bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

CYFIP2 containing WAVE complexes inhibit cell migration

Anna Polesskaya[1], Arthur Boutillon[2], Yanan Wang[1],
Marc Lavielle[3], Sophie Vacher[4], Anne Schnitzler[4],
Nicolas Molinie[1], Nathalie Rocques[1], Artem Fokin[1],
Ivan Bièche[4], Nicolas B. David[2], Alexis M. Gautreau[1,5]

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

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

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

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

[5] School of Biological and Medical Physics, Moscow Institute of Physics and Technology,
141700 Dolgoprudny, Russia

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, Arp2/3.

Abbreviations: MFS, metastasis-free survival; KD, knock-down; KO, knock-out; MO, morpholino.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 revealed that cell migration was 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. The specific CYFIP2 function in inhibiting cell migration was related to its unique ability to down-regulate classical pro-migratory WAVE complexes. The anti-migratory function of CYFIP2 was also revealed in migration of prechordal plate cells during gastrulation of the zebrafish embryo, indicating that the unique function of CYFIP2 is critically important in both physiological and pathophysiological migrations.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 membrane 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,

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 one 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. This role of CYFIP2 was not restricted to pathological situations, but rather appeared as a conserved anti-migratory function of thisunique subunit.

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 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 the cohort, 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 tumors of low SBR histological grade (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 (Table S2). 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, Table S2). Indeed, we previously reported that high levels of *NCKAP1* were associated with poor MFS [18]. Levels

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of *CYFIP2* and *NCKAP1L* might be associated with MFS, although the significance was much lower, 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 transcriptomic analyses using cDNA microarrays or RNAseq 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-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 (Fig. S1). 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 (Table S2).

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 (Table S2). Indeed, high levels of *NCKAP1* were associated with poor MFS, whereas high levels of *CYFIP2* were associated with good MFS (Fig.S2).

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 (Table S3). We sorted the 36 WAVE complexes according to increasing p-values. The 18 combinations with lowest p-values all contained NCKAP1 as the NAP subunit, while the 9 combinations with lowest p-values also contained CYFIP2 as the CYFIP subunit. This result confirms our result using univariate analyses that NCKAP1 is the first order predictor, whereas CYFIP2 is the second order predictor in our cohort. Importantly, this multivariate Cox analysis does not suggest a specific WAVE assembly that would be particularly associated with MFS.

We then evaluated different 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 (Table S4) and thus appeared as the optimal model of MFS in our cohort. In other words, 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

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 our prediction concerning the roles of NCKAP1 and CYFIP2 in controlling MFS, we used a public database of breast cancer patients, where the transcriptome of more than 2500 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. 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. To evaluate loss- and gain-of-function of these proteins, we chose the MDA-MB-231 cell line, whose migration and invasion depend on the Arp2/3 complex [23]. Moreover, we compared the two paralogous subunits CYFIP1 and CYFIP2.

Depletion of the different subunits using RNAi had different impact on WAVE complex levels (Fig.2A). 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 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. ABI1 appeared less affected than other subunits by NCKAP1 or CYFIP1 depletion, perhaps in line with its incorporation in alternative multiprotein complexes [24]. In CYFIP1 depleted cells, CYFIP2 was up-regulated, whereas CYFIP1 levels were unaffected by CYFIP2 depletion. Based on these results, NCKAP1 appears as a necessary component of WAVE complexes, whereas CYFIP2 appears more as a regulatory component, or perhaps even a superfluous component of WAVE complexes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To assess migration of MDA-MB-231 cells, we first used the classical Transwell assay. As expected, NCKAP1 depletion significantly decreased the number of cells able to migrate through the filter (Fig.2B). The depletion of CYFIP1 led to the same reduction in cell migration, whereas the depletion of CYFIP2 had the converse effect of increasing cell migration. This converse effect of CYFIP2 depletion is surprising given the established role of the WAVE complex in mediating cell migration, but in line with our survival analyses of breast cancer patients. The effect of CYFIP2 depletion obtained with a pool of siRNAs was confirmed using a previously published shRNA plasmid [5]. Stable MDA-MB-231 lines expressing either a shRNA targeting NCKAP1 or CYFIP2 were also obtained (Fig.2C). 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. GFP-CYFIP2 displayed cytoplasmic and nuclear localization in these cells, but also appeared enriched at the cell periphery during spreading or migration (Fig. 2D, Movie S1). The overexpression of CYFIP2 decreased cell migration in the Transwell assay, whereas CYFIP2 depletion increased it (Fig.2E). As expected, depletion of NCKAP1 decreased migration. Loss- and gain-of function of CYFIP2 thus yield opposite phenotypes. We sought to confirm Transwell results with other migration assays.

In wound healing, the depletion of CYFIP2 induced a quicker closure of the wound than in control cells (Fig.S3, Movie S2). This again was in contrast to the increased closure time associated with the depletion of NCKAP1 or CYFIP1. We then embedded MDA-MB-231 cells in 3D gels of collagen type I to examine the behavior of these breast invasive cells. In these settings mimicking invasion of the mesenchyme, differences in cell trajectories and migration persistence were more dramatic (Fig.2FG). NCKAP1 depleted cells hardly migrated at all, as evidenced by strongly decreased Mean Squared Displacement (MSD), mostly due to reduced speed (Fig.2HI). NCKAP1 depleted cells ended up entering into apoptosis during the first 24 h (Movie S3). 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.2J). In contrast, CYFIP2 depleted cells often formed protrusions, sometimes more than one, and explored a significantly larger territory than controls. This increased MSD of CYFIP2 depleted cells could be accounted for by the dramatically increased migration persistence. Finally, CYFIP2 depleted cells had no issue of survival, unlike NCKAP1 depleted cells. This assay with MDA-MB-231 in 3D collagen gels thus appears to recapitulate the negative role of NCKAP1 and the positive role of CYFIP2 in MFS of breast cancer patients.

CYFIP2 inhibits cell migration by destabilizing WAVE

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 cells were depleted of NCKAP1, of

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

CYFIP1 or of CYFIP2 using the same pools of siRNAs previously used in MDA-MB231. Depletion of siRNA targets was efficient in MCF10A cells and accompanied by similar destabilization of subunits as seen in MDA-MB-231 cells (Fig.3A).

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. In such cultures, 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 S4). CYFIP1 depletion did not have a pronounced effect on cell morphology. We then recorded MCF10A cells for 24 h 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 (Fig.3D-F). 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.S4), 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 [25].

The ability of the WAVE complex to induce membrane protrusions and to drive cell migration relies on the ability of WAVE2 to activate the Arp2/3 complex. The other subunits of the WAVE complex are thought to provide a regulation on the WAVE molecule that the complex contains [26]. Therefore we decided to examine the stability of WAVE2 when protein synthesis is blocked by cycloheximide (CHX). CHX treatment revealed that WAVE2 was more stable in CYFIP2 depleted cells than in controls (Fig.3G). Increased stability of WAVE2 was associated with increased levels of the other subunits, CYFIP1, NCKAP1, ABI1 and BRK1. We reasoned that the short term depletion of CYFIP2 using siRNAs might be the reason why this stabilization effect does not translate into significantly increased levels of WAVE complexes. To obtain long-term CYFIP2 depletion, we inactivated the gene using CRISPR-Cas9. We screened about 100 independent MCF10A clones using CYFIP2 Western blot and identified two CYFIP2 negative clones, which turned out to be knock-out (KO) on both alleles (Fig.S5). Both clones displayed significantly increased levels of all five subunits CYFIP1, NCKAP1, WAVE2, ABI1 and BRK1 (Fig.4A), forming the most classical so-called ubiquitous WAVE complex [27]. This increase was estimated to be approximately 2-fold for each subunit by densitometry (Fig.4B).

