SUPPLEMENTAL MATERIAL

Complementary proteomics methods identify novel components localizing asymmetrically to centrosomes

Lis Jakobsen[1], Katja Vanselow[1], Marie Skogs[2], Yusuke Toyoda[3], Emma Lundberg[2], Ina Poser[3], Lasse G. Falkenby[1], Martin Bennetzen[1], Jens Westendorf[4], Erich A. Nigg[5], Mathias Uhlen[2], Anthony A. Hyman[3], and Jens S. Andersen[1,*]

[1]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Campusvej 55, DK-5230 Odense, Denmark. [2]School of Biotechnology, AlbaNova University Center, Royal Institute of Technology (KTH), Roslagstullsbacken 21, SE-10691 Stockholm, Sweden. [3]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01307 Dresden, Germany, [4]Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany, [5]Biozentrum, University of Basel, Klingelbergstrasse 50/70, CH-4056 Basel, Switzerland.

*To whom correspondence and requests for materials should be addressed: Jens S. Andersen, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense, Denmark. Tel. +45 6550 2365; Fax. +45 6593 3018; E-mail: jens.andersen@bmb.sdu.dk.

SUPPLEMENTAL FIGURES

Fig. S1. Schematic outline of the double PCP-SILAC method. (A) We explored the ability of PCP-SILAC to profile simultaneously the enrichment of proteins in two independent centrosome preparations using a third preparation as a common internal standard. The correlated distributions between two different gradients were shown to further increase the confidence in organelle classification. In the double PCP-SILAC experiment, three cell populations were labeled with different isotopes. The centrosome-containing fractions prepared from unlabeled cells were mixed and used as the common internal standard for the corresponding fractions prepared from each of the two labeled cell populations. (**B**) SDS-PAGE separation of centrosomal proteins from the PCP-SILAC experiment based on 2 four centrosome containing fractions (see **Fig. 2E, F**). (**C**) The centrosome-containing fractions chosen for the PCP-SILAC experiments were identified by LC-MS analysis of peptide mixtures derived from in-solution digests of aliquots taken from each fraction. In the case shown here, centrosomal proteins were profiled across 10 fractions by averaging the signal intensity of 1158 peptides from 42 centrosomal proteins. These data show that the centrosomes elute in a limited number of fractions.

Fig. S2. Quantitation and organelle classification of proteins by the program MSQuant. (**A**) Quantitation window for Rootletin with a list of peptides assigned to each fraction where the enrichment ratios medium/light and heavy/light for the double PCP-SILAC experiments is shown in the columns labeled 2/1 and 3/1, respectively, followed by their corresponding standard deviation. The program also provide a list of mass spectra data for the quantified peptides, a mass spectrum of the peptide VEAEGQLQQLR with indication of the isotope clusters selected for quantitation, and the LC-elution profiles for each isotope cluster. (**B**) Visualization window with the protein enrichment profiles calculated from the isotope ratios medium/light (dish 2/1) and heavy/light (dish 3/1) in the double PCP-SILAC experiment. Data are shown for five centrosomal proteins as compared to the centrosomal consensus profile in blue and a contaminant in green. The Mahalonobis distance is shown for the non-specific protein.

Fig. S3. Comparison of the PCP-SILAC method with the label-free PCP method. The specificity of the PCP-SILAC methods was compared to the PCP method by plotting classification scores calculated in three different ways from the double PCP-SILAC experiment (experiment A vs experiment B, red: known centrosomal proteins, blue: other proteins). The scores were calculated as the Mahalanobis distance between the profile for each protein and the profile for the centrosomal consensus profile. (**A**) In the case of PCP-SILAC, the profiles were calculated as the median of isotope signal intensity ratios for all peptides representing the same protein in each fraction. (**B**) In the case of PCP, normalized profiles were calculated from individual isotope intensities (medium or heavy) for each protein in each fraction as the sum of intensities for all peptides representing the same protein or (**C**) as the median of normalized profiles determined from the signal intensity (medium or heavy) of each peptide m/z in each fraction. (**D, E**) The superior specificity and sensitivity of the PCP-SILAC method is illustrated by a ROC-curve analysis using the classification scores described above to distinguish true positives (centrosomal proteins) from false positives (non-specific proteins) assuming that the known centrosomal proteins are true-positives (**D**) or that the known and candidate centrosomal proteins are true-positives (**E**). Blue ROC-curve: PCP-SILAC, red ROC-curve: PCP based on the median of profiles for all peptides representing the same protein, green ROC-curve: PCP based on profiles calculated from the sum of intensity of all peptides representing the same protein.

Fig. S4. Organelle classification by the PCP-SILAC methods. (**A, B**) To support the conclusion that the PCP-SILAC method has the ability to classify organellar proteins we performed a more global analysis of the data from the double PCP-SILAC experiment. To this end, we clustered the profiles from the two datasets derived from the double PCP-SILAC experiment into five clusters each using the Mfuzz clustering algorithm. (**C**) Gene Ontology (GO) term enrichment for the proteins observed in each clusters was then performed using hypergeometric testing ($P<0.05$, Benjamini-Hochberg adjusted, enrichment score = -log(P). Clearly, cluster 1 (exp. 3A) and cluster 4 (exp. 3B) perfectly corresponds to each other, since all significant Cellular Compartment (GOCC) terms are

identical. Indeed, cluster 1 (**A**) and cluster 4 (**B**) are highly specific for centrosomal proteins, since they are exclusively enriched (P<0.001, Fishers Exact test) in proteins known to be associated with the GO-term centrosome. Although the gradient centrifugation method was optimized to separate centrosomes from other substructures the non-random GO-term distribution across the different clusters illustrate that other substructures are clearly resolved.

Fig. S5. Images of centrosomal candidate proteins at various stages of the cell cycle. Candidate proteins were stained by antibodies in U-2 OS cells or stably expressed as GFP-fluorescent fusion proteins in HeLa Kyoto cells. (**A**) Antibody-staining of candidates in U-2 OS cells (Alexa 488), the centrosomal marker protein γ-tubulin was stained with Cy3 and DNA with DAPI, (**B, C**) Candidate proteins expressed as GFP-fluorescent fusion proteins in HeLa Kyoto or for CCDC123 in U-2 OS cells. Proteins were stained with γ-tubulin (**B**) or α-tubulin (**C**) and DNA with DAPI. Bars, 10 μm. Additional images are available at http://www.cebi.sdu.dk/CepDB, www.proteinatlas.org, and http://dh.mpi-cbg.de:8080/bac_viewer/search.action.

Fig. S6. Cluster analysis of protein turnover data. (**A**) Known and candidate centrosomal proteins were grouped into proteins with 'low' (<40 %), 'intermediate' (40-75 %) and 'fast' (>75%) turnover (**Table S4**) based on their turnover distribution. The three groups are indicated by the color bar below the distribution histogram. (**B**) Gene Ontology enrichment, using hypergeometric testing, indicate that the three groups of proteins, besides having a number of identical general significant terms, have distinct properties. Proteins with slow turnover are enriched in cytoskeleton associated proteins (P<0.001, Fishers Exact test), proteins with slow and intermediate turnover are enriched in motor protein associated processes as compared to the 'fast' group. The 'intermediate' group is together with the 'fast' group enriched in proteins associated with cell division processes whereas this is not the case for the 'slow' group. The 'fast' group are enriched in protein kinases (CSNK1A1, CSNK1D, PLK1, NEK2, AURKA and CDK5, P<0.01, Fishers exact test) whereas this is not the case for the two other groups. This indicates that even though the centrosomal proteins as a whole is enriched in proteins associated with the cell cycle,

the proteins involved in cell division is more dynamically associated with the centrosome scaffold.

Fig. S7. Subcellular localization of candidates identified by the HPA-screen (A) Antibody-staining of FRMD5 in U-251MG cells indicated that the protein localize to the centrosome. The staining pattern of FRMD5 (Alexa 488) was not supported in U-2 OS cells counterstained for the centrosomal marker protein γ-tubulin (Cy3). Instead, midbody staining was observed. (**B**) BTN3A3 stained a single centrosome at the G1/S phase of the cell cycle in U-2 OS cells as compared to the centrosomal marker protein γ-tubulin (Cy3). Analysis of ciliated RPE cells colored with anti-acetylated tubulin (Cy3) suggests that BTN3A3 associate with the daughter centriole. DNA were stained with DAPI, yellow indicates coincidence of green and red signals. Bars, 5 μm.

Fig. S8. CCDC21 localize to the centrosome in a cell cycle dependent manner. Analysis of CCDC21 in U-2 OS cells stained with anti-CCDC21antibodies (Alexa 488) at various stages of the cell cycle show that CCDC21 reside in nucleoli during interphase and only weakly to centrosomes. During mitosis, stronger staining of CCDC21 is observed in contrast to the diminishing staining of nucleoli. Anti γ-tubulin were stained with Cy3 and DNA with DAPI, yellow indicates coincidence of green and red signals. Bars, 5 μm.

SUPPLEMENTAL TABLES

Table S1. Peptides and proteins identified by PCP-SILAC from centrosome preparation 1. The ratio of enrichment for each fraction is color coded when falling within the boundary of the centrosomal consensus ratio ± 2 × standard deviation.

Table S2. Peptides and proteins identified by PCP-SILAC from centrosome preparation 2. The ratio of enrichment for each fraction is color coded when falling within the boundary of the centrosomal consensus ratio ± 2 × standard deviation.

Table S3. Peptides identified by PCP-SILAC from centrosome preparation 3A and 3B. The ratio of enrichment for each fraction is color coded when falling within the boundary of the centrosomal consensus ratio ± 2 × standard deviation. The relative proteins abundance of is calculated from the peptide intensity signals (see methods).

Table S4. Proteins identified by PCP-SILAC from centrosome preparation 3A and 3B. The ratio of enrichment for each fraction is color coded when falling within the boundary of the centrosomal consensus ratio ± 2 × standard deviation. The proteins are classified as 'centrosome' when previously reported in the literature, 'candidate' when reported in our previous PCP study, and 'new candidate' when having a classification score equal to or below 9 in experiments 3A and 3B. The identified proteins are compared with published proteomics and functional studies of centrosomes and centrosome-related structures: spindle (Sauer et al., 2005), midbody (Skop et al., 2004), cilia (Ostrowski et al., 2002), photoreceptor sensory cilium complex (Liu et al., 2007), cell division defects (Kittler et al., 2007) and cell cycle phenotypes (Neumann et al., 2010). The table contains information about the subcellular localization of proteins to various structures during interphase and mitosis shown in separate columns for the experiments using GFP-tagged proteins and antibodies. The protein turnover data summarizes results presented in **Fig. 7** and **Table S6 and S7**.

