in patient tumors populating the outskirts of the distribution (Fig.1A). The extreme values of *NCKAP1* dominate the predicted MFS when *CYFIP2* values are

intermediate (Fig.1B). In contrast, extreme values of *CYFIP2* significantly oppose the effect of *NCKAP1*, when *NCKAP1* values are not extreme.

To validate the prediction of our statistical model that NCKAP1 and CYFIP2 control MFS, we used a public database of breast cancer patients, where the transcriptome of more than 3900 tumors, was analyzed by Affymetrix chips [22]. Given the large number of patients, more genes encoding WAVE complex subunits were significantly associated with relapse-free survival (RFS) than in our cohort containing slightly more than 500 patients. However, the two most strongly associated ones were *NCKAP1* and *CYFIP2*, as in our analyses. As our model predicted, high levels of NCKAP1 were associated with poor RFS, whereas high levels of CYFIP2 were associated with good RFS (Fig.1C). All these results together indicate that CYFIP2 should have a function at odds with the major function of WAVE complexes, that is to promote cell migration [4].

The WAVE complex subunit CYFIP2 inhibits the migration of mammary carcinoma cells

Since the expression of WAVE subunits CYFIP2 and NCKAP1 are associated with opposite prognoses in breast cancer patients, we sought to compare their function in mammary carcinoma cells. Moreover, we compared the two paralogous subunits CYFIP1 and CYFIP2 in two classical breast cancer cell lines, MCF7, which is HR+ ERBB2-, and MDA-MB-231, which is HR- ERBB2- (triple negative).

Depletion of the different subunits using RNAi had different impact on WAVE complex levels. Indeed, WAVE complexes are stable when fully assembled, providing an explanation as to why depletion of a subunit usually destabilizes the multiprotein complex it should be part of [9]. Depletion of NCKAP1 in MCF7 cells leads to a severe downregulation of WAVE complex subunits, including CYFIP1 and CYFIP2 (Fig.2A). This result shows a key role of NCKAP1 for the stability of CYFIP1- and CYFIP2-containing WAVE complexes in cells. Depletion of CYFIP1 leads to a significant destabilization of the WAVE complex, which can be appreciated on NCKAP1, WAVE2 and BRK1 levels. In contrast, depletion of CYFIP2 does not lead to a visible depletion of the same subunits. Since MCF7 cells have conserved their epithelial organisation, we assessed cell migration in a wound healing assay. CYFIP1 or NCKAP1 depleted cells failed to close the wound before 78-81h compared to 30 h for control cells (Fig. 2A and Movie S1). In sharp contrast, CYFIP2 depleted cells were not impaired in their ability to close the wound and were in fact significantly faster than controls (20 h vs. 30 h).

We then turned to MDA-MB-231 cells, which displayed the same overall pattern of subunit expression upon depletion of NCKAP1 or CYFIP1/2 as MCF7 cells (Fig.2B). Briefly, WAVE complexes were destabilized upon NCKAP1 or CYFIP1 depletion, but not upon CYFIP2 depletion. We thus decided to measure levels of CYFIP1 and 2 using purified CYFIP1- or CYFIP2-containing WAVE complexes as standards, serial dilutions and Western blots in their linear range (Fig.S2). We found that MDA-MB-231 express roughly 6-times more CYFIP1 than CYFIP2 (Fig.2C), providing a first level of explanation as to why CYFIP1 depletion affects more the stability of other subunits than CYFIP2. However, this was not the

only effect, since the depletion of each CYFIP protein resulted in approximately 50 % up-regulation of its remaining CYFIP paralog. These up-regulations of the paralogous CYFIP proteins were not observed at the mRNA level and might represent stabilization of CYFIP subunits with other WAVE complex subunits when the paralogous CYFIP protein is not expressed (Fig.S3).

We first evaluated in the Transwell assay the migration of MDA-MB-231 cells when CYFIP1, 2, or NCKAP1 were depleted using pools of siRNAs. Depletion of NCKAP1 and CYFIP1 significantly decreased the number of cells able to migrate through the filter, whereas the depletion of CYFIP2 had the converse effect (Fig.2D). The effect of transient siRNA-mediated depletions was confirmed using stable MDA-MB-231 lines expressing either a shRNA targeting NCKAP1 or CYFIP2 [5](Fig.S4). We then attempted to obtain stable MDA-MB-231 lines overexpressing NCKAP1 or CYFIP2. We obtained lines expressing GFP-tagged CYFIP2, but repeatedly failed in obtaining clones expressing NCKAP1 in parallel selection schemes. The overexpression of CYFIP2 slightly decreased cell migration in the Transwell assay, whereas CYFIP2 depletion increased it (Fig.S4). Loss- and gain-of function of CYFIP2 thus yield opposite phenotypes.

MDA-MB-231 cells are mesenchymal, unlike epithelial MCF7 cells. Nonetheless, in a wound healing assay, CYFIP2 depleted MDA-MB-231 closed the wound faster than controls, and MDA-MB-231 cells depleted of NCKAP1 and CYFIP1 were significantly delayed in doing so (Fig.S4), exactly as we had observed in MCF7 cells. We then turned to a more physiopathological assay for MDA-MB-231. We seeded isolated MDA-MB-231 cells in 3D gels of collagen type I. In these settings, mimicking invasion of the mesenchyme, differences in cell migration were more dramatic (Fig. 2E). NCKAP1 depleted cells hardly migrated at all, as evidenced by strongly decreased Mean Squared Displacement (MSD), mostly due to reduced speed. NCKAP1 depleted cells ended up entering into apoptosis during the first 24 h (Movie S2). CYFIP1 depleted cells were not significantly affected in their ability to migrate, even though they also appeared prone to die in these settings. CYFIP1 and NCKAP1 depleted cells formed significantly fewer protrusions than controls (Fig.2E). In contrast, CYFIP2 depleted cells often explored a significantly larger territory than controls. The increased MSD of CYFIP2 depleted cells could be accounted for by the dramatically increased migration persistence. The protrusive activity of CYFIP2 depleted cells was significantly increased compared to controls. Finally, CYFIP2 depleted cells had no issue of survival in 3D collagen, unlike cells depleted of NCKAP1 or CYFIP1.

In conclusion, in all the assays performed with the two breast cancer cell lines, the opposite roles of NCKAP1 and CYFIP2 were consistently observed and in line with their association with their prognostic roles in the metastasis-free survival of breast cancer patients.

CYFIP2 inhibits cell migration of untransformed cells

We wondered if the anti-migratory role of CYFIP2 was its normal function or rather associated with cell transformation. To address this question, we used the immortalized, but not transformed, MCF10A mammary cell line. MCF10A expressed approximately 6-fold more CYFIP1 than CYFIP2 (Fig.3A). CYFIP2 is thus less abundant than CYFIP1 in two cell lines, MCF10A and MDA-MB-231. As in MDA-MB-231, siRNA-mediated depletion of NCKAP1 and CYFIP1 from MCF10A cells significantly decreased protein levels of WAVE complex subunits, but not their mRNA levels (Fig.3B, Fig.S3), whereas CYFIP2 depletion did not affect the overall stability of WAVE complexes.

MCF10A cells are more epithelial than MDA-MB-231 cells. They establish cell-cell junctions and form epithelial islets. However, they are plastic epithelial cells. In 2D cultures, in their regular culture medium, which contains EGF, MCF10A cells display cell-cell junctions, but also frequently migrate as single cells. We depleted MCF10A cells with siRNA pools targeting either NCKAP1, CYFIP1 or CYFIP2. Cells depleted of NCKAP1 appeared as small and organized as a tight epithelium, whereas the cells depleted of CYFIP2 appeared larger with membrane protrusions, even if they remained associated with one another (Fig.3B, Movie S3). CYFIP1 depletion did not have a pronounced effect on cell morphology. We then recorded MCF10A cells to analyze cell migration. Trajectories corresponding to single cells were plotted (Fig.3C). NCKAP1 depleted single cells migrated much less than controls, an effect which was mostly due to decreased cell speed. In contrast, CYFIP2 depleted cells did not explore a wider territory than controls, nor did they migrate faster, but they significantly increased migration persistence. Importantly, same results were obtained with two single siRNA sequences for each gene (Fig.S5), indicating that these results were not due to off-targets. Such a phenotype, characterized by increased migration persistence of single MCF10A cells, was previously observed upon activation of RAC1 or upon depletion of the Arp2/3 inhibitory protein ARPIN [23].

To study differentiation of acini in Matrigel, we isolated *CYFIP2* knock-out (KO) clones using CRISPR-Cas9. From about 100 independent MCF10A clones, we selected two CYFIP2 negative clones, which turned out to be KO on both alleles due to insertions/deletions changing the ORF (Fig.S6A). As expected, *CYFIP2* KO clones displayed increased migration persistence (Fig.S6BC). The initial KO of CYFIP2 led to an increase of the level of WAVE complex subunits, as previously shown for siRNA assays, but this effect disappeared in long-term cultures (Fig.S6DE). The differentiation of *CYFIP2* KO clones was then assayed in Matrigel, where MCF10A cells develop acini structures. *CYFIP2* inactivation did not affect the morphogenetic program, nor cell polarity, but resulted in significantly larger 3D structures containing more cells than the control (Fig.3D). Similar results were previously obtained when ARPIN was inactivated [23]. CYFIP2 thus behaves like this well-established inhibitory protein of cell migration, ARPIN. This result on acini structures is consistent with the fact that the RAC1-WAVE-Arp2/3 pathway controls cell cycle progression [23].

To validate the anti-migratory function of CYFIP2 in a physiological system and to test whether this function is specific to breast cells or more general, we turned to the zebrafish embryo, and in particular to prechordal plate cells, which stereotypically migrate during

gastrulation [24,25]. Prechordal plate cells migrate from the fish organizer (shield) to the animal pole of the embryo by forming actin-rich protrusions. These RAC1 dependent protrusions are the 3D equivalents of 2D lamellipodia and are easily distinguished from thin, filopodia-like extensions [26,27]. We assessed the function of CYFIP1, CYFIP2 and NCKAP1 using both morpholino-mediated loss-of-function and mRNA over-expression.

We first analyzed prechordal plate cell trajectories, in embryos injected with morpholino and/or mRNA for CYFIP1, CYFIP2 and NCKAP1 (Fig.4A). Experiments were performed in a goosecoid:GFP transgenic line, allowing easy identification of prechordal plate cells. Nuclei were labelled by expression of a Histone2B–mCherry construct, the cells were tracked (Movie S5), and cell trajectories were plotted. Similar to what was observed using human cell lines, CYFIP2 depletion increased migration persistence as compared to injection of a control morpholino. This effect was rescued by co-injection of a morpholino-insensitive CYFIP2 mRNA, demonstrating the specificity of the phenotype. Consistently, overexpression of CYFIP2, i.e. injection of the same amount of mRNA as for the rescue but without the corresponding morpholino, decreased cell persistence. In contrast to CYFIP2, downregulation of CYFIP1 or NCKAP1 reduced cell persistence, both effects being rescued by the co-injection of the corresponding mRNAs.

We then used cell transplants to look for cell autonomous defects and analyzed cell dynamics and protrusivity. Few prechordal plate cells from a donor embryo injected with morpholino and/or mRNA were transplanted to the prechordal plate of an uninjected host embryo (Fig.4B). Actin-rich protrusions were highlighted by the enrichment of the LifeAct-mCherry marker (Fig.4B, Movie S6). CYFIP2 depletion doubled the number of protrusions compared to cells injected with a control morpholino (Fig.4B). This effect was rescued by a morpholino-insensitive CYFIP2 mRNA. Consistently, CYFIP2 overexpression decreased the number of protrusions, much like the depletion of NCKAP1 and CYFIP1. CYFIP2 depletion also significantly and specifically increased protrusion length (Fig.4B).

The results using zebrafish embryos are thus perfectly in line with those obtained in human breast cells and demonstrate that the unexpected anti-migratory function of CYFIP2 is a general and conserved function of this subunit, at least across vertebrates.

CYFIP2 rescues lamellipodium formation in *CYFIP1/2* double KO cells

To examine whether CYFIP2 was a functional subunit of the WAVE complex, we re-expressed CYFIP2 in B16-F1 *CYFIP1/2* double knock-out cells (DKO) [28]. GFP-CYFIP2 clearly rescued lamellipodium formation in DKO cells, like GFP-CYFIP1, even if CYFIP1 appeared to induce more prominent lamellipodia than CYFIP2 (Fig.5AB, movie S6). We also analyzed two point mutations of CYFIP2, R87C and I664M, that are recurring mutations found in patients affected by intellectual disability [29]. These two point mutations did not impair the ability of CYFIP2 to induce lamellipodia. On the opposite, the mutations seemed to induce more prominent lamellipodia (movie S7).

We quantitatively analyzed the effect of CYFIP2 and of its mutated derivatives on protrusions using line scans. The expression of CYFIP2 was unable to restore the full speed of

protrusions observed in parental B16-F1 cells or in CYFIP1-rescued DKO cells (Fig.5C). R87C and I664M mutations rendered CYFIP2 significantly more efficient at rescuing the protrusion rate, up to the level of parental or CYFIP1 reconstituted cells. We then analyzed the width of lamellipodia, using immunofluorescence of the ARPC2 subunit of the Arp2/3 complex and of cortactin, a protein that stabilizes the Arp2/3 at the branched junction between filaments [4]. In line with their faster protrusions, CYFIP1, R87C and I664M CYFIP2 induced lamellipodia deeper into the cell than wild type CYFIP2 (Fig.5DE). So CYFIP2 is a functional CYFIP protein, but less active than CYFIP1, and point mutations that induce developmental defects in patients alleviate this restrained activity of CYFIP2.

CYFIP2 containing WAVE complexes are poorly activated by active RAC1

CYFIP2 is 88 % identical to CYFIP1. So we replaced CYFIP1 with CYFIP2 in the molecular model of the WAVE complex derived from crystallography of a reconstitued complex containing a WAVE1 form lacking the central proline-rich region and a truncated ABI2 lacking the the disordered C-terminus [30]. We mapped on this model the WIRS binding site that allows the WAVE complex to interact with various transmembrane receptors [31]. None of these binding sites was affected by the substitutions in CYFIP2 (Fig.6A). Consistently, we observed by ultracentrifugation on sucrose gradients that CYFIP2 was incorporated into the native WAVE complex, which sediments at around 11 Swedbergs (Fig.6B)[32]. Two binding sites exist for active RAC1 at the surface of the WAVE complex: the so-called A site, which shares structural homology with the CYRI-B protein [33,34], and the D site [35]. None of the RAC1 binding sites was affected by substitutions in CYFIP2. To examine the binding to RAC1 and activation by RAC1, we reconstituted a WAVE complex with either CYFIP1 or CYFIP2 using a previously described procedure [36]. Both complexes interacted equally well with GTP-bound RAC1 (Fig.6C and D). In pyrene-actin polymerization assays, however, the CYFIP2-containing WAVE complex was poorly activated by RAC1 compared to the CYFIP1-containing WAVE complex (Fig.6E). These in vitro data are thus consistent with the observation that in cells CYFIP2 promotes lamellipodium formation, but not as well as CYFIP1.