CYFIP2 KO clones displayed as expected increased migration persistence, whereas their speed and MSD did not change very significantly (Fig.4C-F). The differentiation of CYFIP2 KO clones was then assayed in matrigel, where MCF10A cells develop acini structures. *CYFIP2* inactivation did not alter the morphogenetic program, but resulted in significantly larger 3D structures containing more cells than the control (Fig.4G-I). The cell polarity, visualized by labeling Golgi apparatus, was not perturbed in these acini (Fig.4J). Similar results were previously obtained when *ARPIN* was inactivated [25]. CYFIP2 thus behaves like the well-established inhibitory protein of cell migration, ARPIN. The RAC1-

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

WAVE-Arp2/3 pathway does not only control cell migration, but also cell cycle progression in these settings [25].

CYFIP2 inhibits cell migration in gastrulating zebrafish embryos

To validate the anti-migratory function of CYFIP2 in a physiological system and to test if this function is specific to breast cells or much more general, we turned to the zebrafish embryo, and in particular to prechordal plate cells, which stereotypically migrate during gastrulation [28,29]. 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 [30,31]. 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.5A). Nuclei were labelled by expression of a Histone2B–mCherry construct, and used to track cell trajectories (Fig.5BC and Movie S5). Experiments were performed in a goosecoid:GFP transgenic line, allowing easy identification of prechordal plate cells. Similarly to what was observed in MDA-MB-231 and MCF10A cells, CYFIP2 depletion increased migration persistence (Fig.5D) 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 (Fig.5EF).

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 a morpholino and/or mRNA were transplanted to the prechordal plate of an uninjected host embryo (Fig.6A, [32]). Actin-rich protrusions were highlighted by the enrichment of the LifeAct-mCherry marker (Fig.6B, Movie S6). CYFIP2 depletion doubled the number of protrusions compared to cells injected with a control morpholino (Fig.6C). 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.6D).

The results using zebrafish embryos are thus perfectly in line with those obtained in 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.

DISCUSSION

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 (MIM #618008) [33,34]. In zebrafish, CYFIP2 loss-of-function mutations result in defective axonal pathfinding in retinal ganglion cells [35]. This function of CYFIP2 is also not redundant with the one of the paralogous subunit, CYFIP1, which is involved in axonal growth [36].

We have been able to explain the protective role of CYFIP2 overexpression in breast cancer by our in vitro or in vivo experiments. 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 [37-40]. CYFIP2 is clearly at odds with other subunits, since it is the first subunit of the WAVE complex that is ever reported to oppose cell migration.

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 [23]. In neuronal growth cones, CYFIP2 was found to localize at the tip of filopodia, structures composed of linear actin and not of branched actin [36,41] in line with an inhibitory function of CYFIP2 on branched actin formation we suggest here. In our stable MDA-MB-231 cell line overexpressing GFP-CYFIP2, the enrichment of CYFIP2 at membrane protrusions is less obvious than the enrichment of other subunits of the complex, consistently with the fact that CYFIP2 impairs membrane protrusions [5,42,43].

CYFIP2 is highly related to CYFIP1, with 88 % identity. Both CYFIP proteins incorporate into WAVE complexes [44-46]. Accordingly, we found here that CYFIP2 depends on NCKAP1 for its stability, like CYFIP1. Importantly, the residues of CYFIP1 that are involved in binding active RAC1 [47,48] are all conserved in CYFIP2. However, CYFIP2 is unique in that it confers an anti-migratory activity to WAVE complexes. The key observation in line with its unique activity is that the depletion of CYFIP2 appears to stabilize WAVE complexes. This effect was subtle in transient transfections of siRNAs and required cycloheximide treatment to reveal it unambiguously. In contrast, in long-term depletion as in CYFIP2 knock-out, the levels of WAVE complexes were found to be enhanced two-fold. It is striking, however, that the effect of CYFIP2 depletion on cell migration is obvious even when total levels of WAVE complexes are not obviously increased, by transient transfections of *CYFIP2* siRNAs. It is thus a possibility that CYFIP2 constantly destabilizes a specific pool of WAVE complexes that is critical for Arp2/3 activation.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The assembly of functional WAVE complexes requires complex intermediates and assembly factors [9,49,50]. Similar processes are at play for the analogous WASH complexes that activate the Arp2/3 complex at the surface of endosomes [51]. When the assembly of these multiprotein complexes is blocked, assembly intermediates do not accumulate, because they are degraded by proteasomes. WAVE complexes are also degraded by proteasomes through direct ubiquitylation of WAVE2, after they have been conformationally activated [52]. WAVE stability is thus tightly controlled, probably as a way to maintain this powerful actin nucleator in check. Our work shows that CYFIP2-dependent controls over the activity and turn-over of WAVE complexes provide an important negative regulation of cell migration. The precise steps that CYFIP2 controls and the detailed structural regulations that the two CYFIP homologs determine provide challenging questions for future investigations.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 [53]. 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'-

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 being 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, [54]).

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 MDA-MB-231 cells expressing CYFIP2 were obtained by transfecting 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; [55]). 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.

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 a total of 6 days.

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:tracrRNA duplex and the purified Cas9 protein by Lipofectamine CRISPRMAX™ Cas9 Transfection Reagent (all reagents from Thermofisher Scientific). The next day, cells

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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. S5 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 [27]. The mouse monoclonal antibody, 231H9, targeting BRK1 was previously described [49]. The antibodies targeting CYFIP-2 (Sigma SAB2701081), NCKAP1 (Bethyl A305-178A) and Tubulin (Sigma T9026) were purchased. Quantification of Western blot was performed by densitometry using the ImageJ software.

Migration assays and live cell microscopy

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. Wound healing was performed using the lifting of Ibidi silicone inserts as previously described [56]. A picture was taken every 10 min for 18 h.

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 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 [57] using DiPer programs [58].

The stable MDA-MB-231 cell line expressing GFP-CYFIP2 was plated on fibronectin-coated slides. Ten min videos (with one frame every 10 sec) were obtained to study GFP localization, while F-actin was labeled using the vital SiR-actin labeling reagent (Cytoskeleton, Inc). Images were acquired using a Leica SP8ST-WS confocal microscope

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

equipped with a HC PL APO 63x/1.40 oil immersion objective, a white light laser, HyD and PMT detectors.

Zebrafish embryos, cell transplantation and imaging

Embryos were obtained by natural spawning of *Tg(-1.8gsc:GFP)ml1* adult fishes [59]. 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 [60]. The control morpholino is a standard control (CCTCTTACCTCAGTTACAATTTATA). mRNAs were synthesized using pCS2+ plasmids containing the human genes described in [27] 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 DiPer [58].

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 [61] 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. Visualisation of images was done on ImageJ, lamellipodia-like actin rich protrusions being quantified on the basis of morphological criteria as described in [30].