Table S5. Subcellular localization of proteins using Human Protein Atlas antibodies. The first part of the table summarizes localization data for proteins identified in the MS-screen comprising known and candidate centrosomal proteins. These proteins were stained with HPA-antibodies in U-2OS cells co-stained with anti-γ-tubulin antibodies as a marker for centrosomes. The second part of the table summarizes localization data for proteins identified in the HPA-screen by the evaluation of images obtained by confocal immunofluorescence microscopy of U-251MG, A-431, and U-2OS cells stained with HPA-antibodies and anti-α-tubulin antibodies as a marker of microtubules. A subset of these candidates was evaluated by γ-tubulin co-localization in U-2OS cells (HPA-tested). Size of structure stained: S=small, M=medium, L=large, VL=very large. Staining of centrosomes is indicated by: Y=yes, Y1= yes but weak, N=No.

Table S6. Turnover of centrosomal proteins**.** Proteins were pulse-labeled for 20 hours and 40 hours with SILAC medium containing Lys-$^{2}H_4$ followed by centrosome isolation, peptide analyses by LC-MS, and protein quantitation.

Table S7. Turnover of centrosomal proteins**.** Proteins were pulse-labeled for 20 hours with SILAC medium containing Lys-$^{13}C_6{}^{15}N_2$ and mixed with fully Lys-$^{2}H_4$ labeled proteins as an internal standard followed by centrosome isolation and peptide analyses by LC-MS. The experiment was repeated with cells fully labeled with Lys-$^{13}C_6{}^{15}N_2$ and pulsed with SILAC medium containing Lys-$^{12}C_6{}^{14}N_2$.

SUPPLEMENTAL MEHHODS

Isolation of centrosomes. Centrosomes were isolated as described by Moudjou and Bornens (Moudjou and Bornens, 1994). Briefly, exponentially growing cells were treated for one hour with nocodazole and cytochalasin D at final concentrations of 60 ng/mLand 1 μg/mL, respectively. Cells were pelleted by centrifugation at 280 ×*g* for 10 min and washed with 10 mM Tris-HCl (pH 7.4), 150 mM NaCl, followed by a second wash with 1 mM Tris-HCl (pH 7.4), 15 mM NaCl, 8% sucrose (w/v). Cells were lysed for 5 minutes in a hypotonic buffer of 1 mM Hepes (pH 7.2), 0.5% NP-40, 0.5 mM $MgCl_2$, 0.1% β-mercaptoethanol containing protease inhibitors. After hypotonic lyses, chromatin and nuclei were pelleted by centrifugation at 2500 ×*g* for 10 min. The lysate was filtered through a medical gauze and remaining chromatin was digested by adjusting the Hepes concentration to 10 mM and by incubation with benzonase at a final concentration of 10 μg/mLfor 30 min at 4^{o}C. Centrosomes were sedimented onto a sucrose cushion (50% sucrose (w/v), 10 mM K-Pipes pH 7.2, 0.1% Triton X-100, 0.1% β-mercaptoethanol) by centrifugation at 22500 × *g* for 20 minutes. Centrosomes from the gradient interface were transferred to a discontinuous sucrose gradient (70%, 50% and 40% sucrose in 10 mM K-Pipes, pH 7.2, 0.1% Triton X-100, 0.1% β-mercaptoethanol) and centrifuged at 116000 ×*g* for 75 minutes. Fractions of either 0.4 or 0.5 mL were collected from the bottom of

the tubes. Centrosome-containing fractions were identified by LC-MS of peptide mixtures derived from the in-solution digest of 25 µl aliquots of each fraction.

Preparation of peptides for mass spectrometry. Aliquots from each sucrose gradient fraction were combined with 100 µL 8 M urea, 100 mM Tris-HCl, pH 8 and incubated at 37^{o}C overnight with endoproteinase Lys-C or trypsin according to the SILAC amino acids used. The resulting peptides were reduced with dithiothreitol (DTT) and alkylated with iodoacetamid for 20 minutes before acidification with 1% triflouroacetic acid (TFA) and desalting on STAGE tips for mass spectrometric analysis. Pelleted centrosomes from the gradient fractions were dissolved in LDS sample buffer, heated for 10 min at 70°C, reduced with DTT, and alkylated with iodoacetamide. Proteins were separated by electrophoresis on NuPAGE Bis-Tris 4-12% gradient gels (Invitrogen) and stained with Comassie Blue. Gel slices (10-12 slices) were cut into small pieces, washed several times with 50 mM ammonium bicarbonate, 50% acetonitrile or ethanol and incubated with 12.5 ng/µL endoprotease LysC in 50 mM ammonium bicarbonate at 37°C overnight. The resulting peptides were extracted with 1% TFA, desalted on STAGE tips and eluted into 96 well plates for mass spectrometric analysis.

Mass spectrometry and peptide identification. Mass spectrometric analysis was performed by LC-MS using an Agilent HP1100 system and a linear ion-trap Fourier-transform ion-cyclotron resonance (LTQ-FT-ICR) or an LTQ-Orbitrap mass spectrometer (Thermo Fisher). Peptides were separated by a 120 min linear gradient of 95% buffer A (0.5% acetic acid in water) to 50% buffer B (80% acetonitrile, 0.5 % acetic acid in water). The LTQ-FT-ICR instrument was operated in the data-dependent mode to acquire high-resolution precursor ion spectra (m/z 300–1,500, resolution 50,000 and ion accumulation to a target value of 3×10^6 ions) in the ICR cell. The three most intense ions were sequentially isolated for accurate mass measurements by selected ion monitoring (SIM) scans (10 Da mass windows, resolution 50,000, and a target accumulation value of 50,000). The ions were simultaneously fragmented in the linear ion trap with a normalized collision energy setting of 27% and a target value of 10,000. The LTQ-Orbitrap mass spectrometer was operated in the data-dependent mode to

acquire high-resolution precursor ion spectra (m/z 300–1,500, resolution 60,000 and ion accumulation to a target value of 5×10^6 ions) in the Orbitrap. The five most intense ions were sequentially isolated for accurate mass measurements and fragmentation in the linear ion trap.

Statistical analysis. The relative enrichment of quantifiable peptides in each of the sucrose gradient fractions was calculated from the peak area of the 'light/heavy' isotope ratio for each single scan mass spectrum. For each experiment, the median of $\log_2$ transformed peptide ratios were computed for each protein in each fraction. Ratios were expressed in a vector form, $x = (x_1, x_2, \ldots, x_n)$, where each dimension (entry) corresponds to an analysis fraction, and n is the total number of fractions. Ratios of a 'consensus set' of 32 known centrosomal proteins, against which all other proteins were compared, were averaged along each dimension to give a vector of expected values for each fraction, $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$, known as the centroid. The distance of a given observation (*e.g.* from another protein) from this point is a measure of the similarity of that protein's profile to the tuning set. This similarity was quantified as a Mahalanobis distance, d_M:

$$d_M = ((x - \mu)^T \Sigma^1 (x - \mu))^{0.5}$$

where S is the variance-covariance matrix of the tuning set. The Mahalanobis distance represents the distance of a given point from the centre (mean) of the tuning proteins in n-dimensional space, normalised by the variance (and covariance) of this group of points. Equal deviations from the centroid in two different dimensions (fractions) may yield different contributions to the Mahalanobis distance, depending upon the breadth of the tuning protein distribution in those dimensions. The metric therefore makes a simple and easily understood measure of the similarity of a given point to the consensus set. Relative protein abundance was estimated from the averaged intensities of all peptide m/z signals associated with each protein in each fraction and calculated as the centroid of the resulting abundance profiles. The percentage of pulse isotope labelling were calculated as

$$(1/((1/\text{ratio})+1))*100$$

References

Kittler, R., Pelletier, L., Heninger, A.K., Slabicki, M., Theis, M., Miroslaw, L., Poser, I., Lawo, S., Grabner, H., Kozak, K., Wagner, J., Surendranath, V., Richter, C., Bowen, W., Jackson, A.L., Habermann, B., Hyman, A.A. and Buchholz, F. (2007) Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat Cell Biol*, **9**, 1401-1412.

Liu, Q., Tan, G., Levenkova, N., Li, T., Pugh, E.N., Jr., Rux, J.J., Speicher, D.W. and Pierce, E.A. (2007) The proteome of the mouse photoreceptor sensory cilium complex. *Mol Cell Proteomics*, **6**, 1299-1317.

Moudjou, M. and Bornens, M. (1994) Isolation of centrosomes from cultured animal cells. In Celis, J.E. (ed.), *Cell Biology: A Laboratory Handbook*. Academic Press, San Diego, CA, pp. 595-604.

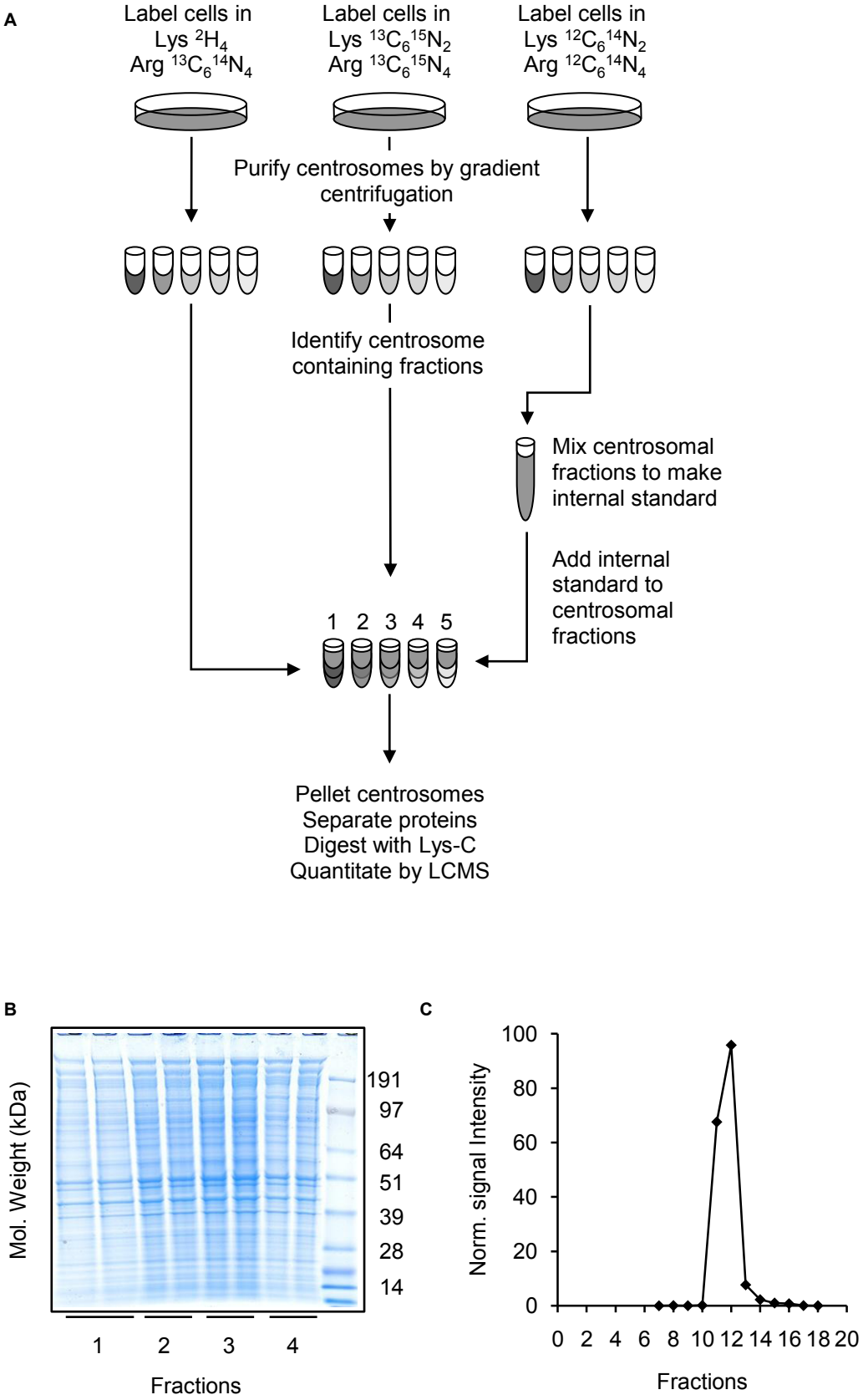
Neumann, B., Walter, T., Heriche, J.K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U., Cetin, C., Sieckmann, F., Pau, G., Kabbe, R., Wunsche, A., Satagopam, V., Schmitz, M.H., Chapuis, C., Gerlich, D.W., Schneider, R., Eils, R., Huber, W., Peters, J.M., Hyman, A.A., Durbin, R., Pepperkok, R. and Ellenberg, J. (2010) Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature*, **464**, 721-727.