DISCUSSION

Here we have systematically analyzed the expression levels of WAVE complex subunits in a cohort of breast cancer patients. *Ad hoc* statistical modeling, taking into account assembly rules among paralogous subunits, increased the statistical power of the analysis and revealed the unique role of the CYFIP2 subunit, whose overexpression is associated with good prognosis for metastasis-free survival. These findings were validated using an independent cohort of breast cancer patients available in public databases. *CYFIP2* had previously been implicated in pathologies, since it is mutated in children affected with intellectual disability and epileptic encephalopathy [37,38]. In zebrafish, CYFIP2 loss-of-function mutations result in defective axonal pathfinding in retinal ganglion cells [39]. This function of CYFIP2 is also not redundant with the one of the paralogous subunit, CYFIP1, which is involved in axon growth [40].

We have experimentally validated the prediction of our model, which implies a protective role of CYFIP2 overexpression in breast cancer. CYFIP2 is at odds with other subunits, since it is the first subunit of the WAVE complex that is ever reported to oppose cell migration. Indeed, we found that CYFIP2 opposes cell migration in a variety of cell systems, MCF10A, MDA-MB-231 and prechordal plate cells from the zebrafish embryo. In these experiments, CYFIP2 depletion enhances cell migration, whereas CYFIP2 overexpression decreases cell migration. We were struck by this anti-migratory role of CYFIP2, which to our knowledge was never reported before, even if depletion of different subunits of the WAVE complex did not always give the same phenotype [41–45].

In all cell systems we studied here, the main parameter that CYFIP2 controls is migration persistence, which relates to the persistence of lamellipodial protrusions [8]. In fact, the role of CYFIP2 is very similar to the Arp2/3 inhibitory protein Arpin that directly inhibits the Arp2/3 complex at the leading edge [46]. In neuronal growth cones, CYFIP2 was found to localize at the tip of filopodia, structures composed of linear actin and not of branched actin [40], in line with an inhibitory function of CYFIP2 on branched actin formation we suggest here.

CYFIP2 is highly related to CYFIP1, with 88 % identity. Both CYFIP proteins incorporate into WAVE complexes [47–49]. Accordingly, we found here that CYFIP2 depends on NCKAP1 for its stability, like CYFIP1 and that CYFIP2 is found into the same WAVE complex migrating at 11 Swedbergs as CYFIP1. Importantly, the residues of CYFIP1 that are involved in binding active RAC1 are all conserved in CYFIP2. However, we found that CYFIP2-containing WAVE complexes were less activatable by active RAC1 than CYFIP1-containing WAVE complexes. This property accounts for the observed phenotypes. Indeed, depletion of CYFIP2 can render available more active RAC1 to activate more CYFIP1-containing WAVE complexes, which are easily activatable (Fig.6F).

Another effect is expected to increase migration upon CYFIP2 depletion and to decrease it upon CYFIP1 depletion. In RNAi experiments, CYFIP1-depleted cells were found to overexpress CYFIP2, whereas CYFIP2-depleted cells were found to overexpress CYFIP1. This compensatory expression of the paralogous CYFIP was marginally observed at the mRNA level in some cases and more significantly at the protein level, suggesting that CYFIP proteins might be stabilized by the availability of partner subunits [9]. However, this compensatory expression of CYFIP1 is lost over time in the *CYFIP2* KO clones that we isolated from MCF10A cells, while enhanced migration persistence of *CYFIP2* KO clones is sustained, strongly suggesting that this effect has a minor contribution compared to the poor activation of CYFIP2-containing complexes by active RAC1.

This modulation of phenotypes based on subunit composition of complexes was previously described for the Arp2/3 complex [50]. The paralogous subunits ARPC1B and ARPC5L assemble Arp2/3 complexes, which are more activatable than the ones assembled around ARPC1A and ARPC5. The situation is perfectly analogous to the one described here for the WAVE complex assembled with CYFIP1 and CYFIP2 paralogous subunits. Vertebrate genomes encoding paralogous subunits for many stable multiprotein complexes thus offer numerous opportunities to fine tune cellular responses. The two examples of WAVE and Arp2/3 complexes illustrate that each cell of a vertebrate organism can regulate levels of cortical

branched actin, polymerized in response to signaling inputs, based on the expression of the paralogous genes that regulate the "activatability" of these molecular machines.

METHODS

Patient cohort for mRNA analysis

All patients (mean age 60.9 years, range 29-91 years) met the following criteria: primary unilateral nonmetastatic breast carcinoma for which complete clinical, histological and biological data were available; no radiotherapy or chemotherapy before surgery; and full follow-up at Institut Curie - Hospital René Huguenin. All patients before 2007 were informed that their tumor samples might be used for scientific purposes and had the opportunity to decline. Since 2007, patients treated in our institution have given their approval by signed informed consent. This study was approved by the local ethics committee (Breast Group of René Huguenin Hospital). Treatment (information available for 524 patients) consisted of modified radical mastectomy in 320 cases (61%) or breast-conserving surgery plus locoregional radiotherapy in 204 cases (39%). The patients had a physical examination and routine chest radiotherapy every 3 months for 2 years, then annually. Mammograms were done annually. Adjuvant therapy was administered to 416 patients, consisting of chemotherapy alone in 130 cases, hormone therapy alone in 178 cases and both treatments in 108 cases. During a median follow-up of 10.5 years (range 1 month to 36.3 years), 210 patients developed metastasis. Sixteen specimens of adjacent normal breast tissue from breast cancer patients or normal breast tissue from women undergoing cosmetic breast surgery were used as sources of normal RNA.

qRT-PCR

Specific mRNAs were quantified from the cycle number (Ct value) at which the increase in the fluorescence signal started to be detected by the laser detector of the ABI Prism 7900 sequence detection system (Perkin-Elmer Applied Biosystems, Foster City, CA) as previously described [52]. Specific transcripts were quantified using the following primers: WASF1-U (5'-CCTCTCATTTTGAAACAAGACCTCAG-3') and WASF1-L (5'-CTAAATGGCAAGGCAGAAAGTGAGT-3') for the *WASF1* gene (PCR product of 79 pb); WASF2-U (5'- AAAGCTGGGGACTTCTGGGTATC-3') and WASF2-L (5'-GTGAAGAAGCAGAGTCTGACTGTGGT-3') for the *WASF2* gene (PCR product of 122 pb); WASF3-U (5'- GAGTGATAAGCCACCGCCTCTG-3') and WASF3-L (5'-GCCCATCCTTCTTGTCATCTCTGTA-3') for the *WASF3* gene (PCR product of 62 pb); ABI1-U (5'-GGGGAACACTGGGACGGAAT-3') and ABI1-L (5'-GCTGTCCTGCCTGGACTATGCT-3') for the *ABI1* gene (PCR product of 124 pb); ABI2-U (5'-CCGTGGGCTCCACGTTCTTACT-3') and ABI2-L (5'-TCCTTCCTGAAAGGACAGCTCATCT-3') for the *ABI2* gene (PCR product of 90 pb); ABI3-U (5'-TGCTGCGGGTCGCTGACTA-3') and ABI3-L (5'-

GCGCCTTCCGCTTGTCTGT-3') for the *ABI3* gene (PCR product of 63 pb); BRK1-U (5'-AAAATCGCAGACTTTCTCAACTCGT-3') and BRK1-L (5'-TTCAAGGGCTGTCAATTTCTCGT-3') for the *BRK1* gene (PCR product of 84 pb); NCKAP1-U (5'-AGTGTACCCTTAGTGACCAGTTGCT-3') and NCKAP1-L (5'-TCAGGTTCCCCTTTCTTACCAGT-3') for the *NCKAP1* gene (PCR product of 106 pb); NCKAP1L-U (5'- GAAAAGTCCATGGAACCATCTCTCA-3') and NCKAP1L-L (5'-GTACTGGTCCTAAATGTTGCGTGCT-3') for the *NCKAP1L* gene (PCR product of 91 pb); CYFIP1-U (5'-CACGAGTACGGCTCTCCTGGTATC-3') and CYFIP1-L (5'-CCGCAGGTTCTGGAAGCACA-3') for the *CYFIP1* gene (PCR product of 102pb); CYFIP2-U (5'-CCCACGTCATGGAGGTGTACTCT-3') and CYFIP2-L (5'-TAATTGTAGCGTGTGGCTCTCTCA-3') for the *CYFIP2* gene (PCR product of 112pb); TBP-U (5'-TGCACAGGAGCCAAGAGTGAA-3') and TBP-L (5'-CACATCACAGCTCCCCACCA-3') for the *TBP* gene (PCR product of 132 bp), which was the reference gene used for normalization. Over and under-expression were defined as >3 and <0.33, respectively, the expression compared to the median expression of normal samples.

Public transcriptomics data on breast cancer [22] were interrogated using the kmplot website (http://kmplot.com) on June 26, 2019 using best cut-offs for JetSet determined best probes (*NCKAP1* 207738_s_at, *CYFIP2* 220999_s_at, [51].

Cell lines, transfection and establishment of stable clones

MCF10A cells were grown in DMEM/F12 medium supplemented with 5% horse serum, 20 ng/mL epidermal growth factor, 10 µg/mL insulin, 500 ng/mL hydrocortisone, and 100 ng/mL cholera toxin. MDA-MB-231 were grown in DMEM medium with 10% FBS. Medium and supplements were from Life Technologies and Sigma. Cells were incubated at 37°C in 5% CO_2. MCF10A and MDA-MB-231 were from the collection of breast cell lines organized by Thierry Dubois (Institut Curie, Paris).

Stable MCF10A cells expressing CYFIP2 were obtained by transfecting MCF10A cells, with the home-made plasmid MXS AAVS1L SA2A Puro bGHpA EF1Flag GFP CYFIP2 Sv40pA AAVS1R, or MXS AAVS1L SA2A Puro bGHpA EF1Flag GFP Blue Sv40pA AAVS1R as a control. Transfection was performed with Lipofectamine 2000 (Invitrogen). To obtain stable integration of the MXS plasmid at the AAVS1 site, cells were cotransfected with two TALEN plasmids inducing DNA double strand breaks at the AAVS1 locus (Addgene #59025 and 59026; [52]. Cells were selected with 1 µg/mL puromycine (Invivogen) and pooled. Stable MCF10A cells expressing shRNA were obtained by transfection with previously described pSUPER-Retro-Puro plasmids [5] and puromycin selection.

The stable 293 Flp-In cell line expressing Flag-HA-CYFIP1 were previously described [53]. An equivalent cell line expressing Flag-HA-CYFIP2 was obtained according to a published procedure[54].

MDA-MB-231 and MCF10A were depleted by siRNAs (OnTarget Smart Pools, Dharmacon), transfected at 20 nM final concentration using Lipofectamine RNAiMAX (Invitrogen), and re-

transfected 72h later, for the total of 6 days. This protocol was necessary due to an unusually long half-life of CYFIP2 protein (AP, unpublished observations).

The MCF10A CYFIP2 knockout cell line was generated with CRISPR/Cas9 system. The targeting sequence 5'-CAUUUGUCACGGGCAUUGCA-3' was used to induce the double strand break. For the negative control the non-targeting sequence 5'-AAAUGUGAGAUCAGAGUAAU-3' was used. Cells were transfected with crRNA:trackRNA duplex and the purified Cas9 protein by Lipofectamine CRISPRMAX™ Cas9 Transfection Reagent (all reagents from Thermofisher Scientific). The next day, cells were subjected to dilution at 0.8 cells/well in 96 well plates. Single clones were expanded and analyzed by CYFIP2 Western blot. 2 positive clones were identified. The PCR products amplified from genomic DNA containing the gRNA recognition site were then cloned (Zero Blunt PCR Cloning Kit, Thermofisher Scientific) and sequenced. A frameshift of +1 and a -1 in the 3rd exon of the CYFIP2 gene in both clones was confirmed by sequencing (see Fig. S6 for details).

Antibodies and Western blot

Cells were lysed in RIPA buffer and analyzed by Western blot. SDS-PAGE was performed using NuPAGE 4-12% Bis-Tris and 3-8% Tris-Acetate gels (Life Technologies). Nitrocellulose membranes were developed with horseradish peroxidase (HRP) coupled antibodies (Sigma) and SuperSignal West Femto chemiluminescent substrate (Thermo Fisher Scientific). Home-made rabbit polyclonal antibodies CYFIP1, ABI1, WAVE2 were previously described [32]. The mouse monoclonal antibody, 231H9, targeting BRK1 was previously described [55]. The antibodies targeting CYFIP-2 (Sigma SAB2701081), NCKAP1 (Bethyl A305-178A), cortactin (Millipore 4F11), ARPC2 (Millipore 07-227) and tubulin (Sigma T9026) were purchased. Quantification of wb was performed by densitometry, using the ImageJ software.

Sucrose gradient

For sucrose gradient analysis of WAVE subunits, Nitrogen cavitation (Parr instruments, 500 Psi for 20 min) followed by centrifugation (16,000 × g, 20 min) and ultracentrifugation (150,000 × g, 60 min) were used to prepare cytosolic extracts from cells trypsinized from two 15 cm dishes and resuspended in the XB buffer (20 mM HEPES, 100mM NaCl, 1mM MgCl2, 0.1 mM EDTA, 1mM DTT, pH 7.7). 200 µL of extract was loaded on the 11 mL 5–20% sucrose gradient in the XB buffer and subjected to ultracentrifugation for 17 h at 197,000 ×g in the swinging bucket rotor SW41 Ti (Beckman). 0.5 mL fractions were collected and concentrated by using trichloroacetic acid precipitation with insulin as a carrier. The samples were washed with acetone, dried and then resuspended in the 1x LDS loading buffer with 2.5% of β-ME for Western blot analysis.

Migration assays

Transwell migration assays were performed using FluoroBlok inserts with 8 µm holes (Corning, 351152), covered with 20 µg/ml fibronectin (Sigma, F1141). MDA-MB-231 cells

were plated in serum-free medium and allowed to migrate towards serum-containing medium for 16 h, incubated with 4 µg/ml calcein AM (Sigma, C1359) for 1 h, and images of fluorescent cells were acquired and quantified using ImageJ software.