Statistical analyses

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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. Statistical analysis was performed using R. Persistence, measured as movement autocorrelation over time is fit for each cell by an exponential decay with plateau ($A = (1 - A_{min}) * e^{-\frac{t}{\tau}} + A_{min}$), where A is the autocorrelation, t the time interval, A_{min} the plateau and $\boldsymbol{\tau}$ the time constant of decay. The plateau value A_{min} is set to zero for cell lines in vitro as they do not display overall directional movement. The time constant $\boldsymbol{\tau}$ of exponential fits were then compared using one-way ANOVA on non linear mixed-effect models for each condition.

Zebrafish cell protrusions. Statistical analysis was performed using linear mixed-effect models to take into account the resampling of the same statistical unit with the R computing environment. Models were compared using one-way ANOVA.

All other comparisons were performed using the Prism software from Graphpad, either with ordinary one-way ANOVA if values were normally distributed or with the non-parametric Kruskal-Wallis otherwise. 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.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Acknowledgements

We thank Nelia Cordeiro for technical support, Theresia Stradal for sharing shRNA plasmids targeting NCKAP1 and CYFIP2, 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 Institut National du Cancer (INCA_6521 for AMG and IB, and INCA_11508 for AMG, IB and ML).

Conflict of interest

The authors declare no conflict of interest

Author contributions

AP performed *in vitro* experiments of cell migration. 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 analyzed GFP-CYFIP2 subcellular localization. AF and YW isolated the CYFIP2 knockout clones and performed experiments with these clones. AMG supervised the study and wrote the manuscript. All authors have commented on the manuscript and approved the submission.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

REFERENCES

[1] Dehal P, Boore JL. Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. PLoS Biol 2005;3:e314–9. doi:10.1371/journal.pbio.0030314.

[2] Ruepp A, Waegele B, Lechner M, Brauner B, Dunger-Kaltenbach I, Fobo G, et al. CORUM: the comprehensive resource of mammalian protein complexes--2009. Nucleic Acids Research 2010;38:D497–501. doi:10.1093/nar/gkp914.

[3] Alberts B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell 1998;92:291–4.

[4] Molinie N, Gautreau A. The Arp2/3 Regulatory System and Its Deregulation in Cancer. Physiol Rev 2018;98:215–38. doi:10.1152/physrev.00006.2017.

[5] Steffen A, Rottner K, Ehinger J, Innocenti M, Scita G, Wehland J, et al. Sra-1 and Nap1 link Rac to actin assembly driving lamellipodia formation. Embo J 2004;23:749–59. doi:10.1038/sj.emboj.7600084.

[6] Lai FPL, Szczodrak M, Block J, Faix J, Breitsprecher D, Mannherz HG, et al. Arp2/3 complex interactions and actin network turnover in lamellipodia. Embo J 2008;27:982–92. doi:10.1038/emboj.2008.34.

[7] Steffen A, Koestler SA, Rottner K. Requirements for and consequences of Rac-dependent protrusion. Eur J Cell Biol 2014;93:184–93. doi:10.1016/j.ejcb.2014.01.008.

[8] Krause M, Gautreau A. Steering cell migration: lamellipodium dynamics and the regulation of directional persistence. Nat Rev Mol Cell Biol 2014;15:577–90. doi:10.1038/nrm3861.

[9] Derivery E, Gautreau A. Generation of branched actin networks: assembly and regulation of the N-WASP and WAVE molecular machines. Bioessays 2010;32:119–31. doi:10.1002/bies.200900123.

[10] Dubielecka PM, Cui P, Xiong X, Hossain S, Heck S, Angelov L, et al. Differential Regulation of Macropinocytosis by Abi1/Hssh3bp1 Isoforms. PLoS ONE 2010;5:e10430. doi:10.1371/journal.pone.0010430.

[11] Suetsugu S, Yamazaki D, Kurisu S, Takenawa T. Differential roles of WAVE1 and WAVE2 in dorsal and peripheral ruffle formation for fibroblast cell migration. Dev Cell 2003;5:595–609.

[12] Iwaya K, Norio K, Mukai K. Coexpression of Arp2 and WAVE2 predicts poor outcome in invasive breast carcinoma. Mod Pathol 2007;20:339–43. doi:10.1038/modpathol.3800741.

[13] Zhang J, Tang L, Shen L, Zhou S, Duan Z, Xiao L, et al. High level of WAVE1 expression is associated with tumor aggressiveness and unfavorable prognosis of epithelial ovarian cancer. Gynecol Oncol 2012;127:223–30. doi:10.1016/j.ygyno.2012.06.008.

[14] Semba S, Iwaya K, Matsubayashi J, Serizawa H, Kataba H, Hirano T, et al. Coexpression of actin-related protein 2 and Wiskott-Aldrich syndrome family verproline-homologous protein 2 in adenocarcinoma of the lung. Clin Cancer Res 2006;12:2449–54. doi:10.1158/1078-0432.CCR-05-2566.

[15] Yang LY, Tao YM, Ou DP, Wang W, Chang ZG, Wu F. Increased Expression of

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Wiskott-Aldrich Syndrome Protein Family Verprolin-Homologous Protein 2 Correlated with Poor Prognosis of Hepatocellular Carcinoma. Clin Cancer Res 2006;12:5673–9. doi:10.1158/1078-0432.CCR-06-0022.

[16] Ji Y, Li B, Zhu Z, Guo X, He W, Fan Z, et al. Overexpression of WAVE3 promotes tumor invasiveness and confers an unfavorable prognosis in human hepatocellular carcinoma. Biomed Pharmacother 2015;69:409–15. doi:10.1016/j.biopha.2014.11.001.

[17] Zhang Y, Guan X-Y, Dong B, Zhao M, Wu J-H, Tian X-Y, et al. Expression of MMP-9 and WAVE3 in colorectal cancer and its relationship to clinicopathological features. J Cancer Res Clin Oncol 2012;138:2035–44. doi:10.1007/s00432-012-1274-3.

[18] Lomakina ME, Lallemand F, Vacher S, Molinie N, Dang I, Cacheux W, et al. Arpin downregulation in breast cancer is associated with poor prognosis. Br J Cancer 2016;114:545–53. doi:10.1038/bjc.2016.18.

[19] Joshi AD, Hegde GV, Dickinson JD, Mittal AK, Lynch JC, Eudy JD, et al. ATM, CTLA4, MNDA, and HEM1 in High versus Low CD38-Expressing B-Cell Chronic Lymphocytic Leukemia. Clin Cancer Res 2007;13:5295–304. doi:10.1158/1078-0432.CCR-07-0283.

[20] Wang C, Tran-Thanh D, Moreno JC, Cawthorn TR, Jacks LM, Wang D-Y, et al. Expression of Abl interactor 1 and its prognostic significance in breast cancer: a tissue-array-based investigation. Breast Cancer Res Treat 2010;129:373–86. doi:10.1007/s10549-010-1241-0.

[21] Zhang J, Tang L, Chen Y, Duan Z, Xiao L, Li W, et al. Upregulation of Abelson interactor protein 1 predicts tumor progression and poor outcome in epithelial ovarian cancer. Hum Pathol 2015;46:1331–40. doi:10.1016/j.humpath.2015.05.015.

[22] Györffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. 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 Res Treat 2009;123:725–31. doi:10.1007/s10549-009-0674-9.

[23] Dang I, Gorelik R, Sousa-Blin C, Derivery E, Guérin C, Linkner J, et al. Inhibitory signalling to the Arp2/3 complex steers cell migration. Nature 2013;503:281–4. doi:10.1038/nature12611.