Ostrowski, L.E., Blackburn, K., Radde, K.M., Moyer, M.B., Schlatzer, D.M., Moseley, A. and Boucher, R.C. (2002) A proteomic analysis of human cilia: identification of novel components. *Mol Cell Proteomics*, **1**, 451-465.

Sauer, G., Korner, R., Hanisch, A., Ries, A., Nigg, E.A. and Sillje, H.H. (2005) Proteome analysis of the human mitotic spindle. *Mol Cell Proteomics*, **4**, 35-43.

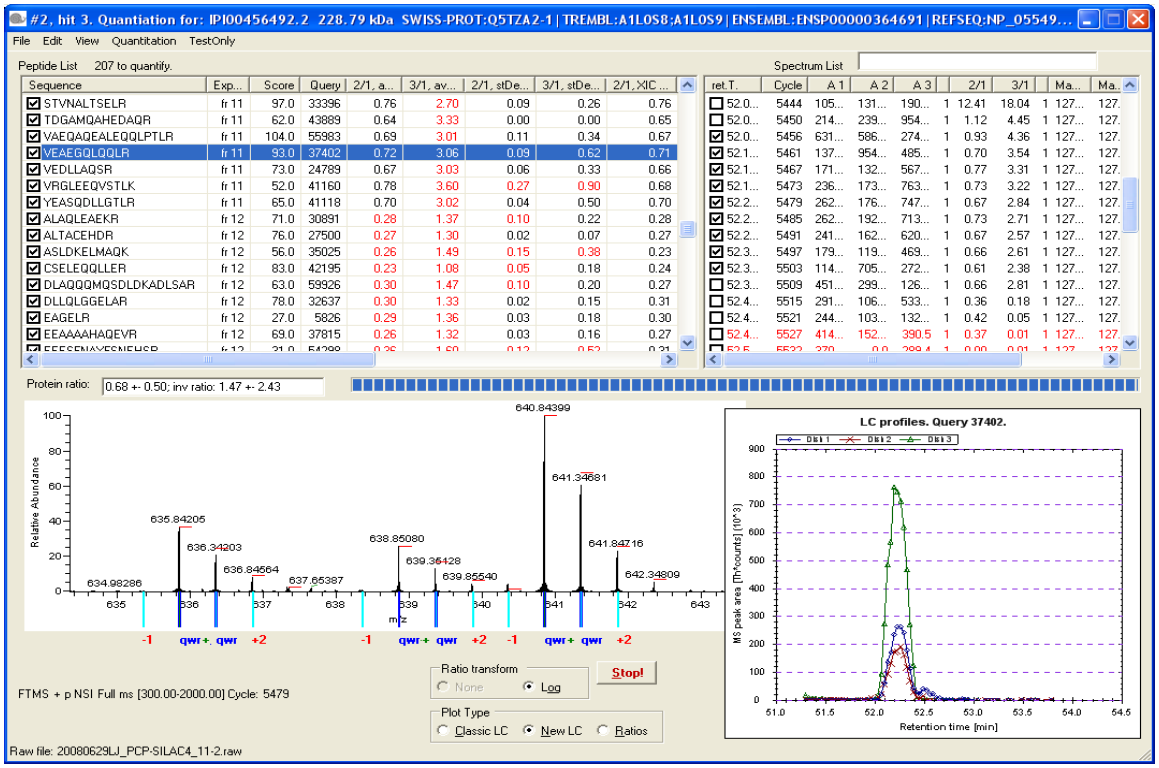
Skop, A.R., Liu, H., Yates, J., 3rd, Meyer, B.J. and Heald, R. (2004) Dissection of the mammalian midbody proteome reveals conserved cytokinesis mechanisms. *Science*, **305**, 61-66.