2D migration was performed using 8 chamber Ibidi dishes (Biovalley 80826) covered with 20 µg/ml fibronectin. 3D migration was performed in 2 mg/ml collagen gel polymerized at 37°C (rat tail collagen type I, Corning 354263), with the cells sandwiched between the two layers of collagen. An inverted Axio Observer microscope (Zeiss) equipped with a Pecon Zeiss incubator XL multi S1 RED LS (Heating Unit XL S, Temp module, CO_2 module, Heating Insert PS and CO_2 cover), a definite focus module and a Hamamatsu camera C10600 Orca-R2 was used to perform videomicroscopy. Pictures were taken every 5 min for 24 h for 2D migration, and every 20 min for 48 h for 3D migration. Random migration of single cells and migration persistence, based on the angular shift between frames, was analyzed as previously described [46] using DiPer programs [56].

Rescue of DKO cells

B16-F1 mouse melanoma cells that are *CYFIP1/2* double KO were a kind gift of Klemens Rottner (Helmholtz-Zentrum für Infektionsforschung, Braunschweig). GFP-tagged human CYFIP1 or CYFIP2 (wild type or mutant) were transiently transfected into the DKO cells, and 48 h later, 10-minute videos (images taken every 10 seconds) were acquired using a confocal laser scanning microscope (TCS SP8, Leica) equipped with a high NA oil immersion objective (HC PL APO 63×/ 1.40, Leica), a white light laser (WLL, Leica) and controlled by the LasX software. Protrusion speed was measured using the Multi Kymograph tool in ImageJ software. For the LineScan analysis, images of fixed, stained cells were obtained, and analyzed as described in [46] and [23].

Zebrafish embryos, cell transplantation and imaging

Embryos were obtained by natural spawning of *Tg(-1.8gsc:GFP)ml1* adult fishes [57]. All animal studies were done in accordance with the guidelines issued by the Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche and were approved by the Direction Départementale des Services Vétérinaires de l'Essonne and the Ethical Committee N°59.

Translation blocking morpholinos (Gene Tool LLC Philomath) were designed against zebrafish *CYFIP1* (AAAAACTATCCGCTTCGACTGTTCA) and *CYFIP2* (CGACACAGGTTCACTCACAAAACAG). The *NCKAP1* morpholino (CCGAGACATGGCTCAAACGACCGTC) was described in [58]. The control morpholino is a standard control (CCTCTTACCTCAGTTACAATTTATA). mRNAs were synthesized using pCS2+ plasmids containing the human genes described in [32] and the mMessage mMachine SP6 kit (Thermo Fischer).

For cell migration quantification, embryos were injected at the one-cell stage with 1.5 nl of a solution containing Histone2B-mCherry mRNA (30 ng/µl) and either control morpholino (0.1, 0.2 or 0.8mM), MoCYFIP1 (0.2mM), MoCYFIP2 (0.1mM) or MoNCKAP1 (0.8mM), with or

without mRNAs encoding either human CYFIP1 (10ng/µl), human CYFIP2 (10ng/µl) or human NCKAP1 (10ng/µl). Injected embryos were mounted in 0.2% agarose in embryo medium and imaged between 60% and 80% epiboly (6.5-8.5 hpf) under an upright TriM Scope II (La Vision Biotech) two photon microscope equipped with an environmental chamber (okolab) at 28°C using a 25x water immersion objective. Visualization of 3D movies and nuclei tracking were done using Imaris (Bitplane). Cell migration parameters were extracted using custom Matlab (Math Works) code and autocorrelation was computed using published Excel macros [56].

For protrusion analysis, embryos were injected in one cell at the four-cell stage with 1.5 nl of a solution containing Lifeact-mCherry mRNA (50 ng/µl) and either control morpholino (0.5 mM), MoCYFIP1 (0.2mM), MoCYFIP2 (0.1mM) or MoNCKAP1 (0.8mM), with or without mRNAs encoding either human CYFIP1 (10ng/µl), human CYFIP2 (10ng/µl) or human NCKAP1 (10ng/µl). Small cell groups were transplanted at shield stage (6 hpf) from the shield of an injected embryo to the shield of an untreated host. Embryos were then cultured in embryo medium [59] with 10 U/mL penicillin and 10 µg/mL streptomycin. Transplanted embryos were mounted in 0.2% agarose in embryo medium and imaged between 60% and 80% epiboly (6.5-8.5 hpf) under an inverted TCS SP8 confocal microscope equipped with environmental chamber (Leica) at 28°C using a HC PL APO 40x/1.10 W CS2 objective. Visualization of images was done on ImageJ, lamellipodia-like actin rich protrusions being quantified on the basis of morphological criteria as described in [26].

Reconstitution of WAVE complexes and in vitro assays

Recombinant WAVE complexes containing full-length human CYFIP1 or CYFIP2, full-length NCKAP1, full-length BRK1, ABI2 (1-158) and WAVE1 (1-230)-$(GGS)_6$-WCA (485-559), referred to as WRC230WCA were purified as previously described [35,36]. CYFIP1- and CYFIP2-containing WAVE complexes behaved similarly during expression and purification by various chromatographic steps. Other proteins, including the Arp2/3 complex, actin, WAVE1 WCA, Tev, GST-RAC1 (Q61L P29S, 1-188), and untagged RAC1 (Q61L P29S, 1-188) were purified as previously described [35].

GST pull-down experiments were performed as previously described [35]. Briefly, 200 pmol of GST-RAC1 and 200 pmol of WAVE complex were mixed with 20 µL of Glutathione Sepharose beads (GE Healthcare) in 1 mL of binding buffer (10 mM HEPES pH 7, 50 or 100 mM NaCl, 5% (w/v) glycerol, 2 mM $MgCl_2$, 1 mM DTT, and 0.05% Triton X100) at 4 °C for 30 min, followed by three washes using 1 mL of the binding buffer in each wash. Finally, the bound proteins were eluted with GST elution buffer (100 mM Tris-HCl pH 8.5, 30 mM reduced glutathione, and 2 mM $MgCl_2$) and examined on SDS-PAGE gels.

GST equilibrium pull-down assays were performed in the EPD buffer (10 mM HEPES pH 7, 50 mM NaCl, 5% (w/v) glycerol, 2 mM $MgCl_2$, and 1 mM DTT) as previous described (Chen et al., 2017). Essentially, each 100 µL of reaction contained 0.1 µM WRC230WCA, varying concentrations of GST-Rac1(Q61L P29S, 1-188), 30 µL of the Glutathione Sepharose beads, and 0.05% Triton X100. All protein samples and beads were first dialyzed or equilibrated in the EPD buffer prior to use. After gentle mixing at 4°C for 30 min, the beads were pelleted by a brief centrifugation, and the supernatant was immediately transferred to SDS loading buffer

and analyzed by Coomassie blue-stained SDS-PAGE gels. Total intensity of the CYFIP1/2 and NCKAP1 bands was quantified by ImageJ to determine the unbound WAVE complex. The derived fractional occupancy from several independent experiments was pooled and globally fitted to obtain the binding isotherms and the apparent dissociation constants K_D.

Actin polymerization assays were performed as previously described [35] with slight modifications. Each reaction (120 µL) contained 4 µM actin (5% pyrene labeled), 10 nM Arp2/3 complex, 100 nM WRC230WCA or WAVE1 WCA, and desired concentration of untagged RAC1 (Q61L P29S, 1-188) in NMEH20GD buffer (50 mM NaCl, 1 mM $MgCl_2$, 1 mM EGTA, 10 mM HEPES pH7.0, 20% (w/v) glycerol, and 1 mM DTT). Pyrene-actin fluorescence was recorded every 5 seconds at 22 °C using a 96-well flat-bottom black plate (Greiner Bio-One™) in a Spark plater reader (Tecan), with excitation at 365 nm and emission at 407 nm (15 nm bandwidth for both wavelengths).

Statistical analyses

Patient cohort. Relationships with mRNA levels and clinical parameters were identified using the χ^2 test. Statistical analyses using univariate and multivariate Cox proportional hazard models were performed with the R computing environment (R Development Core Team, 2017). Codes are available upon request.

Migration persistence. Exponential decay and plateau fit ($y = (1 - b) * e^{-\frac{t}{a}} + b$) was performed for all individual cells. Coefficients were then compared using one-way ANOVA. Statistical analysis was performed in R using linear mixed-effect models to take into account the resampling of the same statistical unit.

Significance. Differences were considered significant at confidence levels greater than 95% ($p < 0.05$). Four levels of statistical significance were distinguished: *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

Acknowledgements

We thank Nelia Cordeiro for technical support, Theresia Stradal for sharing shRNA plasmids targeting NCKAP1 and CYFIP2, Klemens Rottner for the *CYFIP1/2* DKO cells, and Chuang Yu for his help with statistical tools. This work was supported by grants from the Agence Nationale de la Recherche (ANR-15-CE13-0016-01 for AMG and NBD), from the fondation ARC pour la Recherche sur le Cancer (PGA120140200831 for AMG and IB and DOC20170505494 to NM), from National Institute of Health (grant R35 GM128786) to BC, from Institut National du Cancer (INCA_6521 for AMG and IB, and INCA_11508 for AMG, IB and ML). We thank the Polytechnique Bioimaging Facility for assistance with live imaging on their equipment partly supported by Région Ile-de-France (interDIM) and Agence Nationale de la Recherche (ANR-11-EQPX-0029 Morphoscope2, ANR-10-INBS-04 France BioImaging)

Conflict of interest

The authors declare no conflict of interest

Author contributions

AP performed *in vitro* experiments of cell migration and wrote the manuscript. AB and NBD performed *in vivo* experiments in zebrafish embryos. ML performed statistical modeling. SV, AS and IB performed qRT-PCR measurements of WAVE complex expression. NM constructed the integrative plasmid used to overexpress CYFIP2. NR quantified CYFIP1/2 protein expression in cell lines. AF and YW isolated the CYFIP2 knockout clones and performed experiments with these clones. YW performed the sucrose gradient experiments. RG performed the structural comparison of CYFIP1/2 proteins. SR performed the confocal microscopy of B16-F1 cells and helped with the data analysis. SY, YL, and BC reconstituted CYFIP1/2 containing WAVE complexes and performed RAC1 binding and pyrene actin assays. AMG supervised the study and wrote the manuscript. All authors have commented on the manuscript and approved the submission.

REFERENCES

1. Dehal, P., and Boore, J.L. (2005). Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. PLoS Biology *3*, e314-9.

2. Ruepp, A., Waegele, B., Lechner, M., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Mewes, H.W. (2010). CORUM: the comprehensive resource of mammalian protein complexes--2009. Nucleic Acids Research *38*, D497-501.

3. Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell *92*, 291–294.

4. Molinie, N., and Gautreau, A. (2018). The Arp2/3 Regulatory System and Its Deregulation in Cancer. Physiological Reviews *98*, 215–238.

5. Steffen, A., Rottner, K., Ehinger, J., Innocenti, M., Scita, G., Wehland, J., and Stradal, T.E.B. (2004). Sra-1 and Nap1 link Rac to actin assembly driving lamellipodia formation. The EMBO Journal *23*, 749–759.

6. Lai, F.P.L., Szczodrak, M., Block, J., Faix, J., Breitsprecher, D., Mannherz, H.G., Stradal, T.E.B., Dunn, G.A., Small, J.V., and Rottner, K. (2008). Arp2/3 complex interactions and actin network turnover in lamellipodia. The EMBO Journal *27*, 982–992.

7. Steffen, A., Koestler, S.A., and Rottner, K. (2014). Requirements for and consequences of Rac-dependent protrusion. European Journal of Cell Biology *93*, 184–193.

8. Krause, M., and Gautreau, A. (2014). Steering cell migration: lamellipodium dynamics and the regulation of directional persistence. Nature Reviews Mol Cell Biol *15*, 577–590.

9. Derivery, E., and Gautreau, A. (2010). Generation of branched actin networks: assembly and regulation of the N-WASP and WAVE molecular machines. BioEssays : news and reviews in molecular, cellular and developmental biology *32*, 119–131.

10. Dubielecka, P.M., Cui, P., Xiong, X., Hossain, S., Heck, S., Angelov, L., and Kotula, L. (2010). Differential Regulation of Macropinocytosis by Abi1/Hssh3bp1 Isoforms. PLoS ONE *5*, e10430.

11. Suetsugu, S., Yamazaki, D., Kurisu, S., and Takenawa, T. (2003). Differential Roles of WAVE1 and WAVE2 in Dorsal and Peripheral Ruffle Formation for Fibroblast Cell Migration. Dev Cell *5*, 595–609.

12. Iwaya, K., Norio, K., and Mukai, K. (2007). Coexpression of Arp2 and WAVE2 predicts poor outcome in invasive breast carcinoma. Modern Pathology *20*, 339–343.

13. Zhang, J., Tang, L., Shen, L., Zhou, S., Duan, Z., Xiao, L., Cao, Y., Mu, X., Zha, L., and Wang, H. (2012). High level of WAVE1 expression is associated with tumor aggressiveness and unfavorable prognosis of epithelial ovarian cancer. Gynecologic oncology *127*, 223–230.

14. Semba, S., Iwaya, K., Matsubayashi, J., Serizawa, H., Kataba, H., Hirano, T., Kato, H., Matsuoka, T., and Mukai, K. (2006). Coexpression of actin-related protein 2 and Wiskott-Aldrich syndrome family verproline-homologous protein 2 in adenocarcinoma of the lung. Clinical cancer research : an official journal of the American Association for Cancer Research *12*, 2449–2454.

15. Yang, L.Y., Tao, Y.M., Ou, D.P., Wang, W., Chang, Z.G., and Wu, F. (2006). Increased Expression of Wiskott-Aldrich Syndrome Protein Family Verprolin-Homologous Protein 2 Correlated with Poor Prognosis of Hepatocellular Carcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research *12*, 5673–5679.

16. Ji, Y., Li, B., Zhu, Z., Guo, X., He, W., Fan, Z., and Zhang, W. (2015). Overexpression of WAVE3 promotes tumor invasiveness and confers an unfavorable prognosis in human hepatocellular carcinoma. Biomedicine & pharmacotherapy = Biomédecine & pharmacothérapie *69*, 409–415.

17. Zhang, Y., Guan, X.-Y., Dong, B., Zhao, M., Wu, J.-H., Tian, X.-Y., and Hao, C.-Y. (2012). Expression of MMP-9 and WAVE3 in colorectal cancer and its relationship to clinicopathological features. Journal of Cancer Research and Clinical Oncology *138*, 2035–2044.

18. Lomakina, M.E., Lallemand, F., Vacher, S., Molinie, N., Dang, I., Cacheux, W., Chipysheva, T.A., Ermilova, V.D., Koning, L.D., Dubois, T., et al. (2016). Arpin downregulation in breast cancer is associated with poor prognosis. British journal of cancer *114*, 545–553.