[24] Innocenti M, Gerboth S, Rottner K, Lai FPL, Hertzog M, Stradal TEB, et al. Abi1 regulates the activity of N-WASP and WAVE in distinct actin-based processes. Nat Cell Biol 2005;7:969–76. doi:10.1038/ncb1304.

[25] Molinie N, Rubtsova SN, Fokin A, Visweshwaran SP, Rocques N, Polesskaya A, et al. Cortical branched actin determines cell cycle progression. Cell Res 2019;29:432–45. doi:10.1038/s41422-019-0160-9.

[26] Chen Z, Borek D, Padrick SB, Gomez TS, Metlagel Z, Ismail AM, et al. Structure and control of the actin regulatory WAVE complex. Nature 2010;468:533–8. doi:10.1038/nature09623.

[27] Gautreau A, Ho H-YH, Li J, Steen H, Gygi SP, Kirschner MW. Purification and architecture of the ubiquitous Wave complex. Proc Natl Acad Sci U S A 2004;101:4379–83. doi:10.1073/pnas.0400628101.

[28] Montero J-A, Kilian B, Chan J, Bayliss PE, Heisenberg C-P. Phosphoinositide 3-

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

kinase is required for process outgrowth and cell polarization of gastrulating mesendodermal cells. Curr Biol 2003;13:1279–89. doi:10.1016/S0960-9822(03)00505-0.

[29] Dumortier JG, Martin S, Meyer D, Rosa FM, David NB. Collective mesendoderm migration relies on an intrinsic directionality signal transmitted through cell contacts. Proc Natl Acad Sci U S A 2012;109:16945–50. doi:10.1073/pnas.1205870109.

[30] Diz-Muñoz A, Krieg M, Bergert M, Ibarlucea-Benitez I, Muller DJ, Paluch E, et al. Control of Directed Cell Migration In Vivo by Membrane-to-Cortex Attachment. PLoS Biol 2010;8:e1000544–12. doi:10.1371/journal.pbio.1000544.

[31] Petrie RJ, Gavara N, Chadwick RS, Yamada KM. Nonpolarized signaling reveals two distinct modes of 3D cell migration. J Cell Biol 2012;197:439–55. doi:10.1083/jcb.201201124.

[32] Boutillon A, Giger FA, David NB. Analysis of In Vivo Cell Migration in Mosaic Zebrafish Embryos. Methods Mol Biol 2018;1749:213–26. doi:10.1007/978-1-4939-7701-7_16.

[33] Nakashima M, Kato M, Aoto K, Shiina M, Belal H, Mukaida S, et al. De novo hotspot variants in CYFIP2 cause early-onset epileptic encephalopathy. Ann Neurol 2018;83:794–806. doi:10.1002/ana.25208.

[34] Zweier M, Begemann A, McWalter K, Cho MT, Abela L, Banka S, et al. Spatially clustering de novo variants in CYFIP2, encoding the cytoplasmic FMRP interacting protein 2, cause intellectual disability and seizures. Eur J Hum Genet 2019:1–13. doi:10.1038/s41431-018-0331-z.

[35] Pittman AJ, Gaynes JA, Chien C-B. nev (cyfip2) is required for retinal lamination and axon guidance in the zebrafish retinotectal system. Developmental Biology 2010;344:784–94. doi:10.1016/j.ydbio.2010.05.512.

[36] Cioni J-M, Wong HH-W, Bressan D, Kodama L, Harris WA, Holt CE. Axon-Axon Interactions Regulate Topographic Optic Tract Sorting via CYFIP2-Dependent WAVE Complex Function. Neuron 2018;97:1078–1093.e6. doi:10.1016/j.neuron.2018.01.027.

[37] Blagg SL, Stewart M, Sambles C, Insall RH. PIR121 Regulates Pseudopod Dynamics and SCAR Activity in Dictyostelium. Curr Biol 2003;13:1480–7. doi:10.1016/S0960-9822(03)00580-3.

[38] Ibarra N, Blagg SL, Vazquez F, Insall RH. Nap1 Regulates Dictyostelium Cell Motility and Adhesion through SCAR-Dependent and -Independent Pathways. Curr Biol 2006;16:717–22. doi:10.1016/j.cub.2006.02.068.

[39] Pollitt AY, Insall RH. Abi Mutants in Dictyostelium Reveal Specific Roles for the SCAR/WAVE Complex in Cytokinesis. Curr Biol 2008;18:203–10. doi:10.1016/j.cub.2008.01.026.

[40] Litschko C, Linkner J, Brühmann S, Stradal TEB, Reinl T, Jänsch L, et al. Differential functions of WAVE regulatory complex subunits in the regulation of actin-driven processes. Eur J Cell Biol 2017;96:715–27. doi:10.1016/j.ejcb.2017.08.003.

[41] Korobova F, Svitkina T. Arp2/3 complex is important for filopodia formation, growth cone motility, and neuritogenesis in neuronal cells. Mol Biol Cell 2008;19:1561–74. doi:10.1091/mbc.e07-09-0964.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[42] Hahne P, Sechi A, Benesch S, Small JV. Scar/WAVE is localised at the tips of protruding lamellipodia in living cells. FEBS Lett 2001;492:215–20.

[43] Stradal T, Courtney KD, Rottner K, Hahne P, Small JV, Pendergast AM. The Abl interactor proteins localize to sites of actin polymerization at the tips of lamellipodia and filopodia. Curr Biol 2001;11:891–5.

[44] Kumar V, Kim K, Joseph C, Kourrich S, Yoo S-H, Huang HC, et al. C57BL/6N mutation in cytoplasmic FMRP interacting protein 2 regulates cocaine response. Science 2013;342:1508–12. doi:10.1126/science.1245503.

[45] Wan C, Borgeson B, Phanse S, Tu F, Drew K, Clark G, et al. Panorama of ancient metazoan macromolecular complexes. Nature 2015;525:339–44. doi:10.1038/nature14877.

[46] Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, et al. The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell 2015;162:425–40. doi:10.1016/j.cell.2015.06.043.

[47] Chen B, Chou H-T, Brautigam CA, Xing W, Yang S, Henry L, et al. Rac1 GTPase activates the WAVE regulatory complex through two distinct binding sites. eLife 2017;6:W529. doi:10.7554/eLife.29795.

[48] Schaks M, Singh SP, Kage F, Thomason P, Klünemann T, Steffen A, et al. Distinct Interaction Sites of Rac GTPase with WAVE Regulatory Complex Have Non-redundant Functions in Vivo. Curr Biol 2018;28:3674–6. doi:10.1016/j.cub.2018.10.002.

[49] Derivery E, Fink J, Martin D, Houdusse A, Piel M, Stradal TE, et al. Free Brick1 is a trimeric precursor in the assembly of a functional wave complex. PLoS ONE 2008;3:e2462. doi:10.1371/journal.pone.0002462.

[50] Wu S, Ma L, Wu Y, Zeng R, Zhu X. Nudel is crucial for the WAVE complex assembly in vivo by selectively promoting subcomplex stability and formation through direct interactions. Cell Res 2012;22:1270–84. doi:10.1038/cr.2012.47.

[51] Visweshwaran SP, Thomason PA, Guerois R, Vacher S, Denisov EV, Tashireva LA, et al. The trimeric coiled-coil HSBP1 protein promotes WASH complex assembly at centrosomes. Embo J 2018;37. doi:10.15252/embj.201797706.