Supplemental Fig. 1

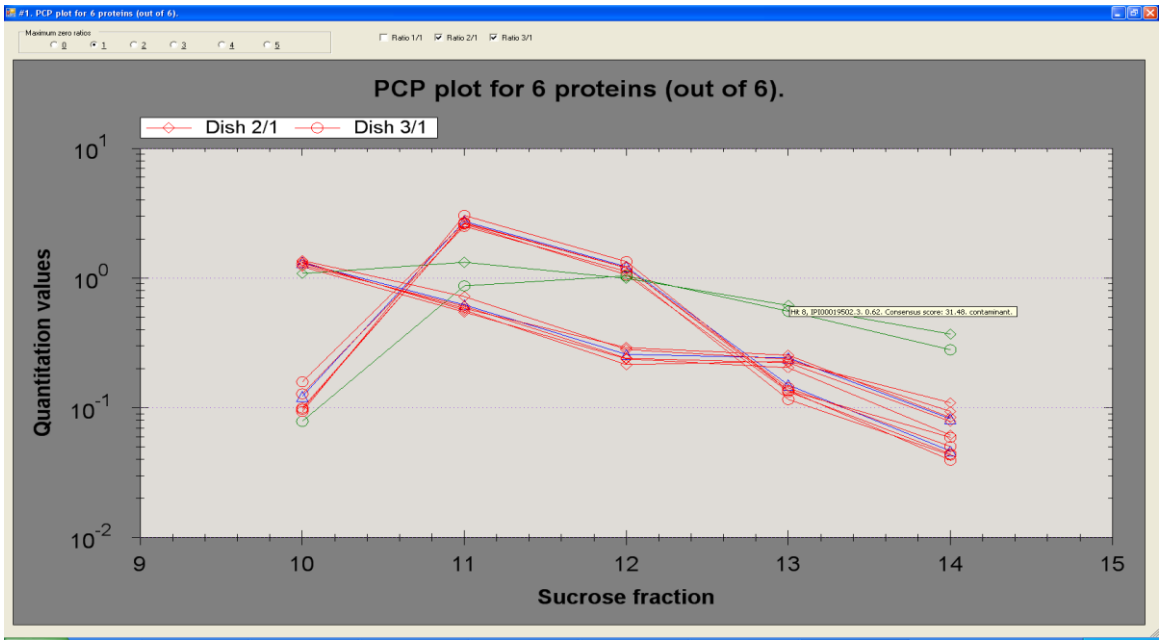


Supplemental Fig. 2

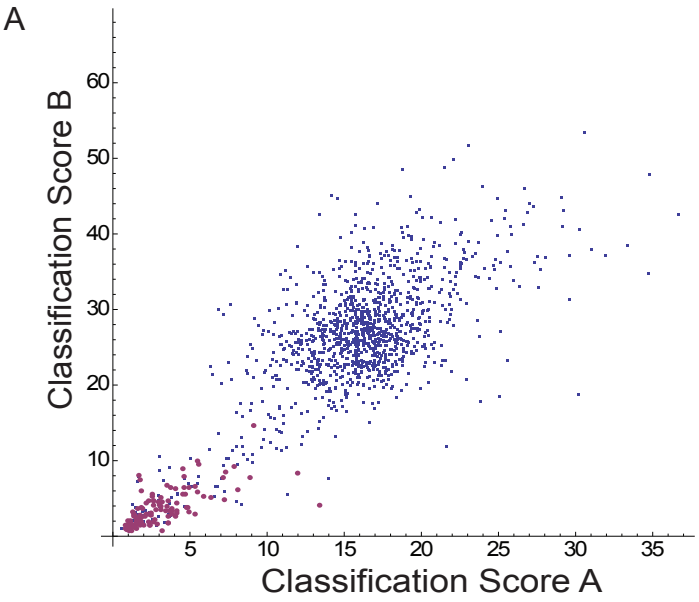
A

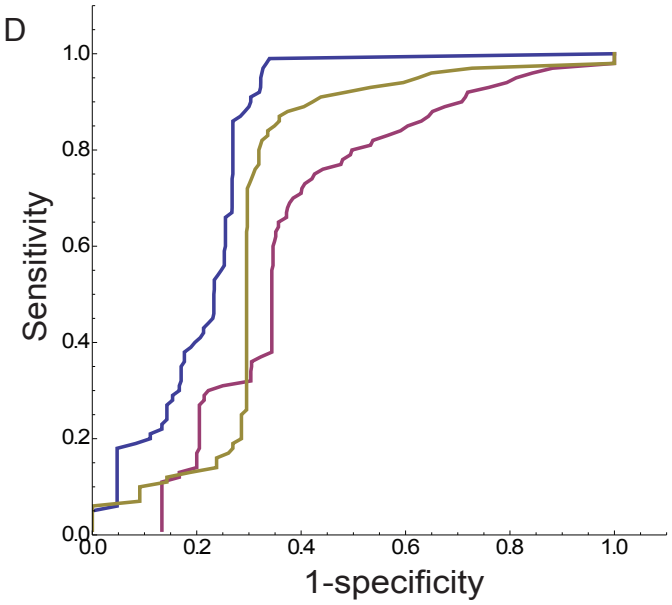


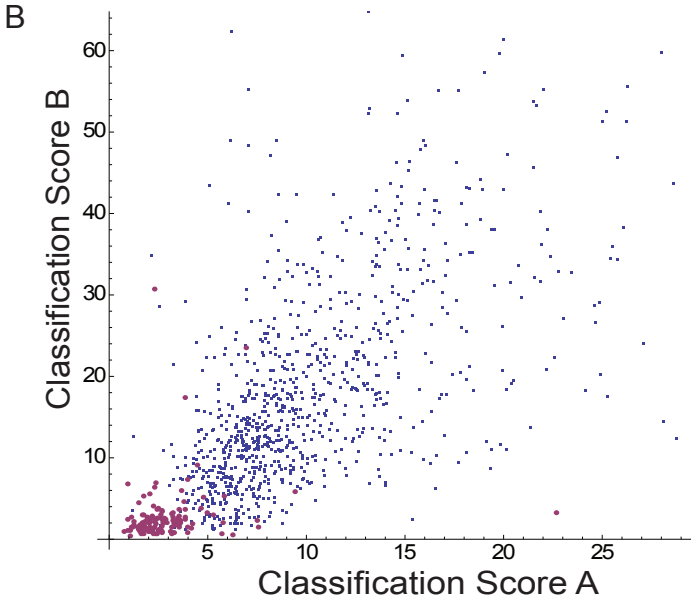
B

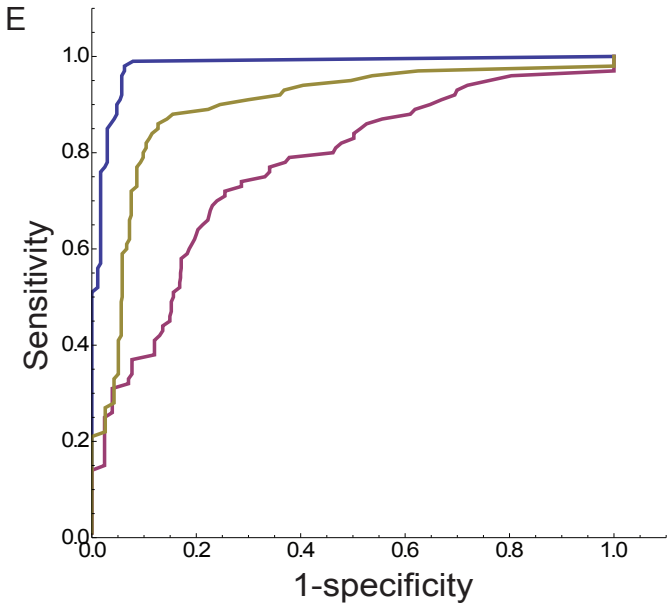