19. Joshi, A.D., Hegde, G.V., Dickinson, J.D., Mittal, A.K., Lynch, J.C., Eudy, J.D., Armitage, J.O., Bierman, P.J., Bociek, R.G., Devetten, M.P., et al. (2007). ATM, CTLA4, MNDA, and HEM1 in High versus Low CD38-Expressing B-Cell Chronic Lymphocytic Leukemia. Clinical cancer research : an official journal of the American Association for Cancer Research *13*, 5295–5304.

20. Wang, C., Tran-Thanh, D., Moreno, J.C., Cawthorn, T.R., Jacks, L.M., Wang, D.-Y., McCready, D.R., and Done, S.J. (2010). Expression of Abl interactor 1 and its prognostic significance in breast cancer: a tissue-array-based investigation. Breast Cancer Research and Treatment *129*, 373–386.

21. Zhang, J., Tang, L., Chen, Y., Duan, Z., Xiao, L., Li, W., Liu, X., and Shen, L. (2015). Upregulation of Abelson interactor protein 1 predicts tumor progression and poor outcome in epithelial ovarian cancer. Human pathology *46*, 1331–1340.

22. Győrffy, B., Lanczky, A., Eklund, A.C., Denkert, C., Budczies, J., Li, Q., and Szallasi, Z. (2009). An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Research and Treatment *123*, 725–731.

23. Molinie, N., Rubtsova, S.N., Fokin, A., Visweshwaran, S.P., Rocques, N., Polesskaya, A., Schnitzler, A., Vacher, S., Denisov, E.V., Tashireva, L.A., et al. (2019). Cortical branched actin determines cell cycle progression. Cell Research *29*, 432–445.

24. Montero, J.-A., Kilian, B., Chan, J., Bayliss, P.E., and Heisenberg, C.-P. (2003). Phosphoinositide 3-kinase is required for process outgrowth and cell polarization of gastrulating mesendodermal cells. Current Biology *13*, 1279–1289.

25. Dumortier, J.G., Martin, S., Meyer, D., Rosa, F.M., and David, N.B. (2012). Collective mesendoderm migration relies on an intrinsic directionality signal transmitted through cell contacts. Proceedings of the National Academy of Sciences of the United States of America *109*, 16945–16950.

26. Diz-Muñoz, A., Krieg, M., Bergert, M., Ibarlucea-Benitez, I., Muller, D.J., Paluch, E., and Heisenberg, C.-P. (2010). Control of Directed Cell Migration In Vivo by Membrane-to-Cortex Attachment. PLoS Biology *8*, e1000544-12.

27. Petrie, R.J., Gavara, N., Chadwick, R.S., and Yamada, K.M. (2012). Nonpolarized signaling reveals two distinct modes of 3D cell migration. The Journal of Cell Biology *197*, 439–455.

28. Schaks, M., Singh, S.P., Kage, F., Thomason, P., Klünemann, T., Steffen, A., Blankenfeldt, W., Stradal, T.E., Insall, R.H., and Rottner, K. (2018). Distinct Interaction Sites of Rac GTPase with WAVE Regulatory Complex Have Non-redundant Functions in Vivo. Current Biology *28*, 3674-3684.e6.

29. Zweier, M., Begemann, A., McWalter, K., Cho, M.T., Abela, L., Banka, S., Behring, B., Berger, A., Brown, C.W., Carneiro, M., et al. (2019). Spatially clustering de novo variants in CYFIP2, encoding the cytoplasmic FMRP interacting protein 2, cause intellectual disability and seizures. Eur J Hum Genet *27*, 747–759.

30. Chen, Z., Borek, D., Padrick, S.B., Gomez, T.S., Metlagel, Z., Ismail, A.M., Umetani, J., Billadeau, D.D., Otwinowski, Z., and Rosen, M.K. (2010). Structure and control of the actin regulatory WAVE complex. Nature *468*, 533–538.

31. Chen, B., Brinkmann, K., Chen, Z., Pak, C.W., Liao, Y., Shi, S., Henry, L., Grishin, N.V., Bogdan, S., and Rosen, M.K. (2014). The WAVE regulatory complex links diverse receptors to the actin cytoskeleton. Cell *156*, 195–207.

32. Gautreau, A., Ho, H.H., Li, J., Steen, H., Gygi, S.P., and Kirschner, M.W. (2004). Purification and architecture of the ubiquitous Wave complex. Proceedings of the National Academy of Sciences of the United States of America *101*, 4379–4383.

33. Fort, L., Batista, J.M., Thomason, P.A., Spence, H.J., Whitelaw, J.A., Tweedy, L., Greaves, J., Martin, K.J., Anderson, K.I., Brown, P., et al. (2018). Fam49/CYRI interacts with Rac1 and locally suppresses protrusions. Nature Cell Biology *20*, 1159–1171.

34. Yelland, T., Le, A.H., Nikolaou, S., Insall, R., Machesky, L., and Ismail, S. (2021). Structural Basis of CYRI-B Direct Competition with Scar/WAVE Complex for Rac1. Structure *29*, 226-237.e4.

35. Chen, B., Chou, H.-T., Brautigam, C.A., Xing, W., Yang, S., Henry, L., Doolittle, L.K., Walz, T., and Rosen, M.K. (2017). Rac1 GTPase activates the WAVE regulatory complex through two distinct binding sites. eLife *6*, W529.

36. Chen, B., Padrick, S.B., Henry, L., and Rosen, M.K. (2014). Biochemical reconstitution of the WAVE regulatory complex. Methods in Enzymology *540*, 55–72.

37. Nakashima, M., Kato, M., Aoto, K., Shiina, M., Belal, H., Mukaida, S., Kumada, S., Sato, A., Zerem, A., Sagie, T.L., et al. (2018). De novo hotspot variants in CYFIP2 cause early-onset epileptic encephalopathy. Annals of neurology *83*, 794–806.

38. Zweier, M., Begemann, A., McWalter, K., Cho, M.T., Abela, L., Banka, S., Behring, B., Berger, A., Brown, C.W., Carneiro, M., et al. (2019). Spatially clustering de novo variants in CYFIP2, encoding the cytoplasmic FMRP interacting protein 2, cause intellectual disability and seizures. European journal of human genetics : EJHG, 1–13.

39. Pittman, A.J., Gaynes, J.A., and Chien, C.-B. (2010). nev (cyfip2) is required for retinal lamination and axon guidance in the zebrafish retinotectal system. Developmental Biology *344*, 784–794.

40. Cioni, J.-M., Wong, H.H.-W., Bressan, D., Kodama, L., Harris, W.A., and Holt, C.E. (2018). Axon-Axon Interactions Regulate Topographic Optic Tract Sorting via CYFIP2-Dependent WAVE Complex Function. Neuron *97*, 1078-1093.e6.

41. Blagg, S.L., Stewart, M., Sambles, C., and Insall, R.H. (2003). PIR121 Regulates Pseudopod Dynamics and SCAR Activity in Dictyostelium. Current Biology *13*, 1480–1487.

42. Ibarra, N., Blagg, S.L., Vazquez, F., and Insall, R.H. (2006). Nap1 Regulates Dictyostelium Cell Motility and Adhesion through SCAR-Dependent and -Independent Pathways. Current Biology *16*, 717–722.

43. Pollitt, A.Y., and Insall, R.H. (2008). Abi Mutants in Dictyostelium Reveal Specific Roles for the SCAR/WAVE Complex in Cytokinesis. Current Biology *18*, 203–210.

44. Litschko, C., Linkner, J., Brühmann, S., Stradal, T.E.B., Reinl, T., Jänsch, L., Rottner, K., and Faix, J. (2017). Differential functions of WAVE regulatory complex subunits in the regulation of actin-driven processes. European Journal of Cell Biology *96*, 715–727.

45. Korobova, F., and Svitkina, T. (2008). Arp2/3 complex is important for filopodia formation, growth cone motility, and neuritogenesis in neuronal cells. Molecular Biology of the Cell *19*, 1561–1574.

46. Dang, I., Gorelik, R., Sousa-Blin, C., Derivery, E., Guérin, C., Linkner, J., Nemethova, M., Dumortier, J.G., Giger, F.A., Chipysheva, T.A., et al. (2013). Inhibitory signalling to the Arp2/3 complex steers cell migration. Nature *503*, 281–284.

47. Stradal, T., Courtney, K.D., Rottner, K., Hahne, P., Small, J.V., and Pendergast, A.M. (2001). The Abl interactor proteins localize to sites of actin polymerization at the tips of lamellipodia and filopodia. Curr Biol *11*, 891–895.

48. Kumar, V., Kim, K., Joseph, C., Kourrich, S., Yoo, S.-H., Huang, H.C., Vitaterna, M.H., Villena, F.P.-M. de, Churchill, G., Bonci, A., et al. (2013). C57BL/6N mutation in cytoplasmic FMRP interacting protein 2 regulates cocaine response. Science *342*, 1508–1512.

49. Wan, C., Borgeson, B., Phanse, S., Tu, F., Drew, K., Clark, G., Xiong, X., Kagan, O., Kwan, J., Bezginov, A., et al. (2015). Panorama of ancient metazoan macromolecular complexes. Nature *525*, 339–344.

50. Abella, J.V.G., Galloni, C., Pernier, J., Barry, D.J., Kjær, S., Carlier, M.-F., and Way, M. (2016). Isoform diversity in the Arp2/3 complex determines actin filament dynamics. Nature Cell Biology *18*, 76–86.

51. Bieche, I., Parfait, B., Laurendeau, I., Girault, I., Vidaud, M., and Lidereau, R. (2001). Quantification of estrogen receptor alpha and beta expression in sporadic breast cancer. Oncogene *20*, 8109–8115.

52. González, F., Zhu, Z., Shi, Z.-D., Lelli, K., Verma, N., Li, Q.V., and Huangfu, D. (2014). An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell stem cell *15*, 215–226.

53. Derivery, E., Lombard, Be., Loew, D., and Gautreau, A. (2009). The Wave complex is intrinsically inactive. Cell Motility and the Cytoskeleton *66*, 777–790.

54. Derivery, E., and Gautreau, A. (2010). Assaying WAVE and WASH complex constitutive activities toward the Arp2/3 complex. Methods in Enzymology *484*, 677–695.

55. Derivery, E., Fink, J., Martin, D., Houdusse, A., Piel, M., Stradal, T.E., Louvard, D., and Gautreau, A. (2008). Free Brick1 is a trimeric precursor in the assembly of a functional wave complex. PLoS ONE *3*, e2462.

56. Gorelik, R., and Gautreau, A. (2014). Quantitative and unbiased analysis of directional persistence in cell migration. Nature Protocols *9*, 1931–1943.

57. Doitsidou, M., Reichman-Fried, M., Stebler, J., Köprunner, M., Dörries, J., Meyer, D., Esguerra, C.V., Leung, T., and Raz, E. (2002). Guidance of primordial germ cell migration by the chemokine SDF-1. Cell *111*, 647–659.

58. Biswas, S., Emond, M.R., and Jontes, J.D. (2010). Protocadherin-19 and N-cadherin interact to control cell movements during anterior neurulation. The Journal of Cell Biology *191*, 1029–1041.

59. Hans, S., Christison, J., Liu, D., and Westerfield, M. (2007). Fgf-dependent otic induction requires competence provided by Foxi1 and Dlx3b. BMC developmental biology *7*, 5.

Table 1: Tumours showing up- or down-regulation of WAVE complex subunits in breast cancer (> 3 or < 0.3)

Generic name	Gene name	All tumours		HR+ ERBB2-		HR+ ERBB2+		HR- ERBB2-		HR- ERBB2+	
		% down	% up	% down	% up	% down	% up	% down	% up	% down	% up
WAVE	***WASF1***	**27**	2	**37**	1	10	2	10	3	21	6
	WASF2	3	1	3	0	3	2	2	3	1	0
	WASF3	**46**	1	**56**	0	48	0	27	2	25	3
ABI	***ABI1***	0	1	0	0	0	2	0	3	0	1
	ABI2	3	0	1	0	2	0	7	0	3	0
	ABI3	1	**12**	1	5	0	17	0	**25**	3	17
BRK	***BRK1***	0	4	0	3	0	10	0	4	0	3
NAP	***NCKAP1***	0	2	0	1	0	2	0	6	0	4
	NCKAP1L	0	**22**	0	15	0	12	0	37	0	**39**
CYFIP	***CYFIP1***	0	4	0	1	0	12	0	6	0	7
	CYFIP2	1	**37**	1	**44**	0	36	0	30	1	17

Figures

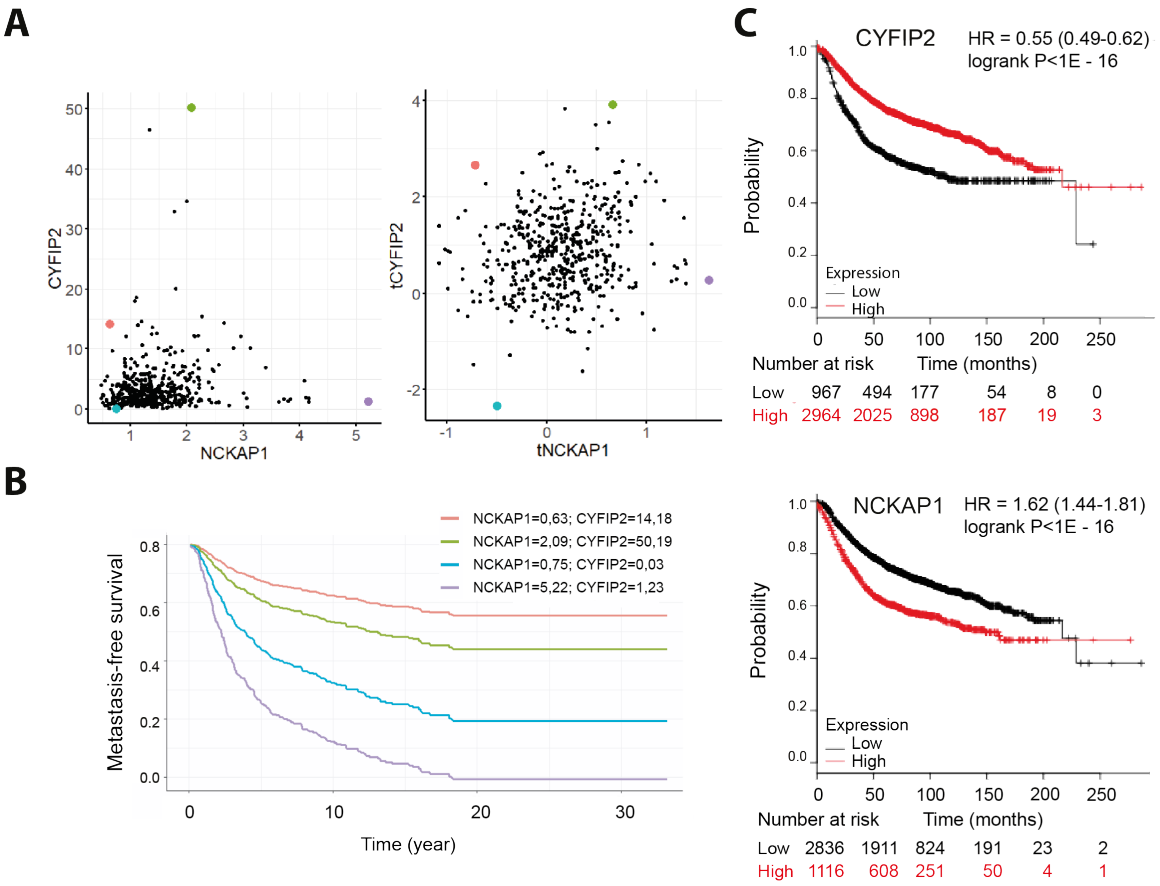


Fig.1, Polesskaya et al.