[52] Joseph N, Biber G, Fried S, Reicher B, Levy O, Sabag B, et al. A conformational change within the WAVE2 complex regulates its degradation following cellular activation. Sci Rep 2017;7:44863. doi:10.1038/srep44863.

[53] Bieche I, Parfait B, Laurendeau I, Girault I, Vidaud M, Lidereau R. Quantification of estrogen receptor alpha and beta expression in sporadic breast cancer. Oncogene 2001;20:8109–15. doi:10.1038/sj.onc.1204917.

[54] Li Q, Birkbak NJ, Györffy B, Szallasi Z, Eklund AC. Jetset: selecting the optimal microarray probe set to represent a gene. BMC Bioinformatics 2011;12:474. doi:10.1186/1471-2105-12-474.

[55] González F, Zhu Z, Shi Z-D, Lelli K, Verma N, Li QV, et al. An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell 2014;15:215–26. doi:10.1016/j.stem.2014.05.018.

[56] Molinie N, Gautreau A. Directional Collective Migration in Wound Healing Assays. Methods Mol Biol 2018;1749:11–9. doi:10.1007/978-1-4939-7701-7_2.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[57] Dang I, Gautreau A. Random Migration Assays of Mammalian Cells and Quantitative Analyses of Single Cell Trajectories. Methods Mol Biol 2018;1749:1–9. doi:10.1007/978-1-4939-7701-7_1.

[58] Gorelik R, Gautreau A. Quantitative and unbiased analysis of directional persistence in cell migration. Nature Protocols 2014;9:1931–43. doi:10.1038/nprot.2014.131.

[59] Doitsidou M, Reichman-Fried M, Stebler J, Köprunner M, Dörries J, Meyer D, et al. Guidance of primordial germ cell migration by the chemokine SDF-1. Cell 2002;111:647–59.

[60] Biswas S, Emond MR, Jontes JD. Protocadherin-19 and N-cadherin interact to control cell movements during anterior neurulation. J Cell Biol 2010;191:1029–41. doi:10.1083/jcb.201007008.

[61] Hans S, Christison J, Liu D, Westerfield M. Fgf-dependent otic induction requires competence provided by Foxi1 and Dlx3b. BMC Dev Biol 2007;7:5. doi:10.1186/1471-213X-7-5.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure legends

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 that predicts Metastasis-Free Survival (MFS) based on *NCKAP1* and *CYFIP2* mRNA levels as the only two inputs was derived. The 4 tumors representing the different outskirts of the gene expression in the cohort, highlighted in (A), 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 containing 2519 breast cancer patients.

Figure 2. CYFIP2 inhibits the migration of human breast cancer cells. (**A**) MDA-MB-231 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. (**B**) Quantification of Transwell migration efficiency of cells shown in (**A**), n=9 technical repeats of an experiment representative of 3 independent ones. (**C**) Western blot analysis of NCKAP1 and CYFIP2 expression in stable MDA-MB-231 lines expressing either the indicated shRNAs or overexpressing the GFP-CYFIP2 protein (GFP-C2). (**D**) Subcellular localization of GFP-CYFIP2 stably expressed in live MDA-MB-231 cells. Still image extracted from Movie S1 acquired at the base of the cell using confocal microscopy. F-actin is labeled using the vital SiR-actin labeling reagent. (**E**) Quantification of Transwell migration efficiency of cells shown in (**C**), n=9 technical repeats of an experiment representative of 3 independent ones. (**F-J**) Single cell trajectories, migration persistence, speed, Mean Square Displacement (MSD) and frequency of protrusions in siRNA-transfected MDA-MB-231 cells embedded in 3D collagen gels, n=30. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

Figure 3. CYFIP2 inhibits the migration of normal breast epithelial human cells. (**A**) 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. (**B**) Phase-contrast images of cells described in (A). **(C-F)**, Trajectories, migration persistence, speed, MSD of MCF10A cells transfected with the indicated siRNAs in 2D, n=25. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001. (**G**) Western blot analysis of WAVE subunits in MCF10A cells transfected with the indicated siRNAs and

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

treated or not with 100 μM cycloheximide (CHX) for 16 h. Left and right parts were from the same membrane. The cut indicates where intervening lanes were removed.

Figure 4. A KO of CYFIP2 augments the migration capacity of MCF10A cells by increasing the levels of all WAVE complex subunits. (**A)** A typical western blot of WAVE complex subunits expression in CYFIP2 KO cells compared to controls. (**B**) Quantification of three independent Western blots. (**C-F**) Trajectories, migration persistence, speed, MSD of CYFIP2 KO cell lines compared to CRISPR-Cas9 negative control and non-treated MCF10A cells. (**G-J**) Acini formation in Matrigel by the control and CYFIP2 KO cell lines (**G**) Confocal microscopy images of typical acini labeled in blue (DAPI) and red (GM130, Golgi apparatus). (**H-I**) Quantification of acini volume ($x10^6$ μm^3) and the number of cells per acini, n=20. (**J**), Quantification of cell polarity within the acini, n=130 *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

Figure 5. CYFIP2 inhibits migration persistence 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. (**B**) Isolated slice from a volume acquisition of a Tg(Goosecoid:GFP) zebrafish embryo. GFP-expressing notochord and prechordal plate cells are labelled in green while nuclei express histone2B-mCherry (in magenta). Nuclei of prechordal plate cells are 3D-tracked over time (Movie S5). (**C**) Trajectories of 10 first time points (20 min) for 50 randomly selected cells for each condition, plotted at the same origin (axes in μm). (**D-F**) Migration persistence of prechordal plate cells transfected with the indicated MO and/or mRNA.

Figure 6. CYFIP2 inhibits actin rich protrusions in zebrafish embryos during gastrulation. (**A**) 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. (**B**) Images of representative cells for each condition; red arrowheads indicate actin-rich protrusions. (**C**) Quantification of the average number of protrusions per frame, n=17 to 32 cells from 4 to 5 embryos per condition. (**D**) Quantification of protrusion length, n=95 (randomly selected protrusions per condition). ns P>0.05; * P<0.05; ** P<0.01; *** P<0.001. In panels C and D, p-values without a bar refer to comparisons with the control condition.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.02.184655; this version posted July 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1: 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