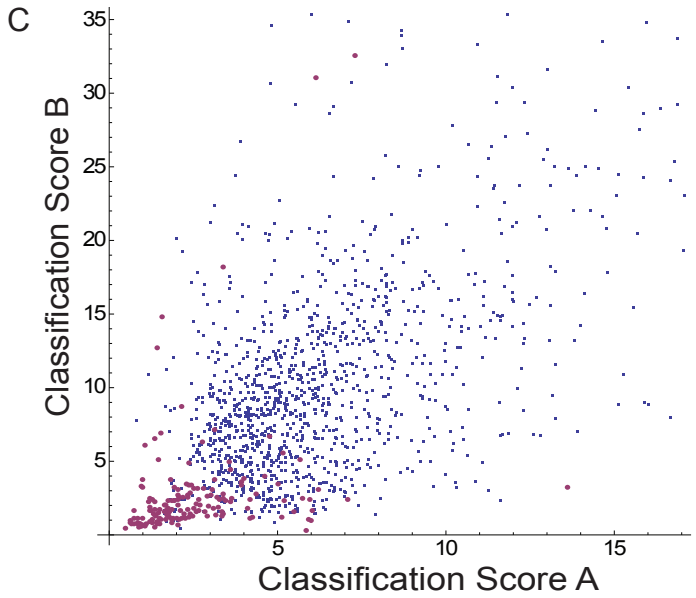
Supplemental Figure 3





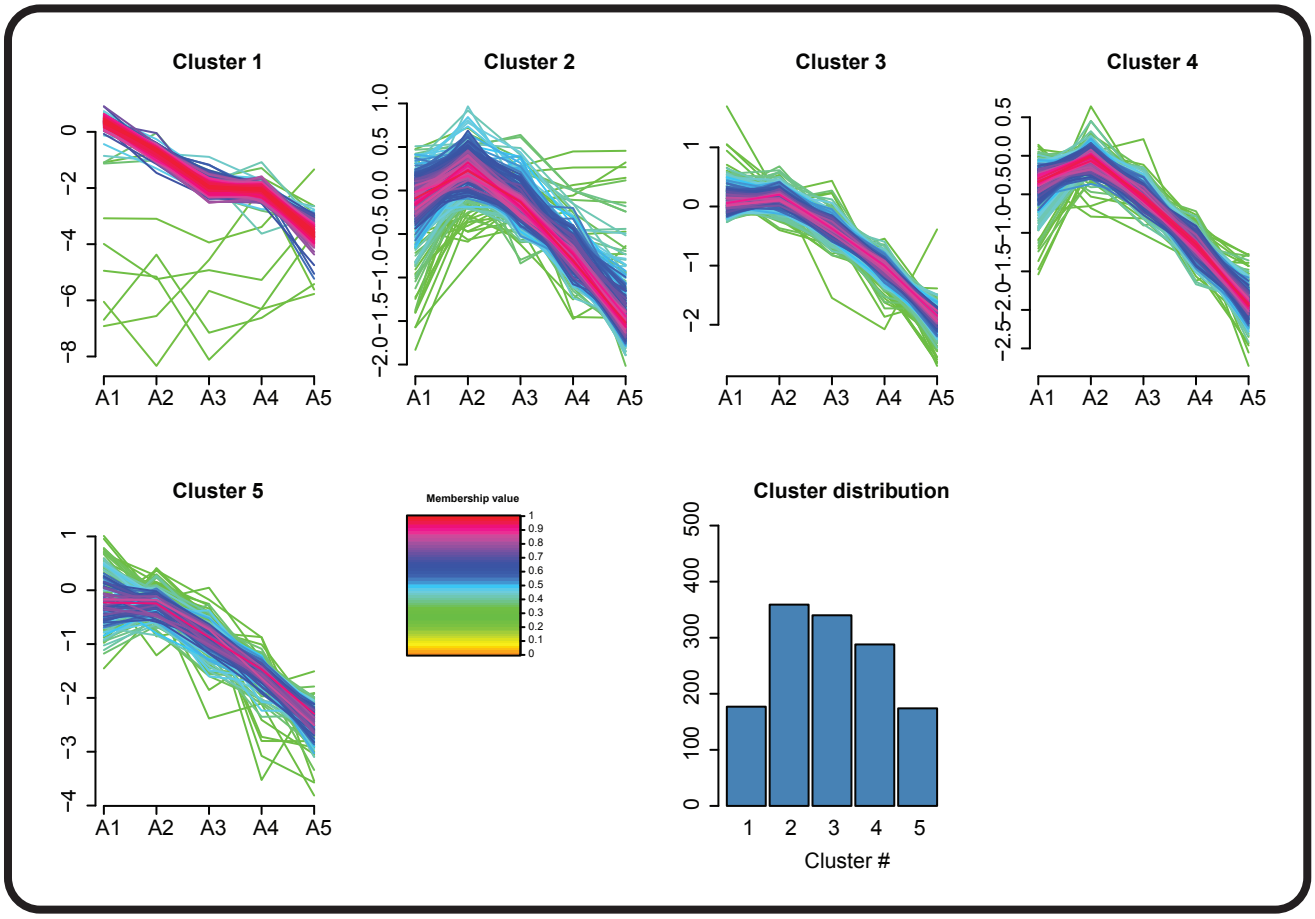




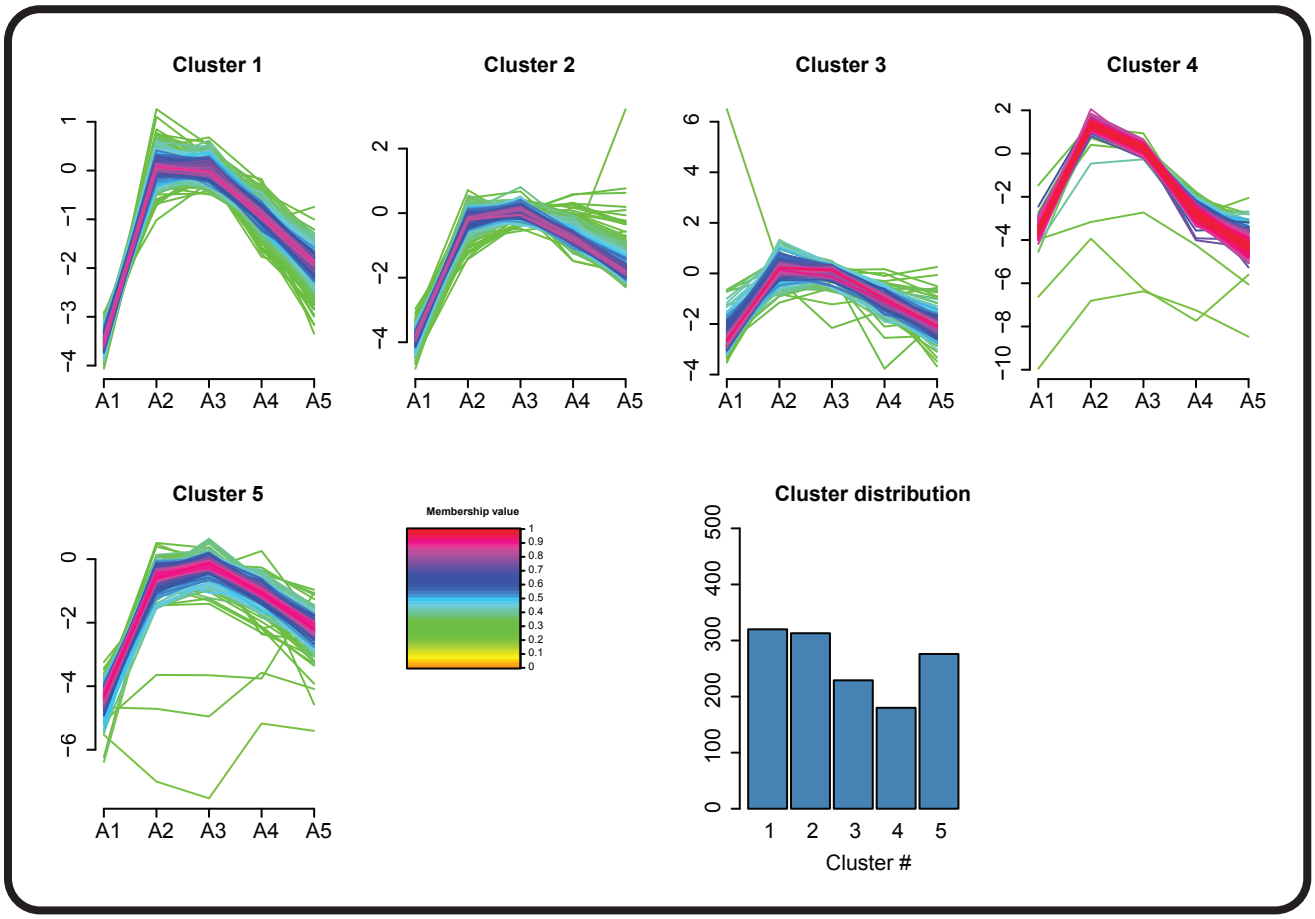


Supplemental Fig. 4

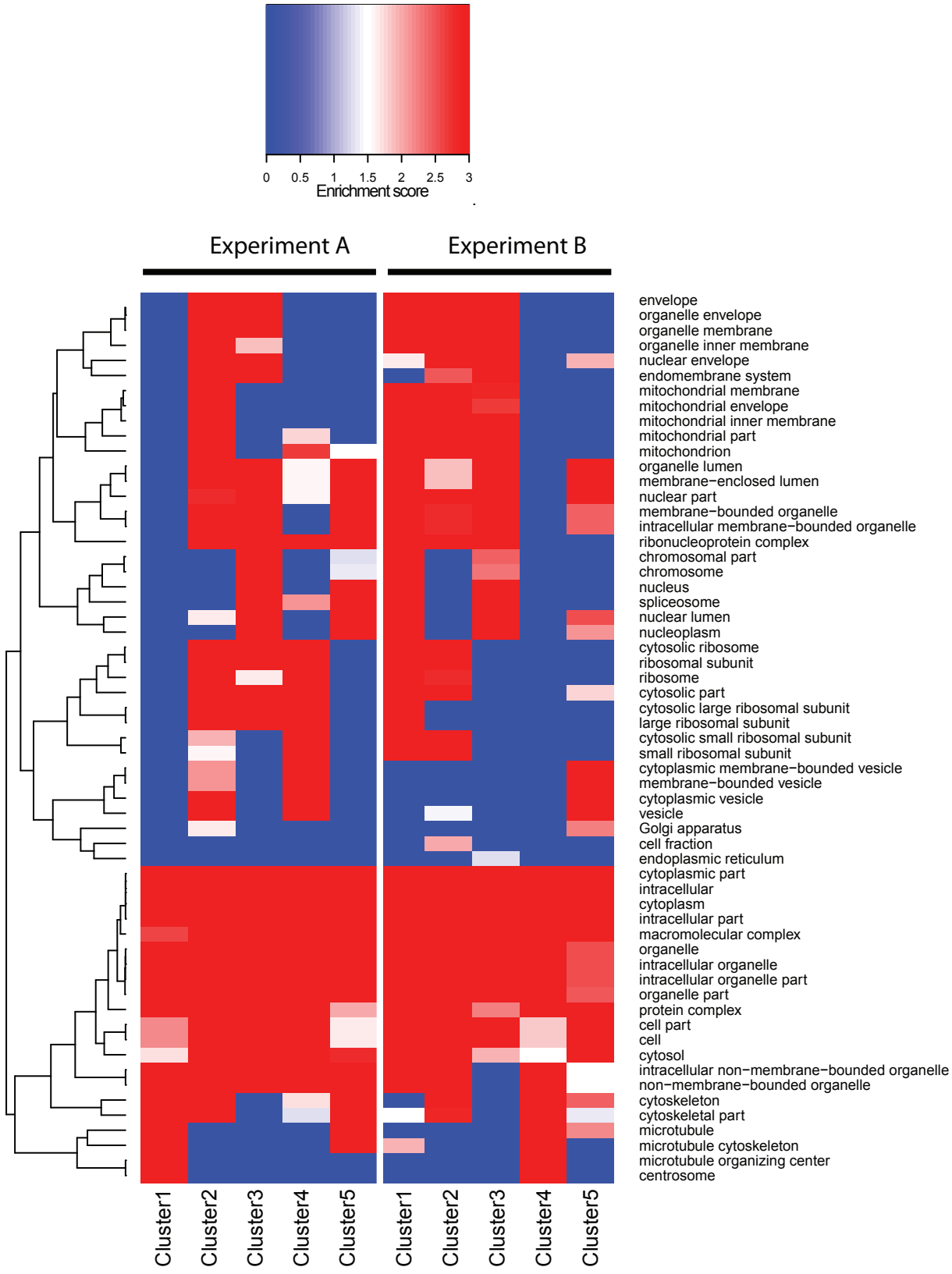
A