Figure 1. CYFIP2 overexpression is associated with good prognosis in breast cancer patients.

(A) Distribution of *NCKAP1* and *CYFIP2* mRNA levels in mammary carcinomas from a cohort of 527 breast cancer patients, before (left panel) or after transformation and normalization (right panel) **(B)** A multivariate Cox model predicting metastasis-free survival (MFS) based on *NCKAP1* and *CYFIP2* mRNA levels as the only two inputs was derived. The 4 highlighted tumors representing the different outskirts of gene expression in the cohort were chosen to run the model. The purple and turquoise patients developed metastases that were diagnosed after 922 and 1487 days, respectively. The red and green patients did not develop metastasis and survived for 4248 and 4146 days, respectively. Even though extreme *NCKAP1* values drive MFS in the red and purple patients, the extreme values of *CYFIP2* rule the outcome of the green and turquoise patients at intermediate values of *NCKAP1*. The model thus predicts that high levels of *NCKAP1* are associated with poor prognosis, whereas high levels of *CYFIP2* are associated with good prognosis. **(C)** Validation of the prediction using a public database, kmplot.com, containing more than 3900 breast cancer patients. Kaplan-Meier representations.

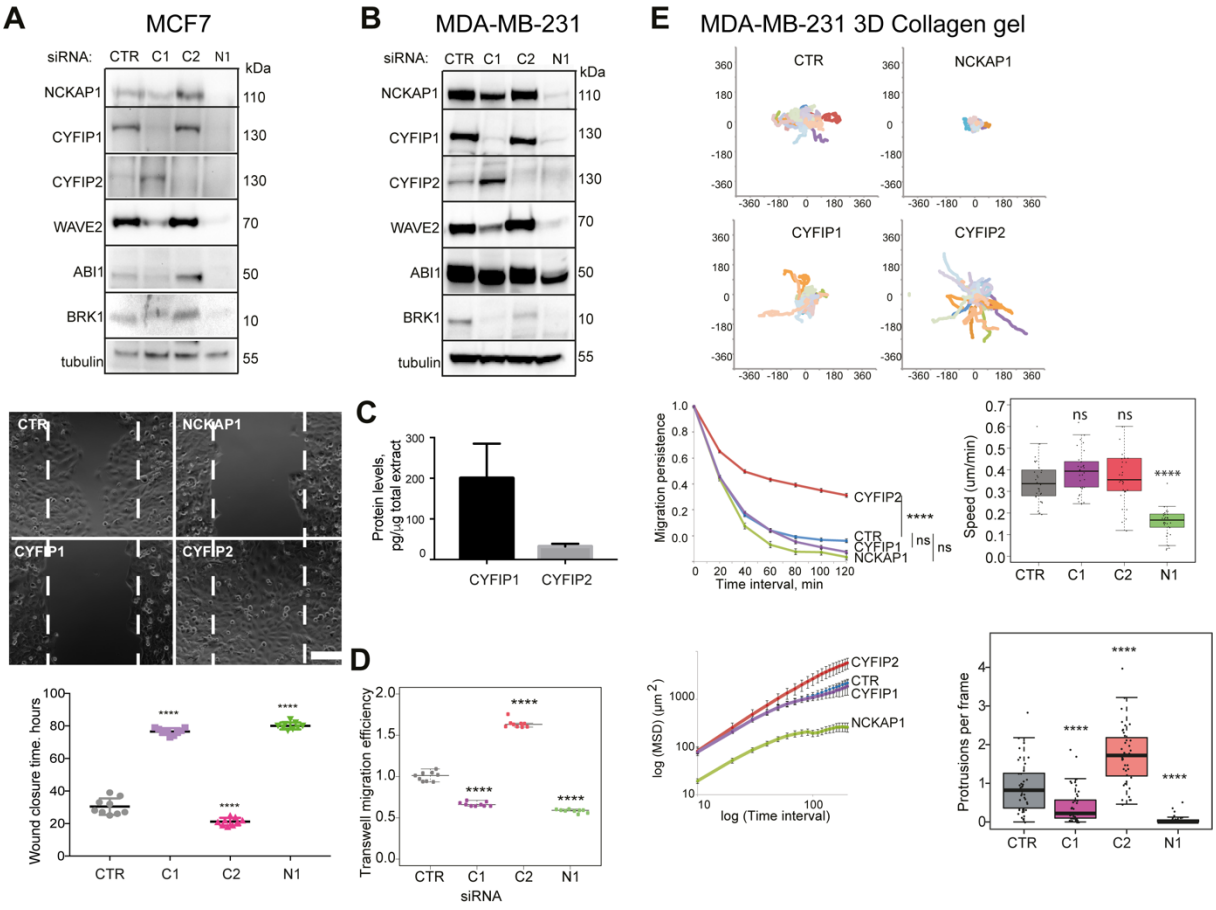


Fig.2, Polesskaya et al.

Figure 2. CYFIP2 inhibits the migration of human breast cancer cells. (**A**) MCF7 cells were transfected with pools of siRNAs targeting CYFIP1 (C1), CYFIP2 (C2), NCKAP1 (N1) or non-targeting ones (CTR). Western blots of WAVE complex subunits and tubulin as a loading control. Wound healing of MCF7 cells. Still images corresponding to the time, where the first wound is healed (CYFIP2). Quantification of nine technical repeats. Scale bar: 400 μm. (**B**) MDA-MB-231 cells were transfected with pools of siRNAs and analyzed by Western blots as above. (**C**) Levels of CYFIP1 and CYFIP2 proteins in MDA-MB-231. Mean ± SD of 3 biological repeats. (**D**) Quantification of Transwell migration efficiency of MDA-MB-231 cells, n=9. (**E**) Depleted MDA-MB-231 cells depleted of the indicated proteins were embedded in 3D collagen type I gels and recorded by videomicroscopy. Trajectories, migration persistence, Mean Square Displacement (MSD), and average number of protrusions per frame are plotted, n=30. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

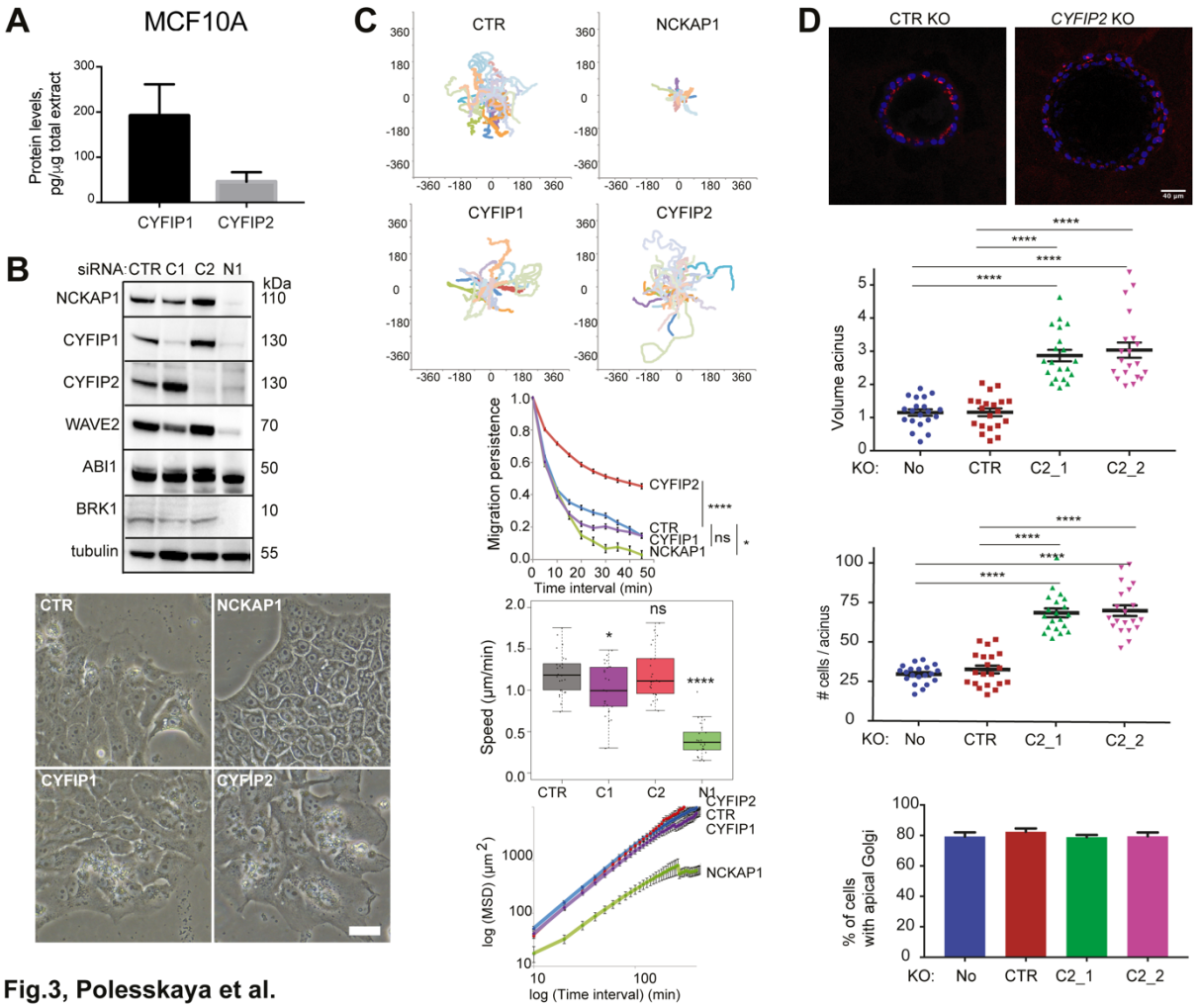


Fig.3, Polesskaya et al.

Figure 3. CYFIP2 inhibits the migration of normal breast epithelial cells and the growth of 3D acini. (**A**) Levels of CYFIP1 and CYFIP2 proteins in MDA-MB-231. Mean ± SD of 3 biological repeats. (**B**) MCF10A cells were transfected with pools of CYFIP1 (C1), CYFIP2 (C2), NCKAP1 (N1) or non-targeting siRNAs (CTR). WAVE complex subunits and tubulin as a loading control were analyzed by Western blot. Phase-contrast images of depleted cells. Scale bar: 50 µm. **(C)** Trajectories, migration persistence, speed, MSD extracted from random migration of single MCF10A cells. 2D migration, Fibronectin coating, n=25. (**D**) *CYFIP2* KO cells or parental MCF10A cells were differentiated at the surface of matrigel. Confocal microscopy of acini labeled with DAPI (blue) and the Golgi marker GM130 (red). Scale bar: 40 µm. Quantification of acinus volume and the number of cells per acinus, n=20. (**H**), Quantification of cells' polarity within acini, n=130. *P<0.05; ****P<0.0001.

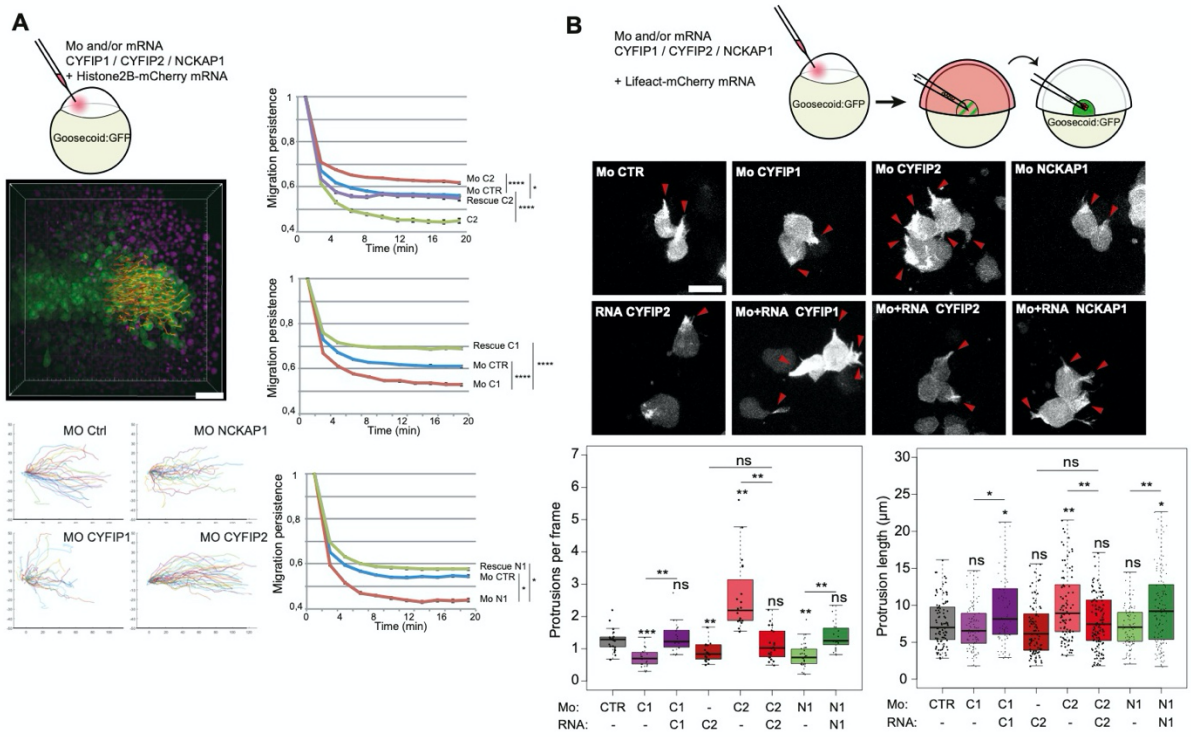


Figure 4. CYFIP2 inhibits migration persistence and actin rich protrusions in zebrafish embryos during gastrulation. (A) Scheme of the experimental design. Embryos were injected with Histone2B-mCherry mRNA and morpholinos (Mo) targeting a control sequence (CTR), CYFIP1 (C1), CYFIP2 (C2), NCKAP1 (N1), alone or in combination with mRNAs encoding the same proteins (rescue). Dorsal view of a volume acquisition of a Tg(Gsc:GFP) zebrafish embryo. Scale bar is 50 µm. Animal pole is located at the right. Notochord and prechordal plate cells express GFP (green) and nuclei express histone2B-mCherry (in magenta). Nuclei of prechordal plate cells are 3D-tracked over time (color coded) (Movie S5). Trajectories of 10 first time points (20 min) for 50 randomly selected cells for each condition, plotted at the same origin (axes in µm). Migration persistence of prechordal plate cells injected with the indicated MO and/or mRNA. (B) CYFIP2 inhibits actin rich protrusions in zebrafish embryos during gastrulation. Scheme of the experimental design. Donor embryos were injected with the actin filament marker LifeAct-mCherry mRNA and morpholinos (Mo) targeting a control sequence (CTR), CYFIP1 (C1), CYFIP2 (C2), NCKAP1 (N1), alone or in combination with mRNAs encoding the same proteins. Labeled prechordal plate cells from a donor embryo were transplanted into an uninjected embryo and recorded. Imaged of typical cells described in (A), red arrowheads indicate actin-rich protrusions. Scale bar: 20µm. Quantification of the average number of protrusions per frame, n=17 to 32 cells from 4 to 5 embryos per condition. Quantification of protrusion length, n=95 (randomly selected protrusions per condition). ANOVA on linear mixed model accounting for the sampling biases. ns $P>0.05$; * $P<0.05$; ** $P<0.01$; *** $P<0.001$. The p-values without a bar refer to comparisons with the control condition.