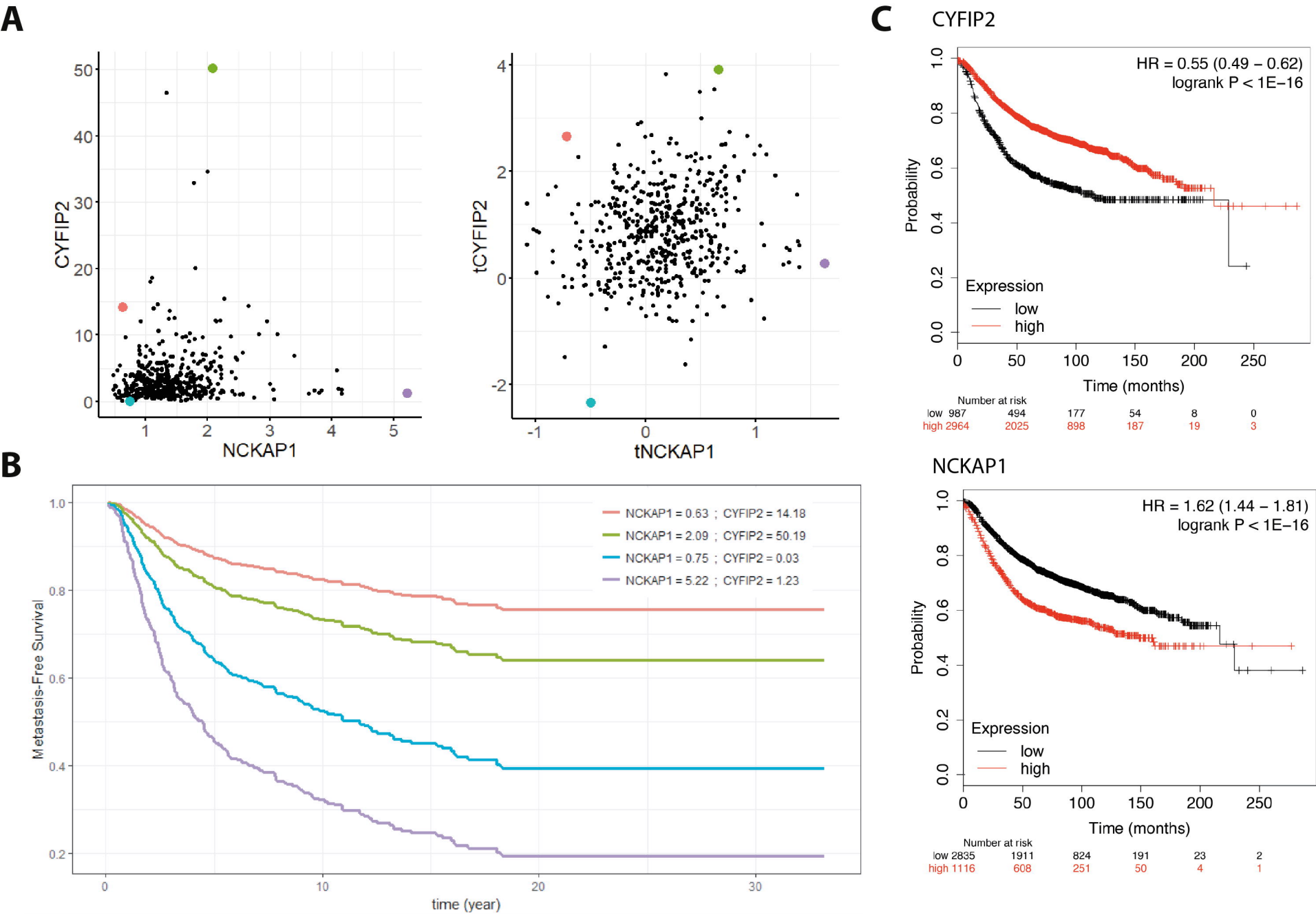

Fig.1, Polesskaya et al.

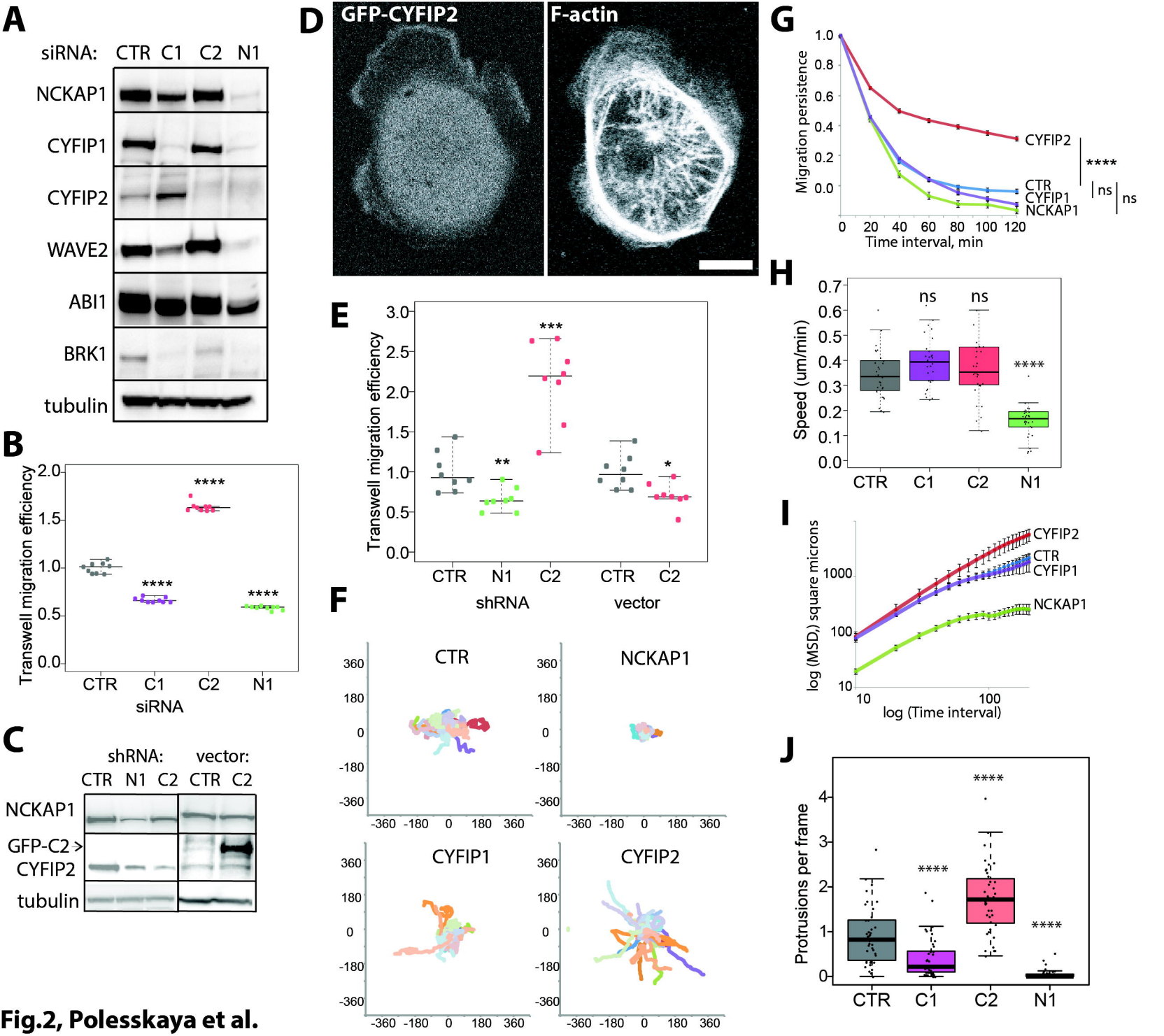

Fig.2, Polesskaya et al.