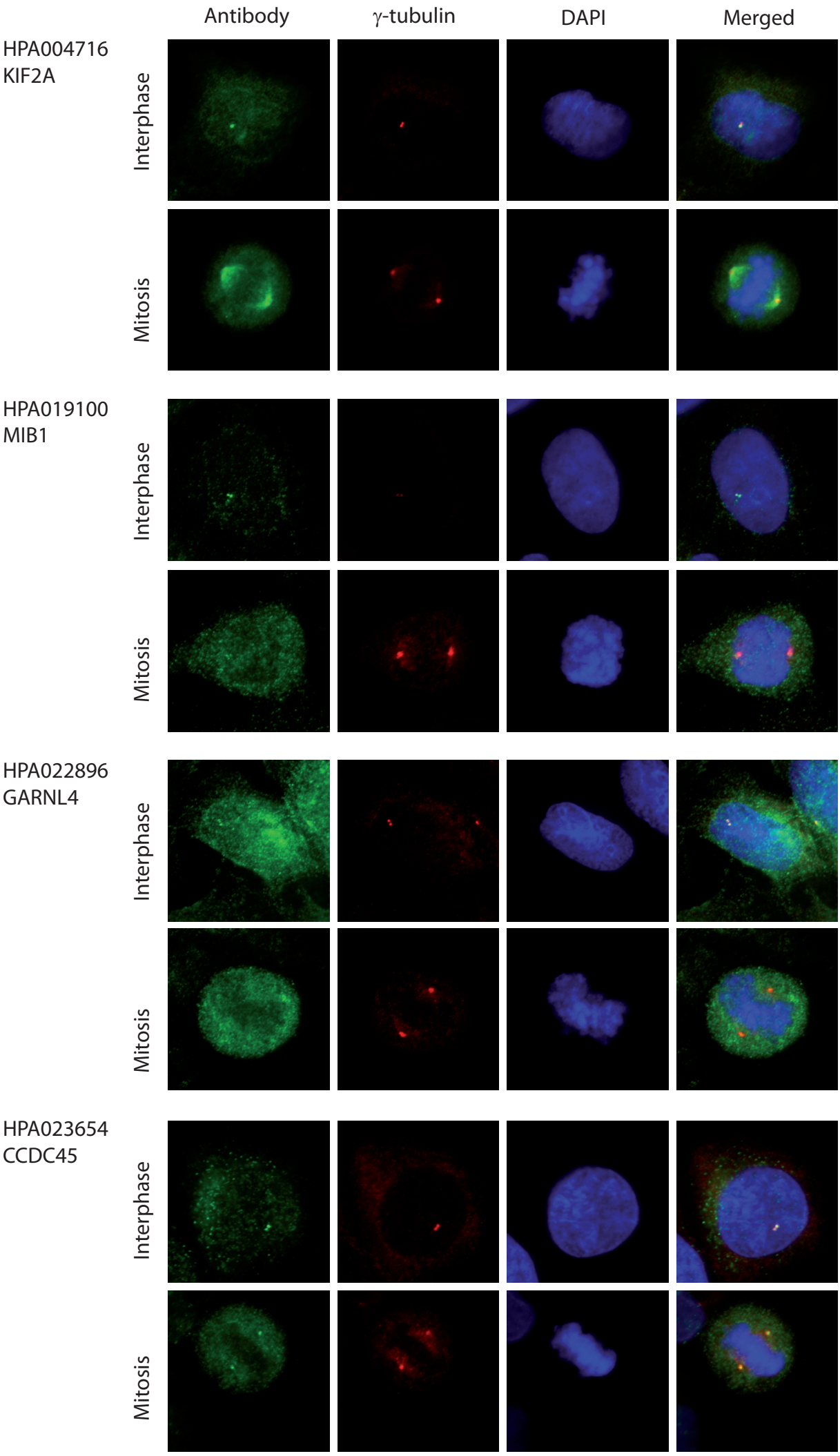
B



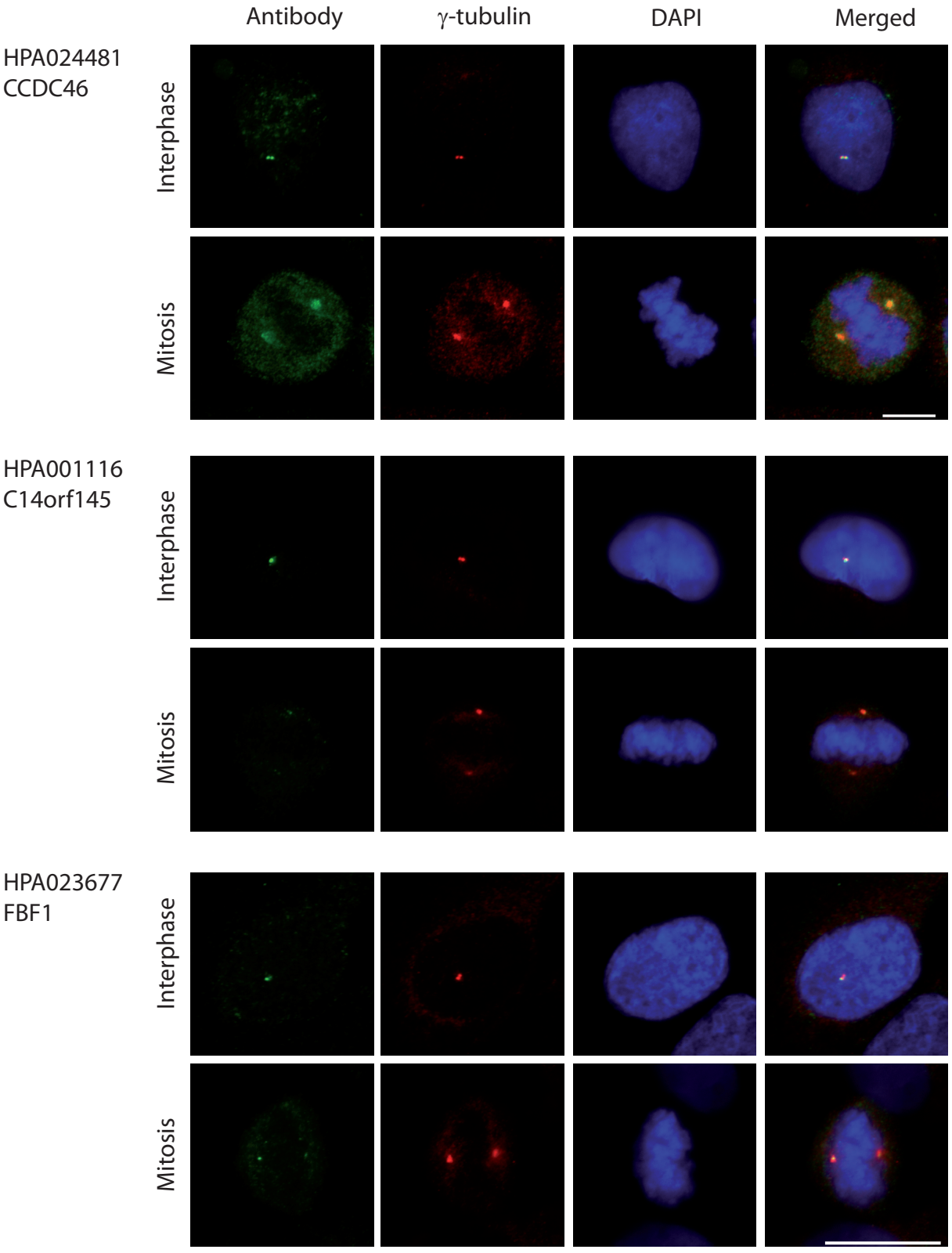
Supplemental Fig. 4C



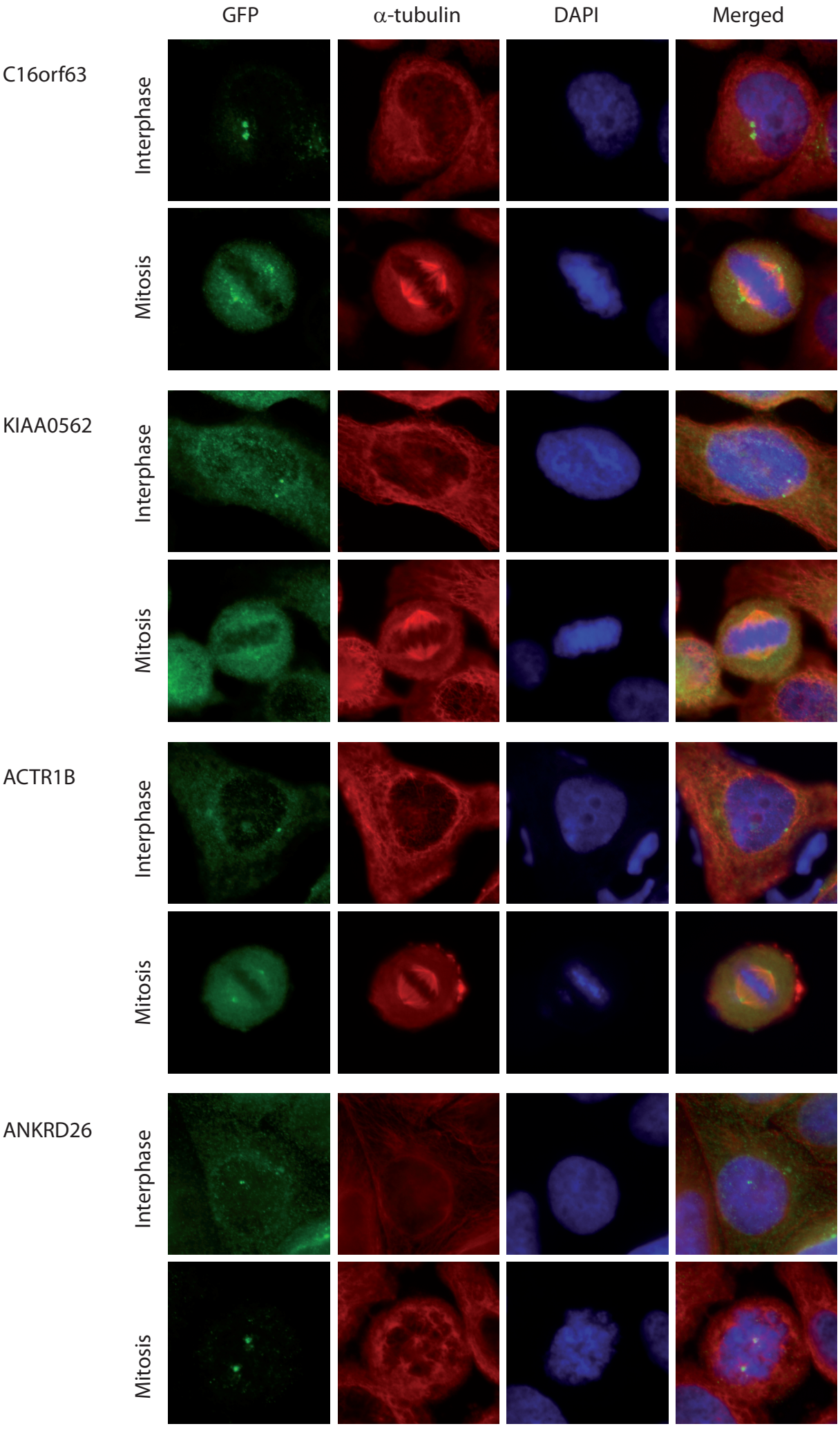
Supplemental Fig. 5A

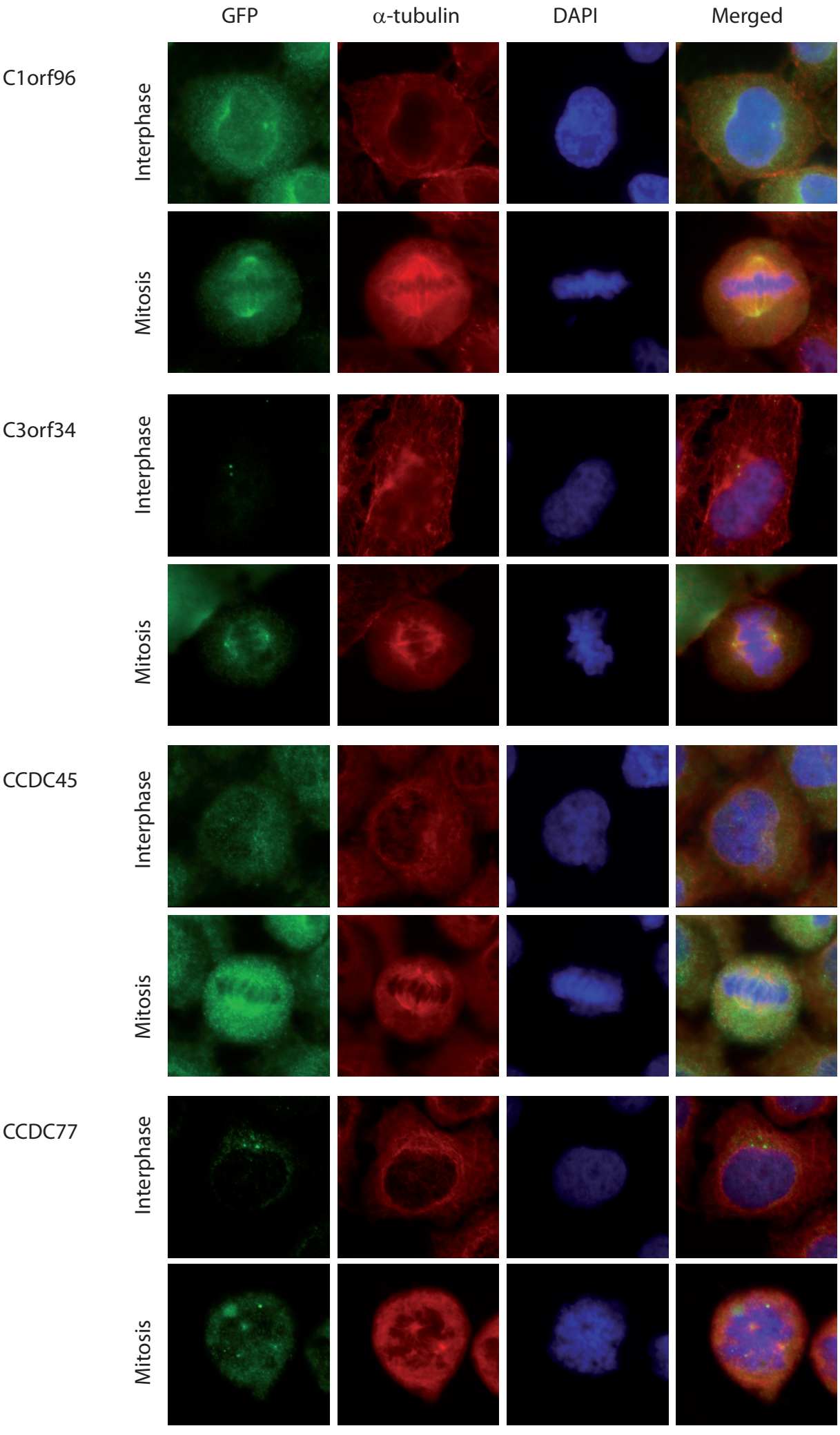


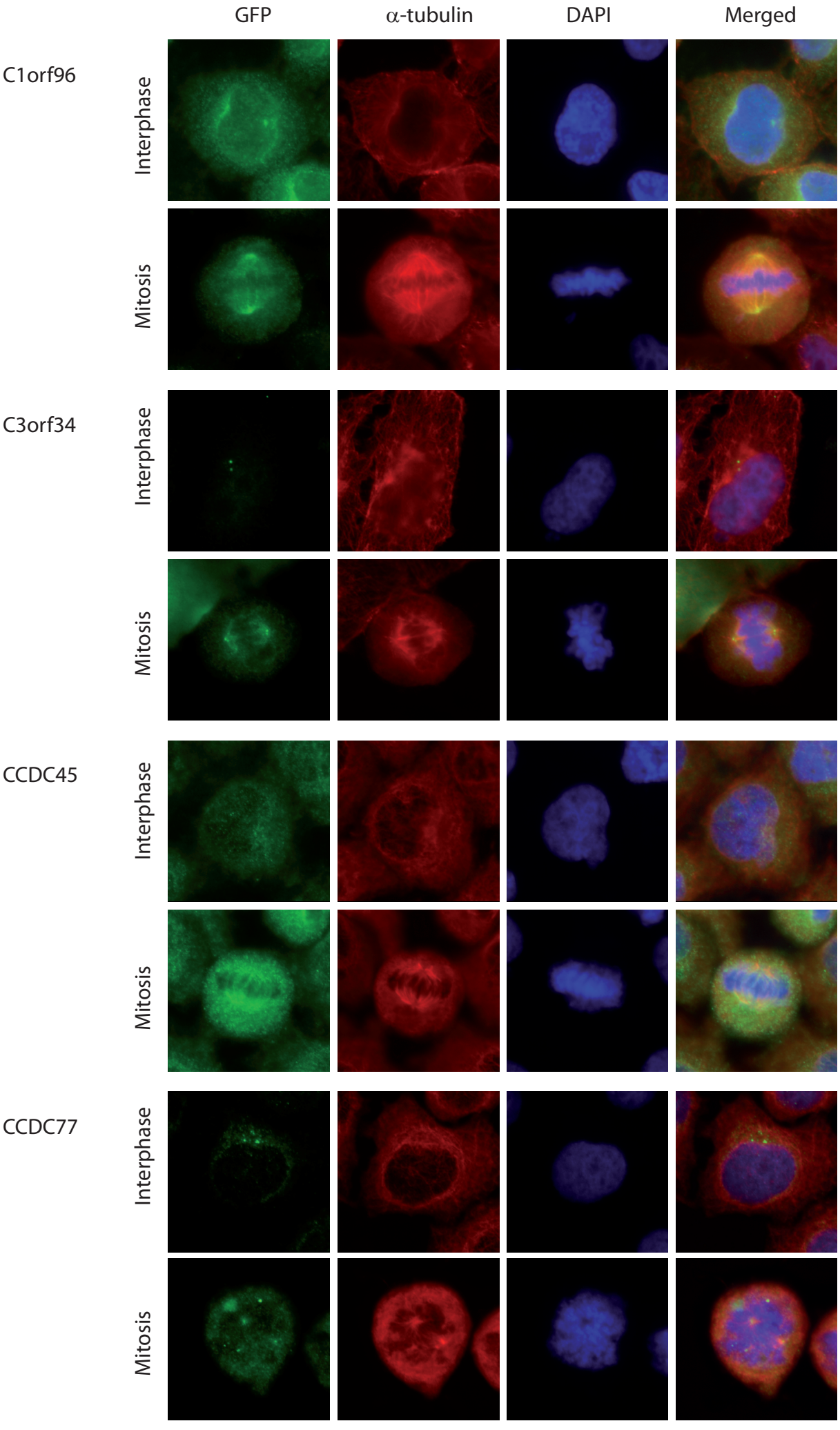
Supplemental Fig. 5A