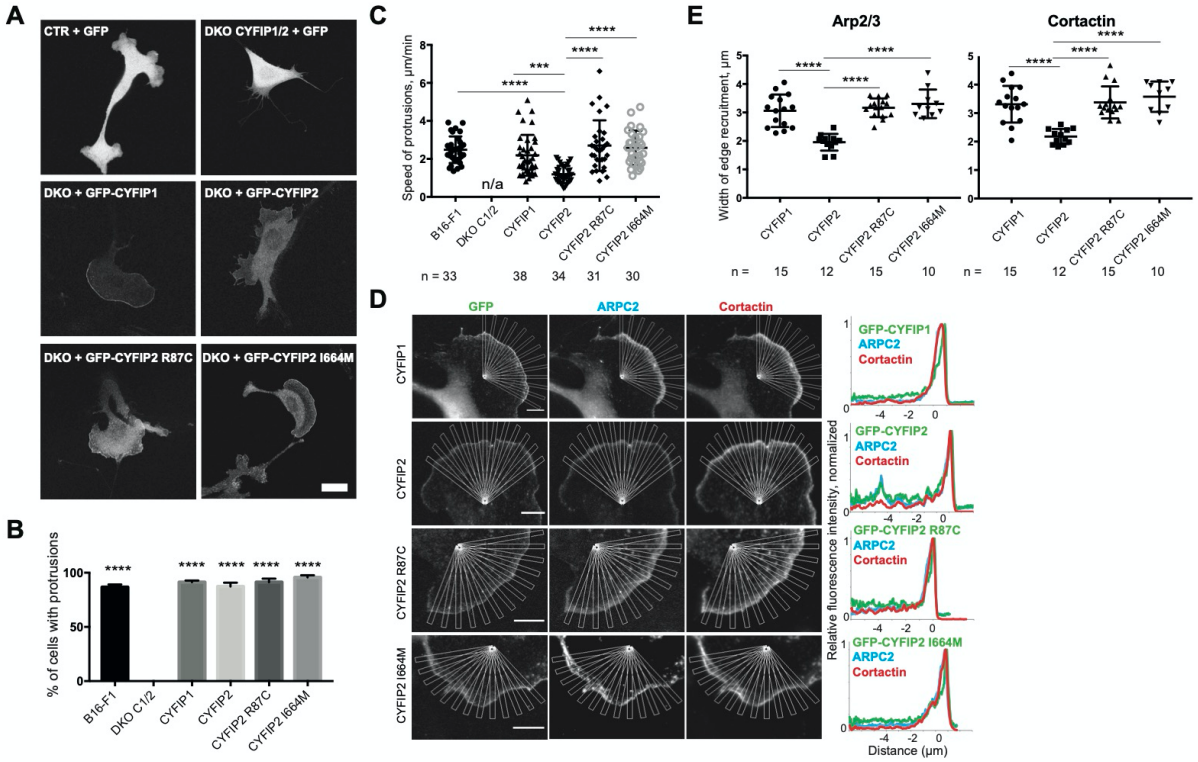


Fig.5, Polesskaya et al.

Figure 5. CYFIP2 rescues *CYFIP1/2* Double Knock-Out (DKO). GFP-tagged CYFIP1, CYFIP2, and two CYFIP2 mutants were expressed in B16-F1 control cells and in DKO cells. (A) Distribution of GFP fusion proteins and morphology of transfected cells. Scale bar: 20 µm. (B) Percentage of transfected cells forming protrusions, n=100, analysis by one-way ANOVA and Dunnett's multiple comparisons test. (C) Average speed of protrusions. Only the significant differences as determined by one-way ANOVA and Tukey's multiple comparisons test are indicated. (D) Recruitment of GFP-tagged CYFIP1, CYFIP2, and mutant CYFIP2 assessed by multiple radial line scans. Average profiles of the indicated markers upon registering line scans. Scale bars: 5µm. (E) Width of Arp2/3 and cortactin recruitment. ***P<0.001; ****P<0.0001.

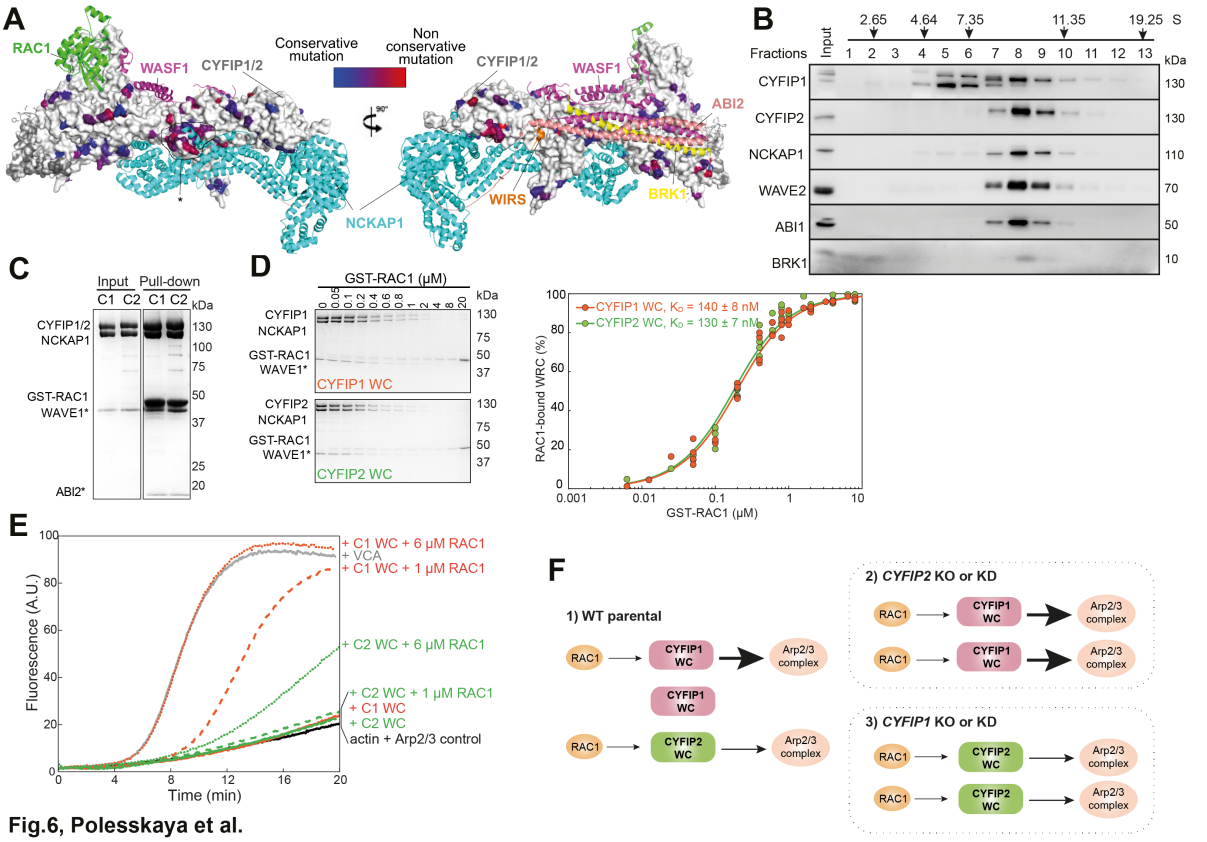


Fig.6, Polesskaya et al.

Figure 6. CYFIP2-containing WAVE complexes are less activatable by RAC1 than CYFIP1-containing WAVE complexes. (A) Structural models of CYFIP1 and CYFIP2. Sequence identity between CYFIP1 and CYFIP2 is 88% and non-conserved positions are colour-coded. The vast majority of non conserved residues fall outside of binding sites for known protein partners. WAVE complex subunits and WIRS peptide were obtained from PDB:4N78. RAC1 binding was modelled using the CYRI-B Rac1 complex as template (PDB:7AJK). (B) Ultracentrifugation of MCF10A lysate on a sucrose gradient. WAVE complex subunits are revealed by Western blots. The CYFIP1 antibody cross-reacts with a lower molecular weight band. (C) Coomassie-blue stained SDS-PAGE gels showing reconstitution of WAVE complexes containing CYFIP1 or CYFIP2 and pull-down with GTP-bound RAC1 (GST-RAC1 Q61L P29S). ABI2* and WAVE1* are not full length proteins (See Methods). (D) WAVE complexes containing supernatants upon pull-down with increasing amounts of GST-Rac1 Q61L P29S. Dissociation constants K_D and standard errors are derived from fitting of quantification of 4 independent experiments at various concentrations. (E) Pyrene-actin polymerization assay of CYFIP1- or CYFIP2-containing WAVE complexes. Conditions: 4 μM actin (5% pyrene-labeled), 10 nM Arp2/3 complex, 100 nM WAVE complexes (WC) or WAVE1 WCA, and indicated amounts of untagged Rac1 Q61L P29S. Results are representative of two independent experiments. (F) Model: CYFIP2-containing WAVE complexes activate less Arp2/3 upon RAC1 binding than CYFIP1-containing WAVE complexes. Upon depletion of CYFIP2, Arp2/3 activity increases, because more CYFIP1 containing complexes are activated by RAC1, leading to increased migration. On the opposite, upon depletion of CYFIP1, Arp2/3 activity decreases, despite compensatory increase in CYFIP2 levels, leading to reduced migration.

Table S1: Characteristics of the breast tumors relative to *CYFIP2* mRNA levels

	Number of patients (%)	**Number with metastases (%)**	***p-value*** [a]	***CYFIP2* mRNA normally expressed**	***CYFIP2* mRNA over expressed (> 3)**	***p-value*** [f]
Total	527 (100)	210 (39.8)		332 (63.0)	195 (37.0)	
Age						
<50	125 (23.7)	52 (41.6)	0.52 (NS)	82 (65.6)	43 (34.4)	0.49 (NS)
>50	402 (76.3)	158 (39.3)		250 (62.2)	152 (37.8)	
SBR histological grade [b,c]						
I	60 (11.7)	12 (20.0)	**0.0019**	34 (56.7)	26 (43.3)	**0.011**
II	241 (47.1)	100 (41.5)		141 (58.5)	100 (41.5)	
III	211 (41.2)	94 (44.5)		150 (71.1)	61 (28.9)	
Lymph node status [d]						
0	159 (30.5)	48 (30.2)	**<0.0000001**	96 (60.4)	63 (39.6)	0.66 (NS)
1-3	250 (47.9)	88 (35.2)		162 (64.8)	88 (35.2)	
>3	113 (21.6)	72 (63.7)		70 (61.9)	43 (38.1)	
Macroscopic tumor size [e]						
<25mm	248 (48.0)	77 (31.0)	**0.000015**	154 (62.1)	94 (37.9)	0.66 (NS)
>25mm	269 (52.0)	132 (49.1)		172 (63.9)	97 (36.1)	
ERα status						
Negative	181 (34.3)	76 (42.0)	0.10 (NS)	138 (76.2)	43 (23.8)	**0.0000052**
Positive	346 (65.7)	134 (38.7)		194 (56.1)	152 (43.9)	
PR status						
Negative	255 (48.4)	110 (43.1)	**0.025**	186 (72.9)	69 (27.1)	**0.0000047**
Positive	272 (51.6)	100 (36.8)		146 (53.7)	126 (46.3)	
ERBB2 status						
Negative	397 (75.3)	153 (38.5)	0.17 (NS)	235 (59.2)	162 (40.8)	**0.0016**
Positive	130 (24.7)	57 (43.8)		97 (74.6)	33 (25.4)	
Molecular subtypes						
HR- ERBB2-	102 (19.4)	38 (37.3)	0.054 (NS)	71 (69.6)	31 (30.4)	**0.00011**
HR- ERBB2+	72 (13.7)	36 (50.0)		60 (83.3)	12 (16.7)	
HR+ ERBB2-	295 (56.0)	115 (39.0)		164 (55.6)	131 (44.4)	
HR+ ERBB2+	58 (11.0)	21 (36.2)		37 (63.8)	21 (36.2)	

Supplemental figures

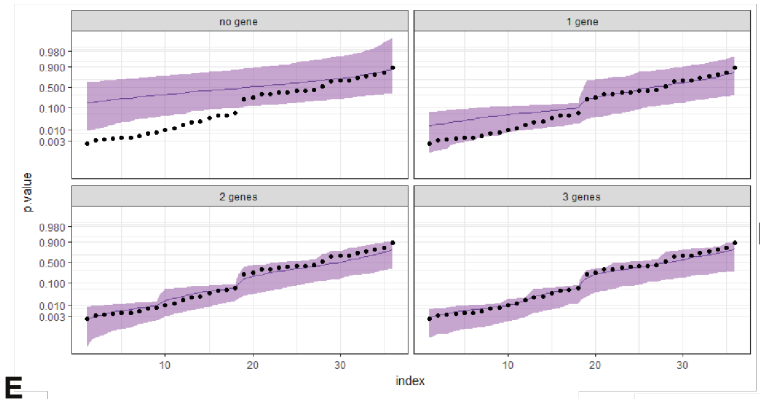
A

	Original values			Transformed and normalized values			
gene	coefficient	p-value	adjusted p-value	gene	coefficient	p-value	adjusted p-value
NCKAP1	0.27	0.012	0.132	*NCKAP1*	0.195	0.005	0.055
CYFIP2	-0.027	0.138	1	*CYFIP2*	-0.134	0.059	0.59
NCKAP1L	-0.051	0.288	1	*NCKAP1L*	-0.058	0.397	1
WASF2	-0.162	0.377	1	*WASF3*	-0.045	0.519	1
ABI3	-0.022	0.439	1	*ABI1*	0.042	0.546	1
WASF3	-0.066	0.573	1	*WASF2*	-0.041	0.561	1
ABI1	0.054	0.644	1	*ABI3*	-0.036	0.603	1
ABI2	-0.068	0.749	1	*BRK1*	0.035	0.609	1
CYFIP1	-0.01	0.866	1	*CYFIP1*	0.036	0.612	1
WASF1	0.009	0.930	1	*ABI2*	-0.034	0.624	1
BRK1	0.004	0.962	1	*WASF1*	-0.007	0.919	1