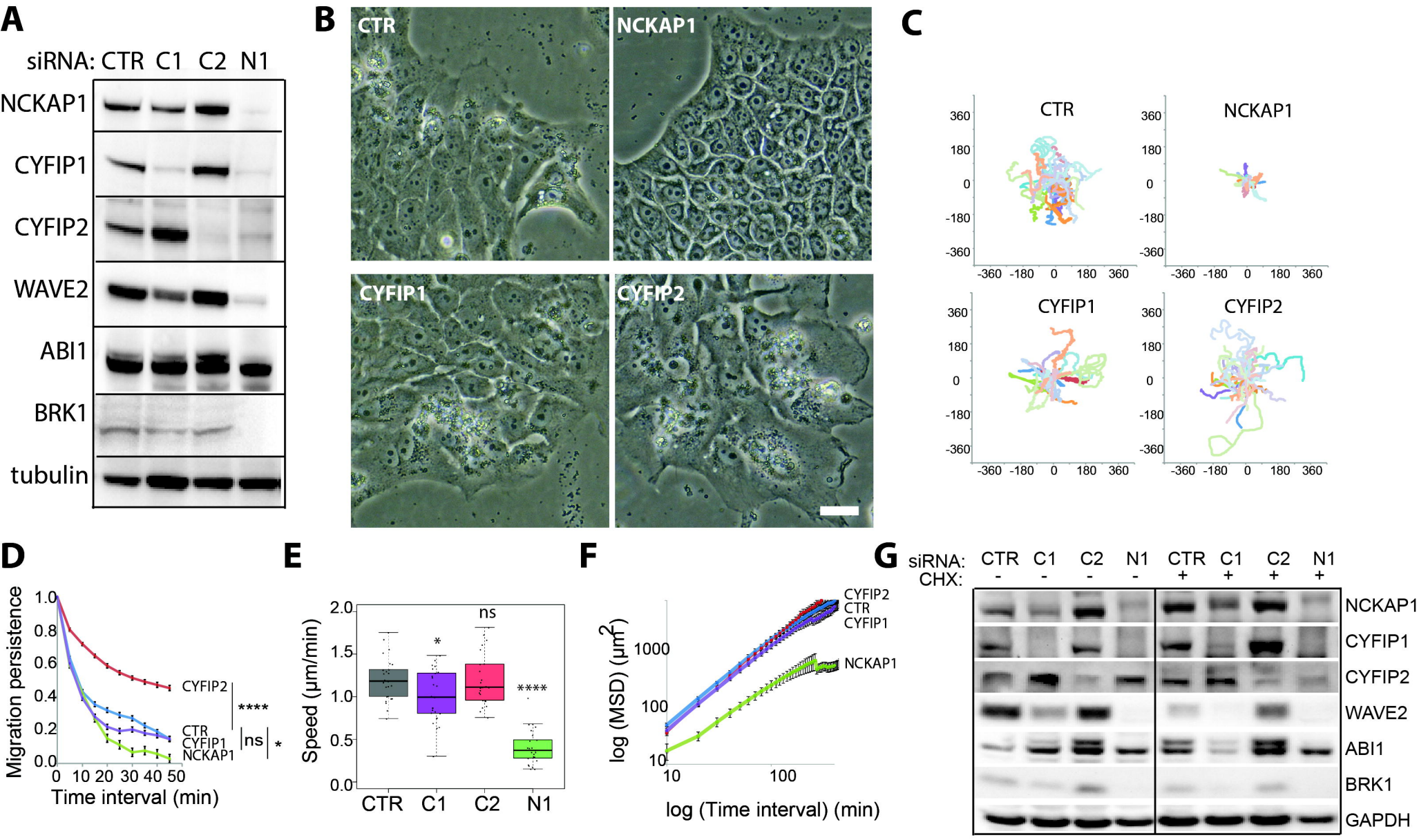

Fig.3, Polesskaya et al.

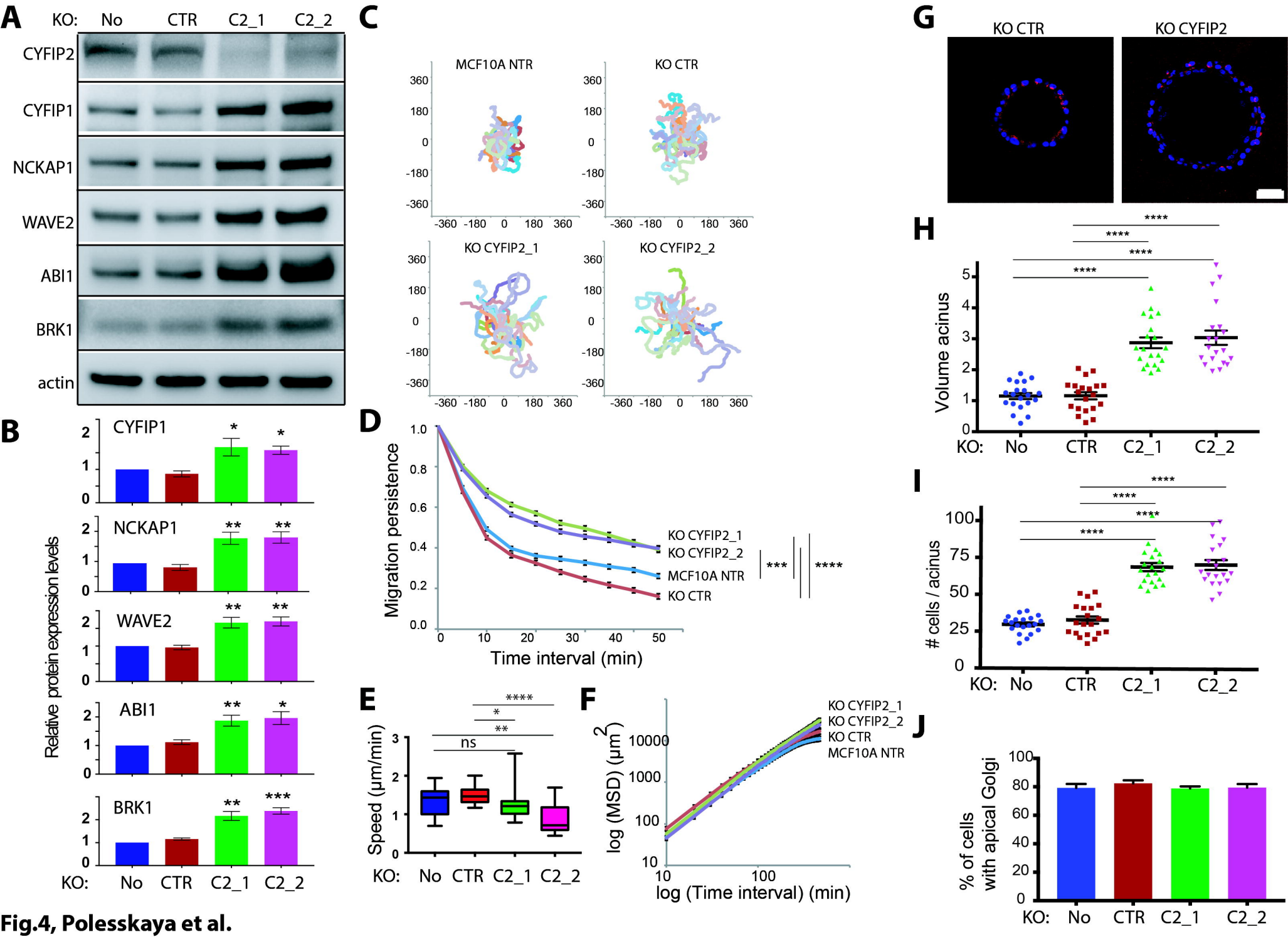

Fig.4, Polesskaya et al.