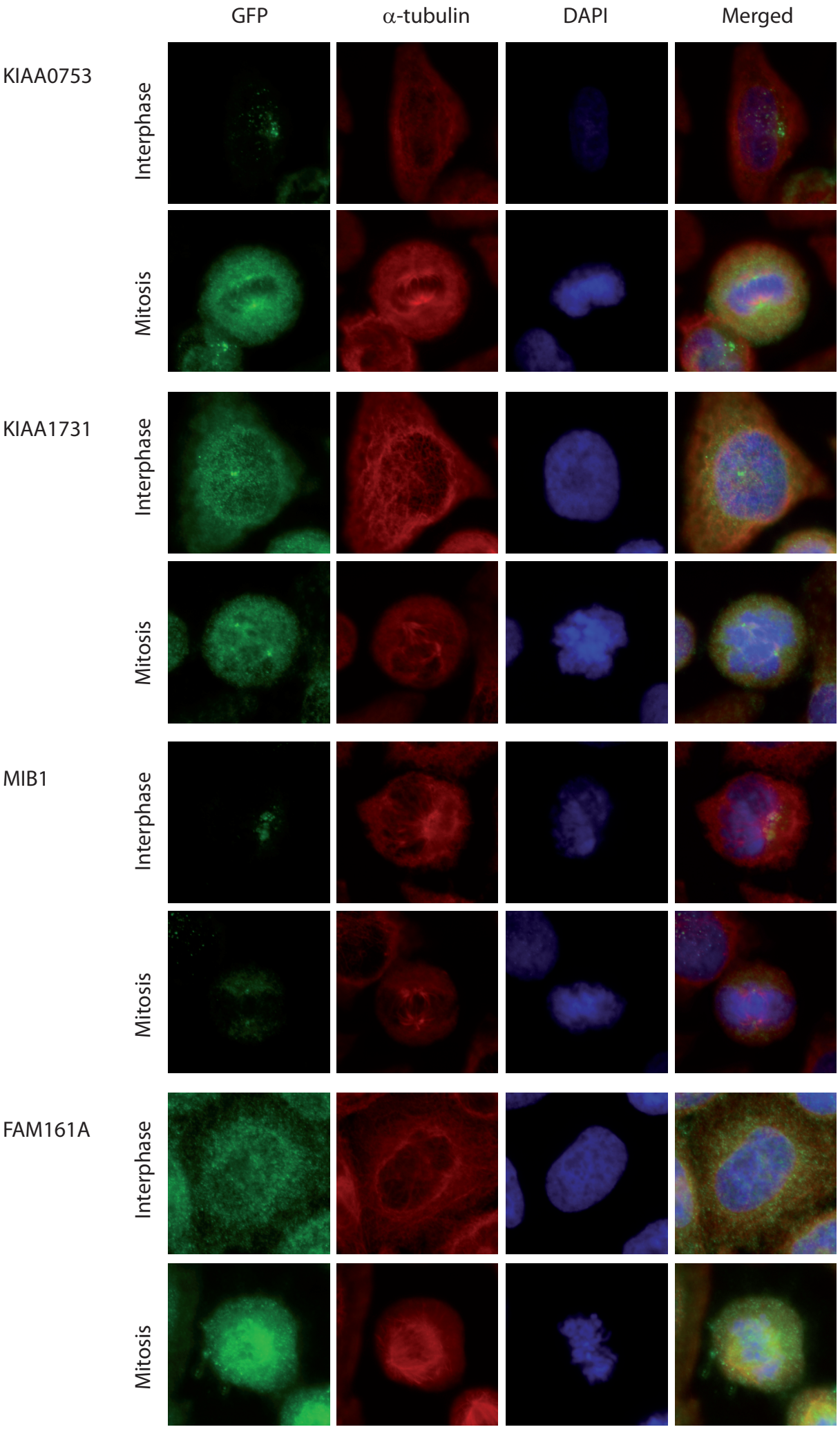


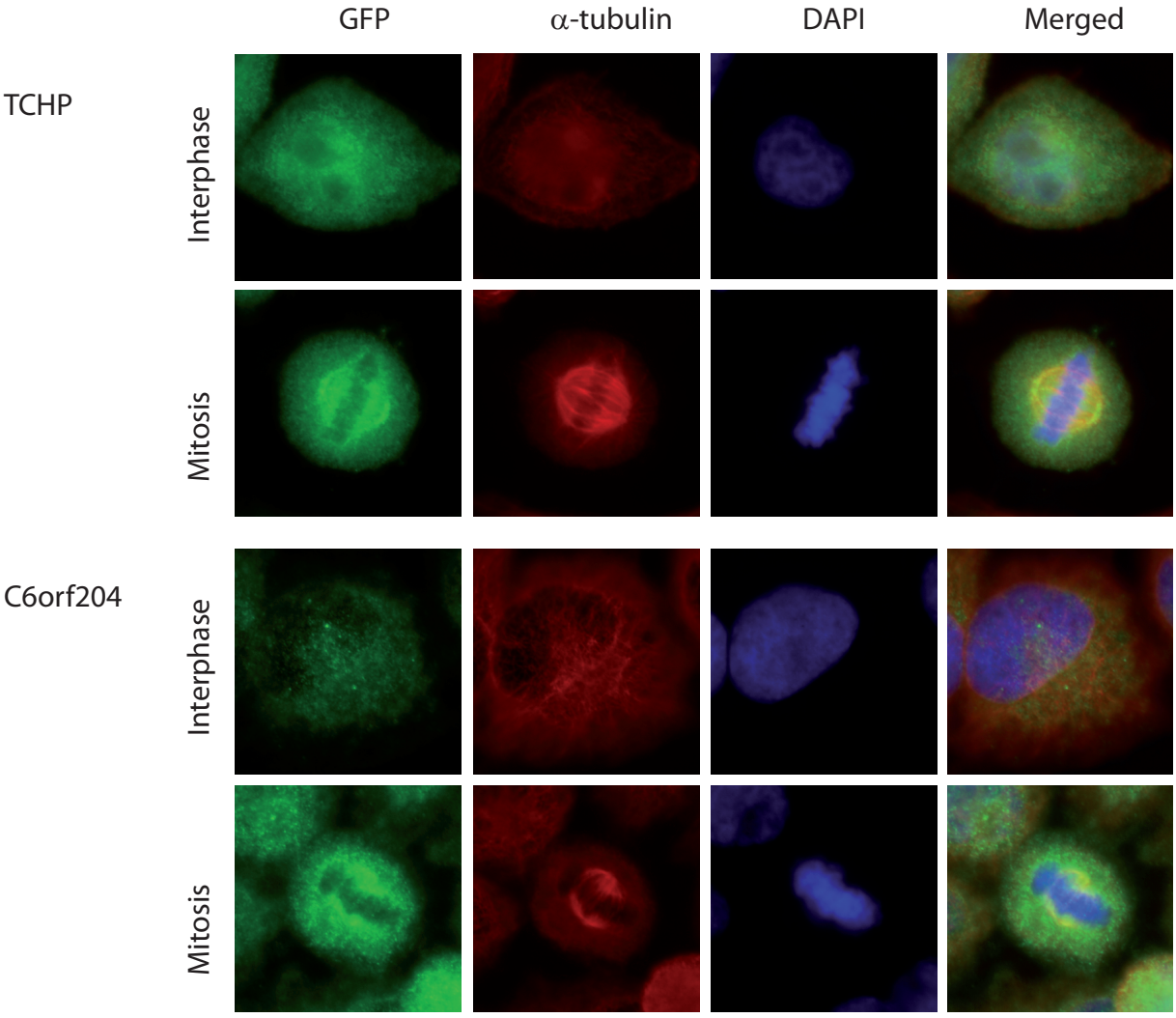
Supplemental Fig. 5C




GFP
α-tubulin
DAPI
Merged
C1orf96
Interphase
Mitosis
C3orf34
Interphase
Mitosis
CCDC45
Interphase
Mitosis
CCDC77
Interphase
Mitosis