B

Subunit assembled					
NAP	CYFIP	ABI	WAVE	BRK	p-value
1	**2**	2	3	1	0,0023
1	**2**	2	2	1	0,0032
1	**2**	1	3	1	0,0036
1	**2**	3	3	1	0,0038
1	**2**	1	2	1	0,0044
1	**2**	3	2	1	0,0044
1	**2**	2	1	1	0,0050
1	**2**	1	1	1	0,0069
1	**2**	3	1	1	0,0076
1	1	2	2	1	0,0098
1	1	2	3	1	0,0115
1	1	2	1	1	0,0167
1	1	3	2	1	0,0220
1	1	1	2	1	0,0246
1	1	3	3	1	0,0352
1	1	1	3	1	0,0444
1	1	3	1	1	0,0470
1	1	1	1	1	0,0569
2	2	1	3	1	0,2212
2	2	1	2	1	0,2464
2	2	1	1	1	0,3183
2	2	3	3	1	0,3259
2	2	2	3	1	0,3649
2	2	3	2	1	0,3764
2	2	3	1	1	0,4103
2	2	2	2	1	0,4154
2	2	2	1	1	0,4287
2	1	1	2	1	0,5184
2	1	2	2	1	0,6576
2	1	2	3	1	0,6746
2	1	1	3	1	0,6847
2	1	3	2	1	0,7330
2	1	2	1	1	0,7767
2	1	3	3	1	0,7998
2	1	1	1	1	0,8247
2	1	3	1	1	0,8945

C



D

Genes	log.likelihood	BIC
-	-1234.7	2469.3
NCKAP1	-1230.7	2466.8
NCKAP1, CYFIP2	-1227.6	**2465.8**
NCKAP1, CYFIP2, WASF3	-1226.2	2468.4
NCKAP1, CYFIP2, WASF3, ABI2	-1225.4	2472.2
NCKAP1, CYFIP2, WASF3, ABI2, BRK1	-1224.9	2476.6

E

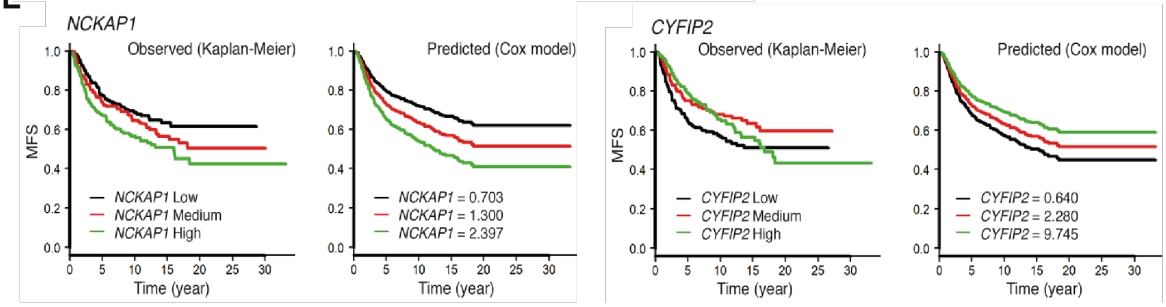


Figure S1. Statistical modeling of the association of WAVE complexes with metastasis-free survival (MFS). A Cox univariate analysis before and after transformation and normalization of mRNA levels of WAVE complex subunits. **B** Association of each WAVE assembly with MFS. **C** Random permutations. Experimental p values associated with the 36 WAVE assemblies (black dots) are compared with computed p-values corresponding to hypotheses (purple; 1000 simulations to derive the 90 % confidence interval): all subunits can be permuted (0 gene); all subunits except the most significant one, *NCKAP1*, can be permuted (1 gene); all subunits except the two most significant ones, *NCKAP1* and *CYFIP2*, can be permuted (2 genes); All subunits except the 3 most significant ones (3 genes) can be permuted. The small p-values obtained for the combination *NCKAP1* and *CYFIP2* are not obtained by chance, since computer simulations graphically illustrate the good agreement between what is observed and what is expected according to models when at least two of the most significant genes are fixed. **D** Comparison of the different models with Bayesian Information Criteria (BIC). The model with 2 fixed genes corresponds to the optimal statistical model, <hcih is used below and for Fig.1. **E** Kaplan-Meier of MFS as a function of NCKAP1 or CYFIP2 mRNA levels. Observed survival is modeled with the 2-variable Cox model.

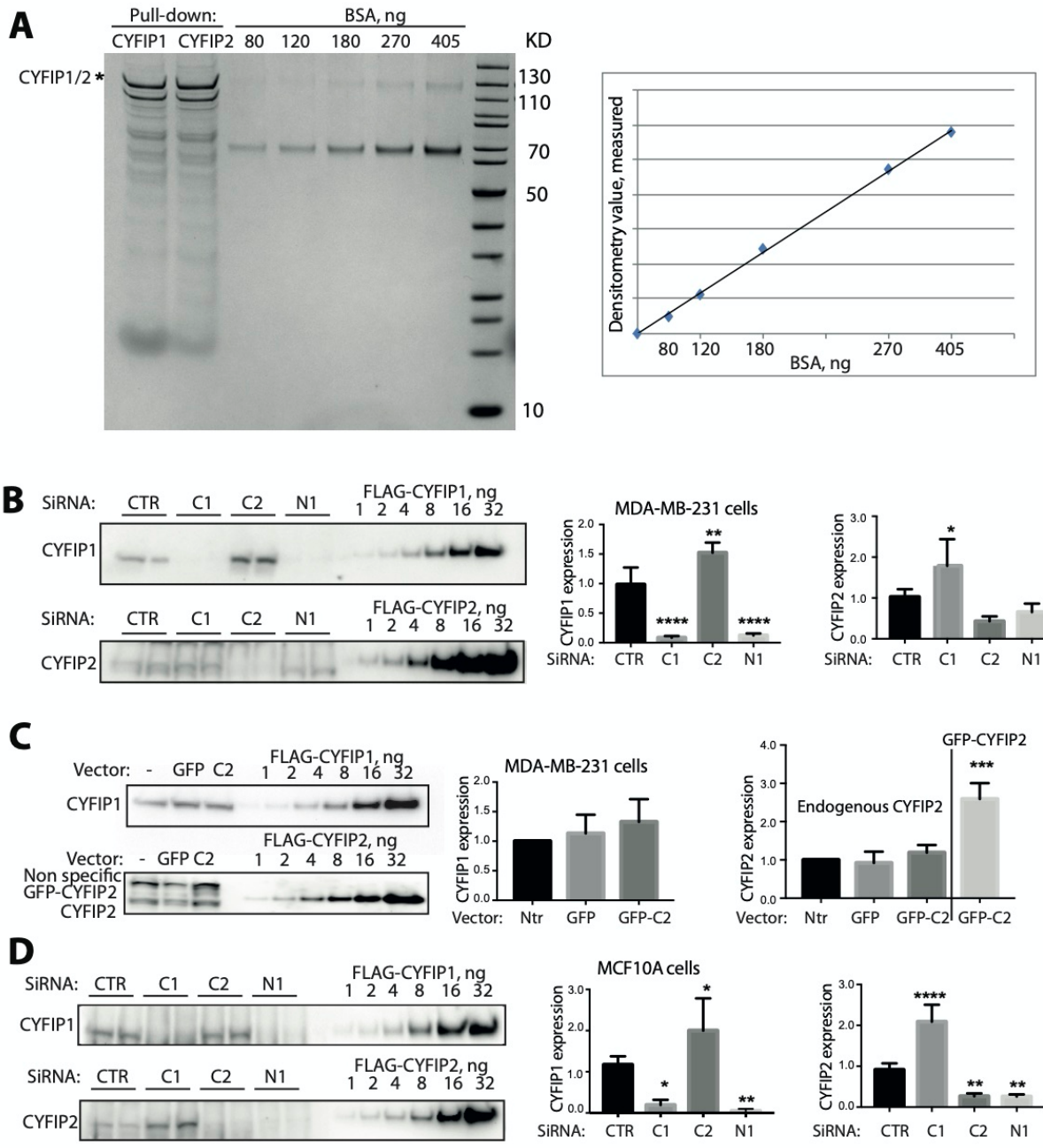


Figure S2. Quantification of CYFIP1/2 proteins in MDA-MB-231 and MCF10A cells. (A) Colloidal coomassie-stained gel with FLAG-tagged, purified CYFIP1/2-containing WAVE complexes, and BSA standard. We have purified tagged CYFIP1 or CYFIP2-containing WAVE complexes from stable 293 Flp-In cell lines. (B) Quantification of CYFIP1, CYFIP2 and NCKAP1 levels in MDA-MB-231 cells treated with siRNAs. Duplicate transfections of the siRNA smartpool were analyzed for each gene, the experiment was repeated twice (total n=4). (C) Quantification of CYFIP1, CYFIP2 and NCKAP1 levels in MDA-MB-231 stable cell lines expressing GFP or GFP-CYFIP2, n=3. (D) Quantification of CYFIP1, CYFIP2 and NCKAP1 levels in MCF10A cells treated with siRNAs. Duplicate transfections of the siRNA smartpool were analyzed for each gene, the experiment was repeated twice (total n=4). Mean ± SD. Shown are the statistically significant differences (one-way ANOVA). *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

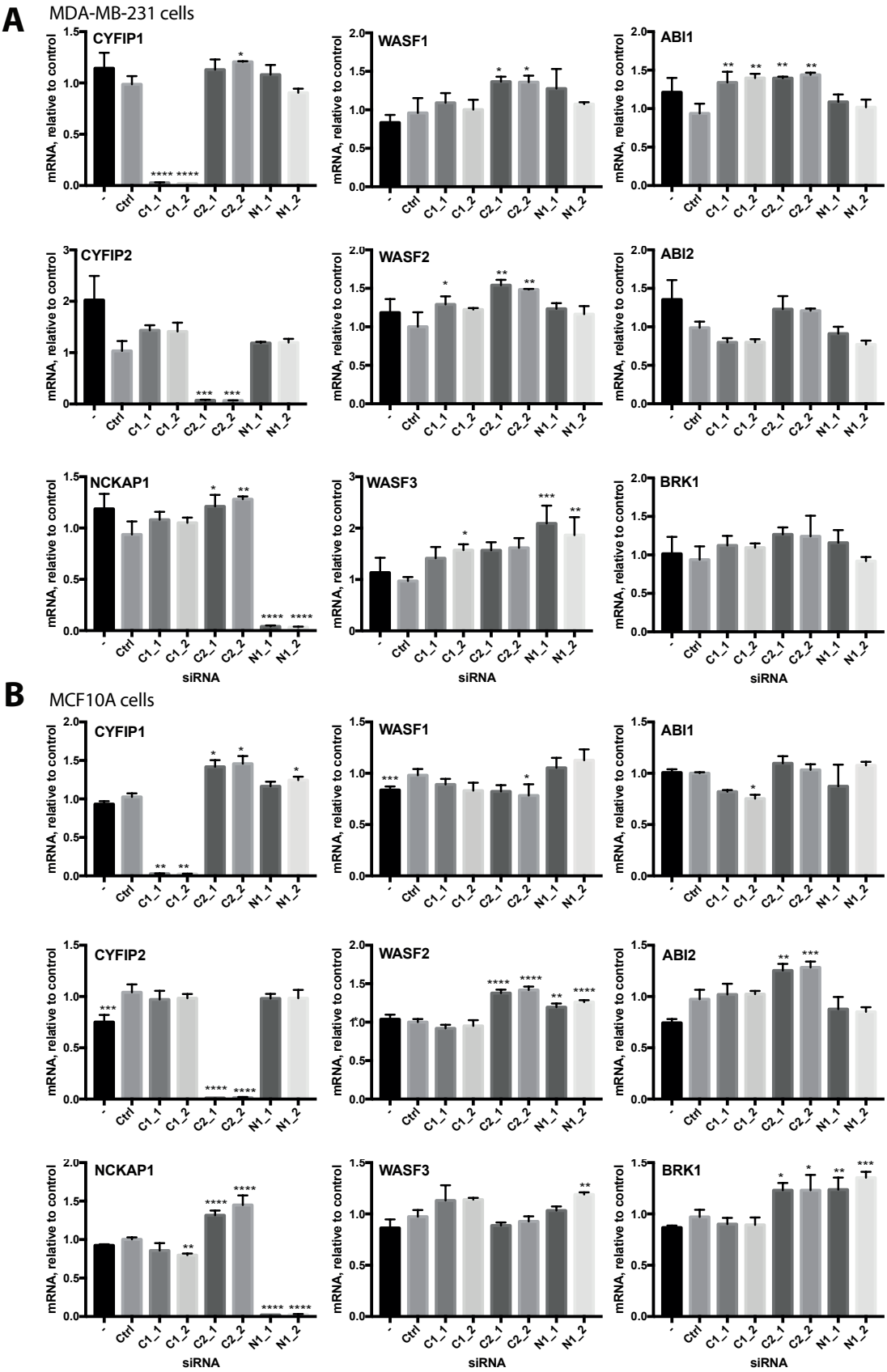


Figure S3. qRT-PCR analysis of mRNA levels of WAVE subunits in MDA-MB-231 and MCF10A cells transfected with two independent siRNAs targeting either CYFIP1 (C1),

CYFIP2 (C2), or NCKAP1 (N1). Mean ± SD of n=3 independent biological experiments, shown are the statistically significant differences (one-way ANOVA). *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