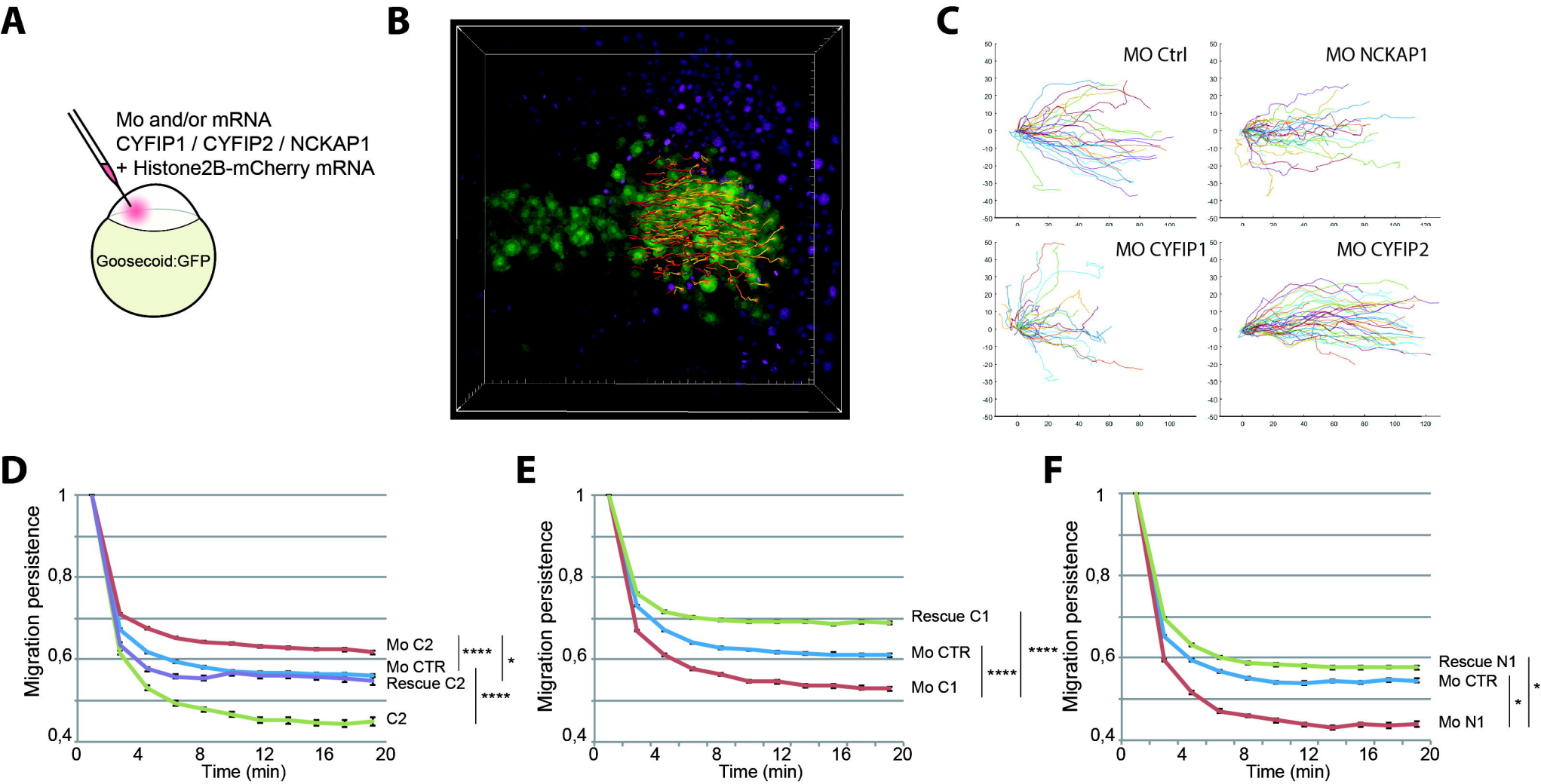

Fig.5, Polesskaya et al.

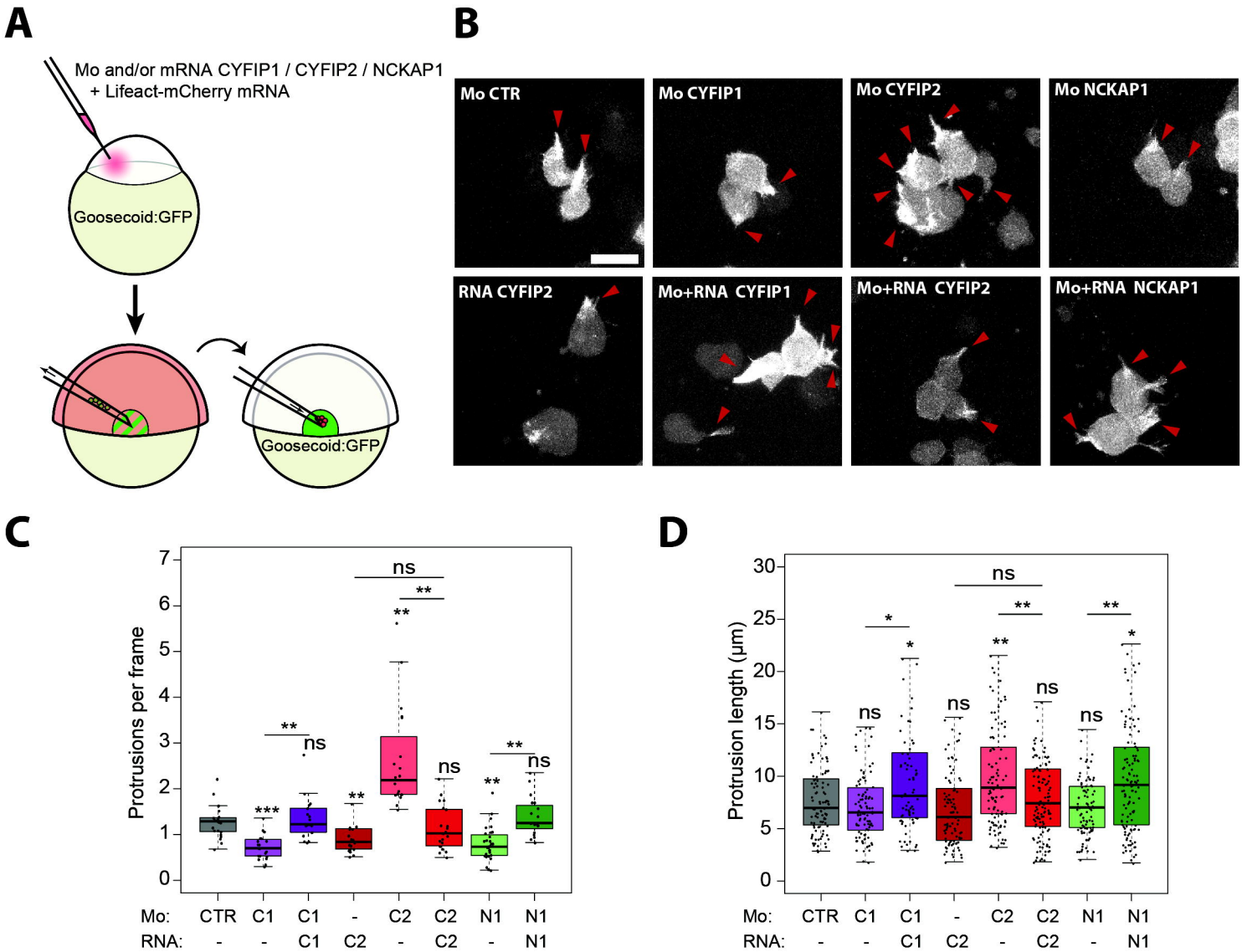

Fig.6, Polesskaya et al.