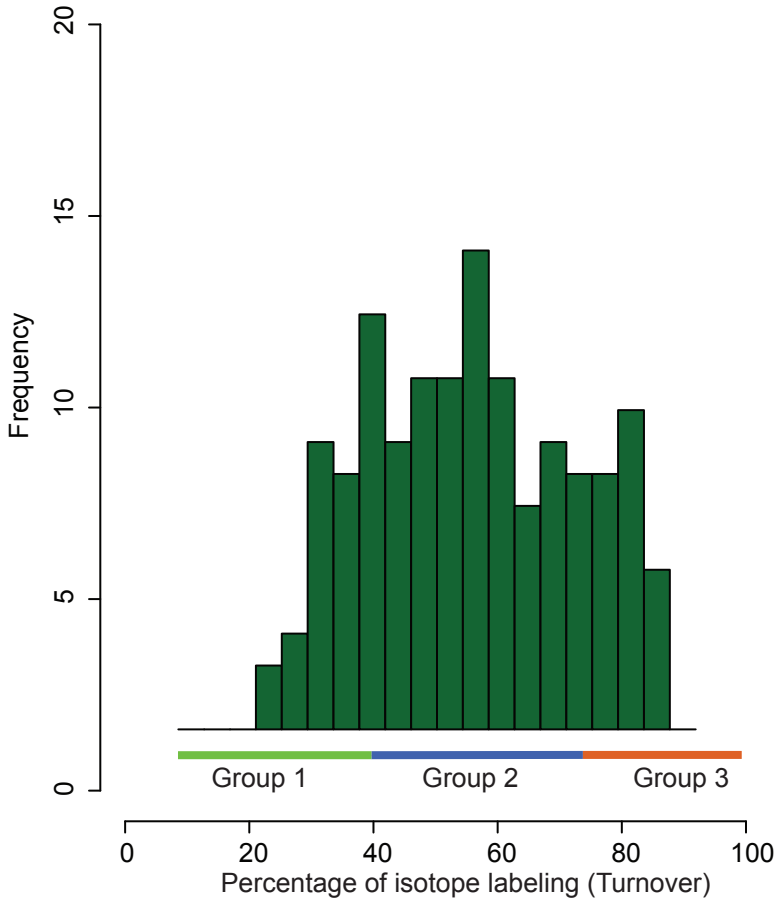

GFP
α-tubulin
DAPI
Merged
C1orf96
Interphase
Mitosis
C3orf34
Interphase
Mitosis
CCDC45
Interphase
Mitosis
CCDC77
Interphase
Mitosis


GFP
α-tubulin
DAPI
Merged
KIAA0753
Interphase
Mitosis
KIAA1731
Interphase
Mitosis
MIB1
Interphase
Mitosis
FAM161A
Interphase
Mitosis

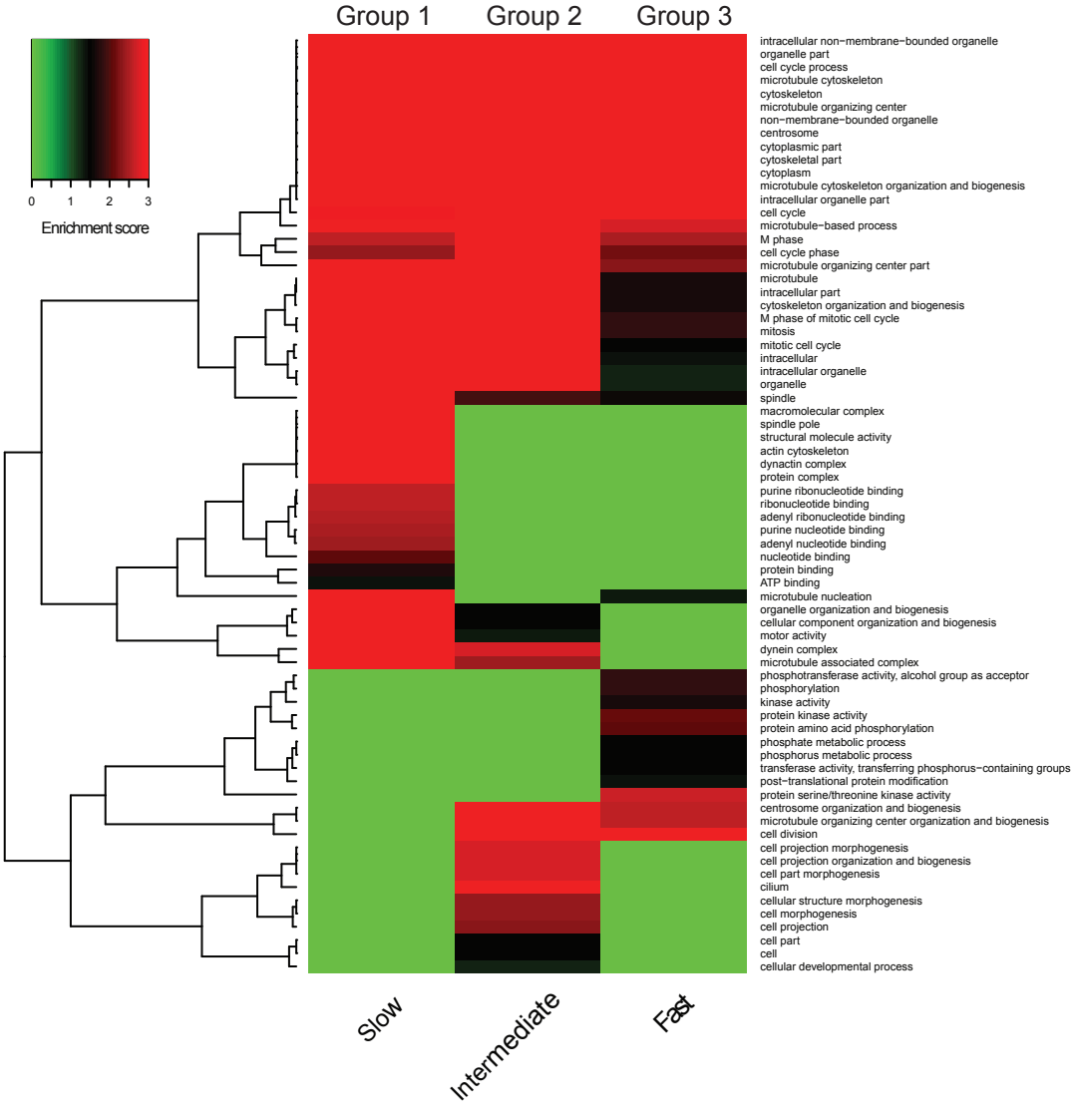

GFP
α-tubulin
DAPI
Merged
TCHP
Interphase
Mitosis
C6orf204
Interphase
Mitosis

Supplemental Fig. 6

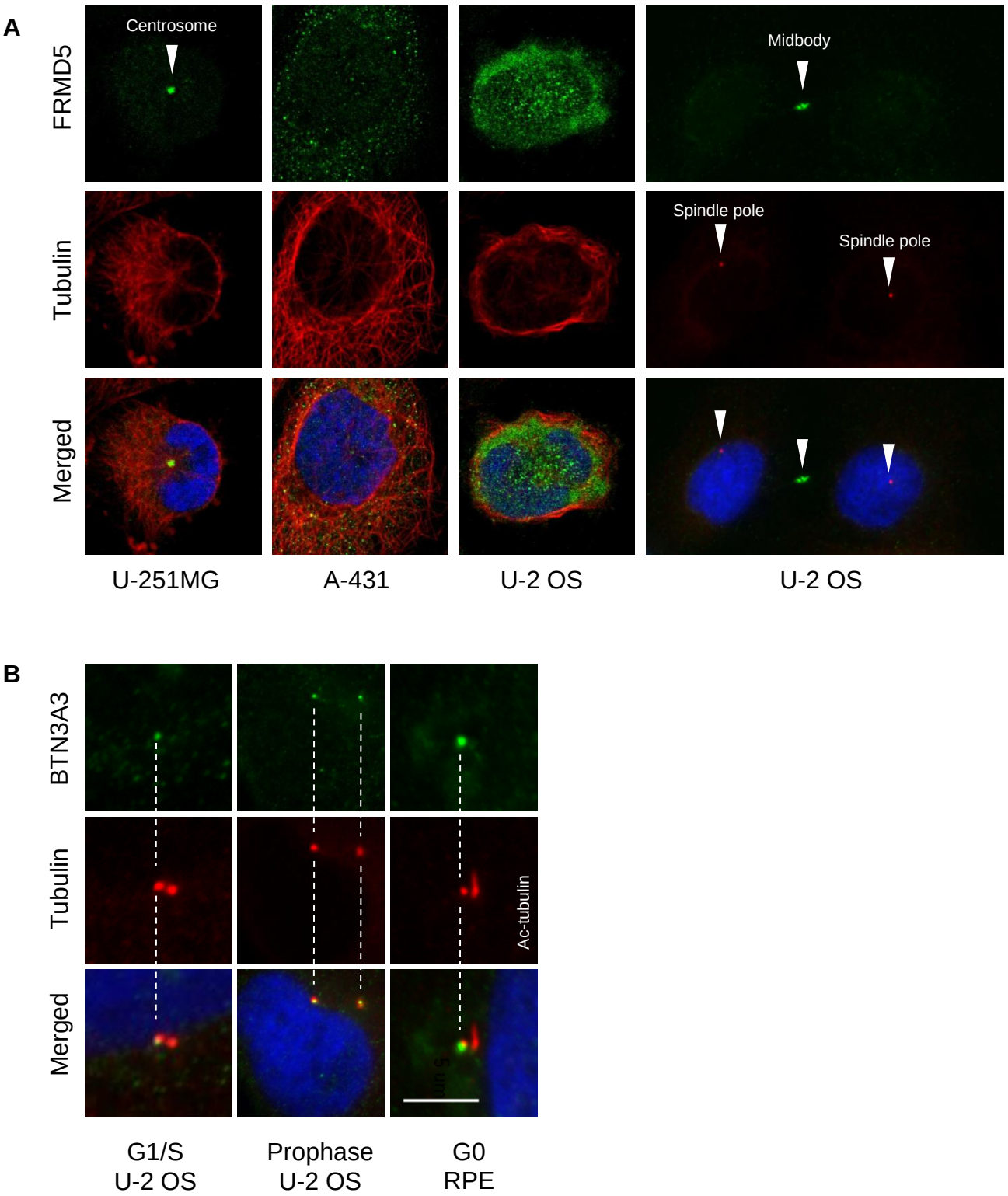
A



B



Supplemental Fig. 7



Supplemental Fig. 8