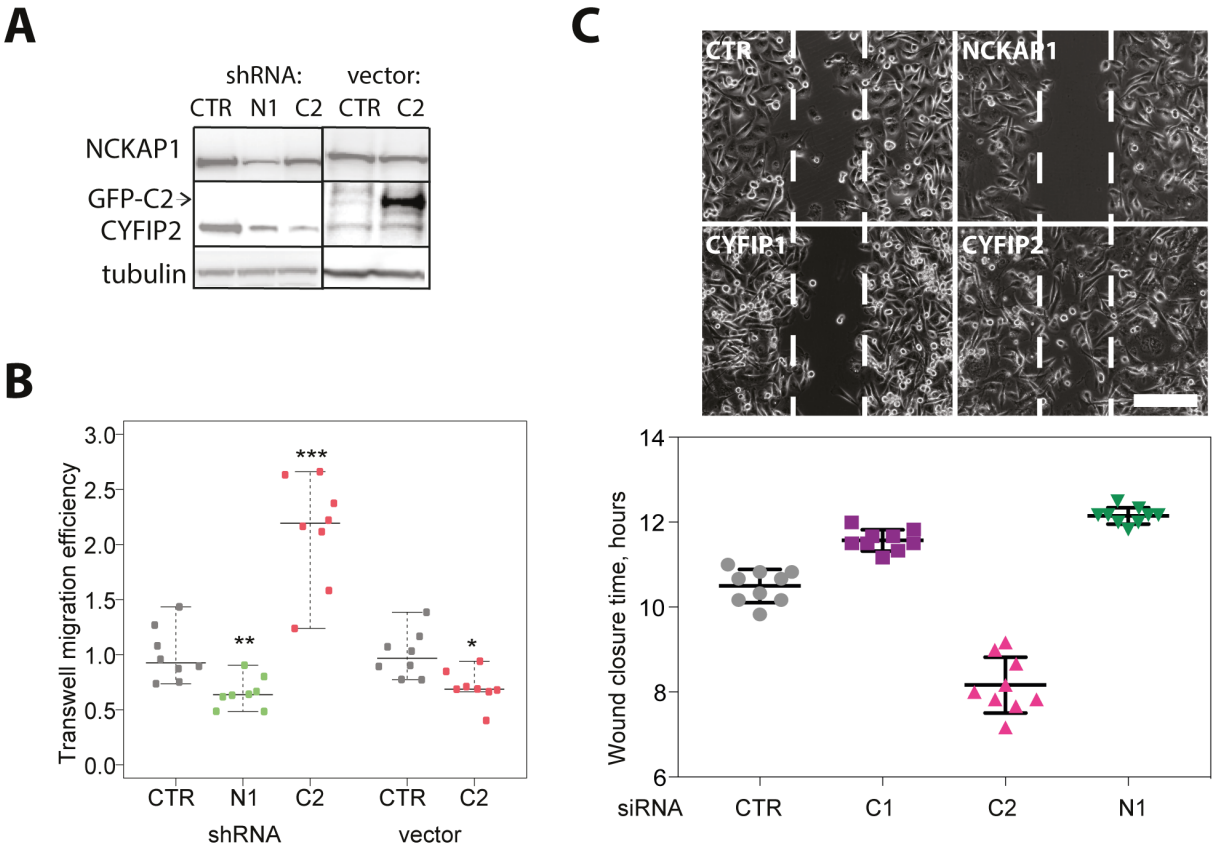
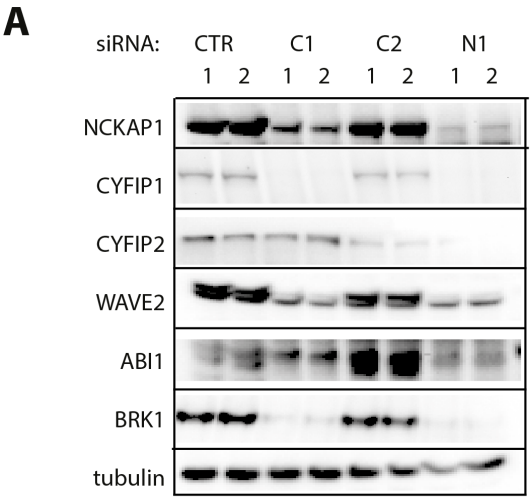
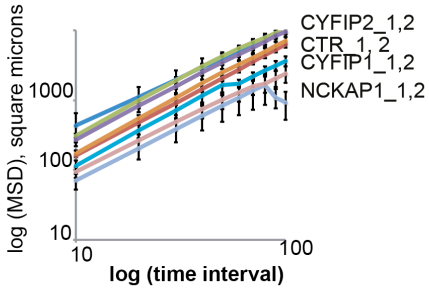
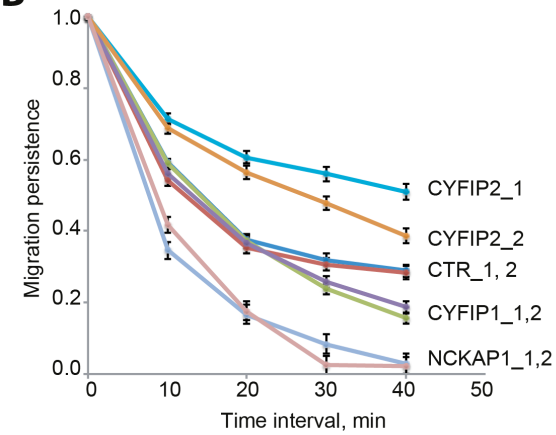


Figure S4. Wound healing and Transwell migration assays of MDA-MB-231 cells. (A) stable MDA-MB-231 cell lines expressing either the indicated shRNAs or overexpressing the GFP-CYFIP2 protein (GFP-C2) were analyzed by Western blots with NCKAP1 and CYFIP2 antibodies. (B) Quantification of Transwell migration efficiency of cells shown in (A), mean ± SD of n=9 (technical repeats), statistical analysis by one-way ANOVA. (C) Cells were transfected with indicated siRNAs targeting NCKAP1 (N1), CYFIP1 (C1), CYFIP2 (C2) or non-targeting controls (CTR) for 5 days, and plated on Ibidi dishes with inserts (Molinie & Gautreau, 2018). 24 h later, the inserts were removed, and wound healing was monitored by video microscopy as previously described. A picture was taken every 10 min for 18 h. Still images taken when the first wound is closed. Scale bar, 400 µm. Quantification of the closure time, mean ± SD of n=9 technical repeats, statistical analysis by one-way ANOVA.







	comparison	p-value	
1	Ctr_1/C1_1	0.46	ns
2	Ctr_2/C1_1	0.82	ns
3	Ctr_1/C1_2	0.38	ns
4	Ctr_2/C1_2	0.71	ns
5	Ctr_1/C2_1	0.0012	**
6	Ctr_2/C2_1	0.00023	***
7	Ctr_1/C2_2	0.044	*
8	Ctr_2/C2_2	0.014	*
9	Ctr_1/N1_1	0.00020	***
10	Ctr_2/N1_1	0.00077	***
11	Ctr_1/N1_2	0.00049	***
12	Ctr_2/N1_2	0.0016	***

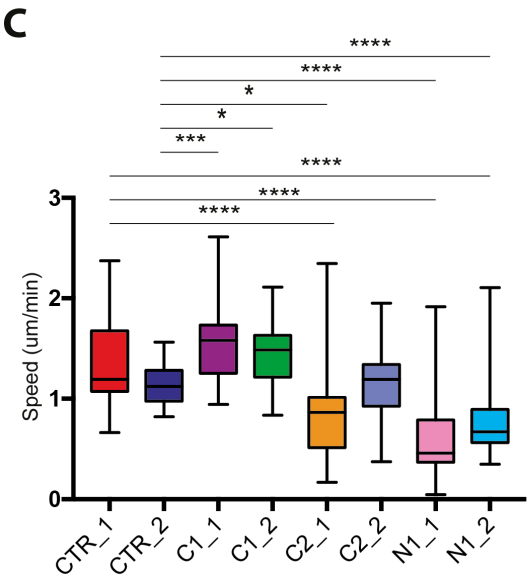


Figure S5. Two independent siRNAs targeting *CYFIP1*, *CYFIP2*, and *NCKAP1* were used to deplete the corresponding proteins from MCF10A cells. Depletion was evaluated by Western blot (A). From single cell trajectories of 2D random migration assays, migration persistence, speed and MSD were calculated and plotted (n=30 cells) (B-D). Statistical analysis was carried out by one-way ANOVA. In panel (C), only the significant differences are indicated.

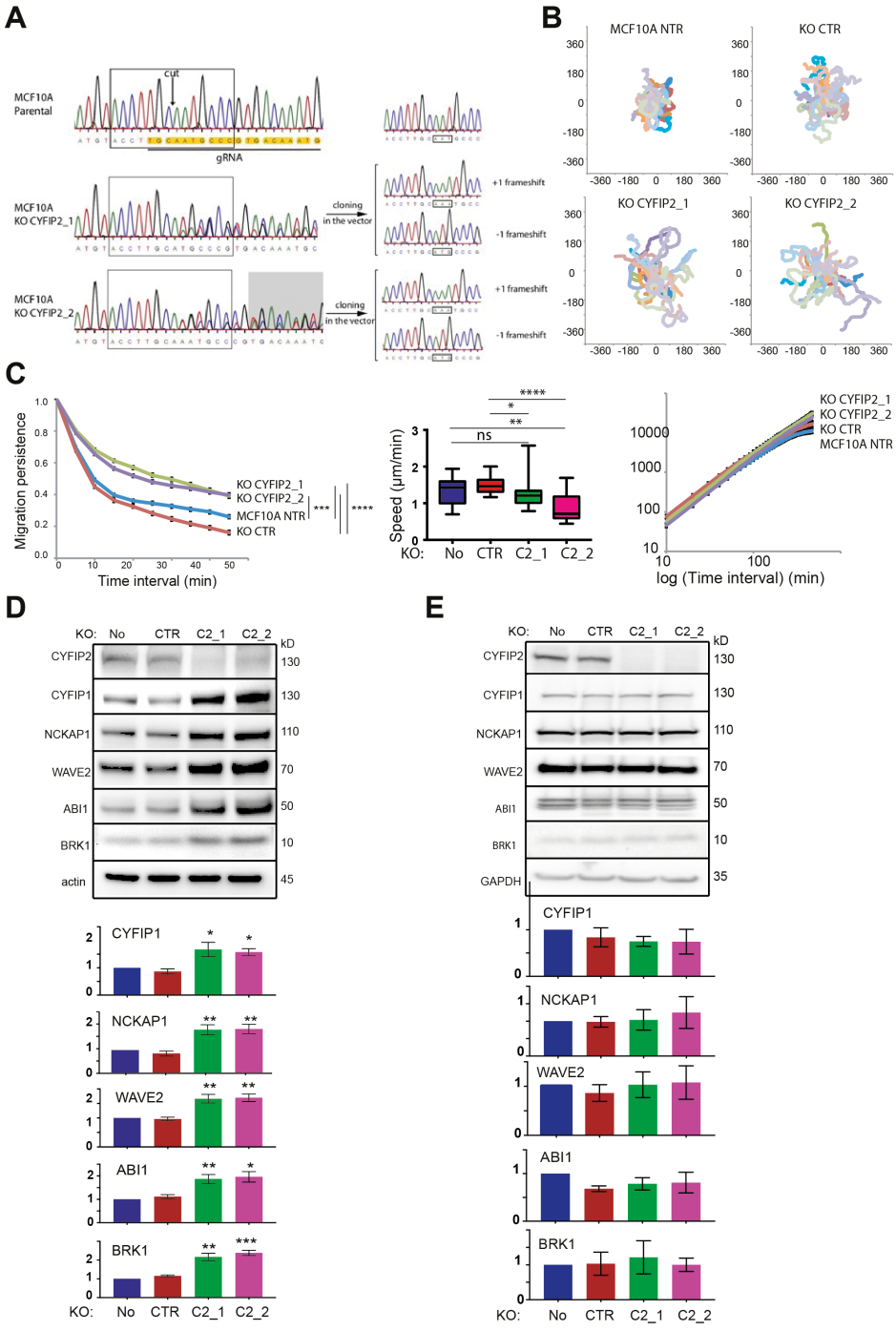


Figure S6. Characterization of CYFIP2 KO clones 1 and 2 in MCF10A cells. (**A**) The only two KO clones that were isolated appeared to have the same genetic alterations. Since they were truly independent, they were both analyzed in parallel in all subsequent assays. (**B**) Trajectories of CYFIP2 KO and control cells in 2D random migration assay, n=30. (**C**) Analysis of migration persistence, speed and MSD of cells shown in (**B**). (**D-E**) Expression of WAVE complex subunits and a quantification of three independent western blots in the above-shown cell lines immediately after selection (**D**) and after two months of culture (**E**).

Supplemental Reference

Molinie N & Gautreau A (2018) Directional Collective Migration in Wound Healing Assays. *Methods in molecular biology (Clifton, NJ)* 1749: 11–19

Legend to supplemental movies

Movie S1. CYFIP2 inhibits migration of MCF7 cells in a wound healing assay. Cells were transfected with indicated siRNAs targeting NCKAP1, CYFIP1, CYFIP2 or non-targeting controls (CTR) for 5 days, and plated on Ibidi dishes with inserts. 24 h later, inserts were removed, and wound healing was monitored by video microscopy. A picture was taken every 20 min. Scale bar: 400 μm.

Movie S2. CYFIP2 inhibits migration of MDA-MB-231 cells in 3D collagen gel. Cells transfected with the indicated siRNAs were recorded by phase contrast optics for 48 h with one frame every 20 min. Scale bar: 50 μm.

Movie S3. CYFIP2 inhibits migration of MCF10A cells. Cells transfected with the indicated siRNAs were recorded by phase contrast optics for 24 h with one frame every 5 min. For the calculation of migration parameters, only single cells were analyzed. Scale bar: 50 μm.

Movie S4. Four-dimensional tracking of prechordal plate nuclei. Nuclei of a Tg(gsc:GFP) embryo were labeled with Histone2B-mCherry (magenta). A Z-stack was acquired every 2 min. Nuclei of prechordal plate cells (identified by GFP expression and morphological criterion), not visible here, were 3D-tracked in time (white squares). Tracks are building up as cells are moving. Animal pole is to the right.

Movie S5. CYFIP2 inhibits actin rich protrusions in zebrafish embryos during gastrulation. Donor embryos were injected with the actin filament marker LifeAct-mCherry mRNA and morpholinos (Mo) targeting CYFIP1 (C1), CYFIP2 (C2), NCKAP1 (N1), alone or in combination with mRNAs encoding the same proteins. 1 frame per minute. Scale bar: 25 μm.

Movie S6. GFP-CYFIP2 restores lamellipodium protrusion and is recruited to the lamellipodial edge. B16-F1 mouse melanoma cells, control and CYFIP1/2 double knock-out (DKO) were transfected with the indicated GFP plasmids and green fluorescence was recorded every 10 s for 10 min.

Movie S7. R87C and I664M CYFIP2 restore lamellipodium protrusion and are recruited to the lamellipodial edge. CYFIP1/2 double knock-out (DKO) B16-F1 cells were transfected with the indicated GFP plasmids and green fluorescence was recorded every 10 s for 10 min.