#### **Supplemental Data:**



#### **Supplemental Information 1: Tilting of kMT axis with respect to the spindle axis**

The average tilt in the axis of a kMT with respect to the spindle axis was estimated from a published serial section EM reconstruction of a metaphase budding yeast spindle (Fig. S1a). The average of angles for all the kMTs from this reconstruction was found to be  $6.3 \pm 5^{\circ}$  (Fig. S1b). This angle would result in a negligible underestimation (0.001) in distances measured along the spindle axis.

We found that the average separation between the two kinetochore clusters in a strain expressing Nuf2p-GFP & Ndc80p-tdTomato from our metaphase dataset is  $1.05 \pm 0.26$  µm (N = 75) and  $5.01 \pm 1.07$  $\mu$ m (N = 80) for the anaphase dataset. These values are typical for all of our data, and they correctly predict the expected spindle length (assuming a mean kMT length of  $\sim$  300 nm) from previous experiment. We can thus very reliably distinguish metaphase cells from cells in early/late anaphase based on the position of the kinetochore spots and the distance separating the two spots within a cell.





The accuracy in centroid localization for an imaging and image acquisition setup is determined by the signal-to-noise ratio of the image. To estimate the centroid localization accuracy for the budding yeast kinetochore clusters in metaphase, we first determined the average GFP and tdTomato signal and background values for strains containing Ctf19p-GFP, Ndc80p-tdTomato, and Ndc8op-GFP. Ctf19p-GFP represents kinetochore proteins with low copy numbers, and also a higher background signal, whereas Ndc80p represents proteins with high copy number per kinetochore. Ndc80p-tdTomato (or Nuf2ptdTomato) was used as the reference point on the kinetochore for all the measurements. For each strain, the maximum signal intensity as well as the standard deviation of the background (from intra-nuclear regions) was determined. The spread of the intensity distribution in the image plane was determined through 1-D Gaussian curve fitting. Images of kinetochore clusters were then simulated by randomly distributing fluorophores in an area of appropriate size, and then convolving this object space with a 2-D Gaussian point spread function. For each signal and background condition, 100 image pairs were simulated such that one image had the centroid of kinetochore distribution in the object space displaced by a known, fixed amount with respect to its partner image. The images included background signal containing noise with Gaussian characteristics (modeled after the camera dark noise) and noise due to

Poisson statistics (shot noise dependent on signal magnitude). Such image pairs were then analyzed by least squares fitting the intensity distribution with a 2-D Gaussian surface. The standard deviation in the measured separation between the centroids of each image pair was designated as the expected precision and plotted as a function of the brightest pixel value within the spot (the dotted lines mark the approximate range for our measurements). The results show that for the lowest signal-to-noise ratio conditions (Ctf19p-GFP), the expected precision in determining the distance is  $\sim 10$  nm. Due to a more compact distribution of kinetochores in anaphase cells, the signal-to-noise ratio improves in anaphase.

**Effect of the spatial staggering of kinetochores on centroid localization accuracy**



Nanometer-scale distances can be accurately measured along the axis of the kinetochore, even if the kinetochores themselves are staggered along the spindle axis as shown in the schematic at the top (variation in kMT lengths in metaphase is less than 150 nm; ref. [1]). The accuracy of such measurements directly depends only on the accuracy with which the location of the centroid of the image of a cluster of kinetochores can be determined, and not directly on the staggering of the kinetochores. The extent of kinetochore staggering can have an indirect effect, since for large spatial staggering, the image of a kinetochore cluster does not appear as one spot as shown in the bottom figure. To illustrate this point, we simulated images of kinetochore clusters by randomly dispersing only six kinetochores over a square area of the indicated dimension. Each kinetochore is represented by a pair of green and red fluorophores separated by a distance of 56 nm – the length of the Ndc80 complex molecule. An overlay of the images made by such a cluster of six kinetochores is shown in the middle panel. The bottom panel displays an overlay of the 2-D Gaussian fits obtained for the simulated images. These representative images clearly show that as the staggering of the kinetochore cluster exceeds 200 nm, the corresponding microscope image acquires a distinct structure with multiple peaks, instead of the single peak seen for the first two cases. The least squares fitting routine for determining the centroid location with nanometer accuracy described in the previous section makes the assumption that the observed intensity distribution can be described by a 2-D Gaussian function. The last two images show that this key assumption breaks down for kinetochore staggers over a 200x200 nm region, which is much larger than the expected spatial staggering of the sixteen kinetochores in a metaphase as well as anaphase spindle is much smaller. Therefore, the accuracy with which distances can be determined is unaffected by the staggering of the kinetochore along the axis of the budding yeast spindle.

## **Supplemental Data 3: Image registration procedure**



Registration between the EGFP and tdTomato images was carried out prior to each experiment as follows [2]. Tetraspek microspheres (diameter = 100 nm, cat. # T-7279, Invitrogen) were suspended in 5% poly-acryl amide gel (refractive index ~ 1.37) and sandwiched between a slide and a coverslip. Single, isolated beads were selected, and scanned over a 150x150 pixel image area in the center of the field of view with a spacing of 1  $\mu$ m or 1.5  $\mu$ m using an XY translation stage (MS-2000, ASI, Eugene, OR). An image of the bead was recorded at each position. Irregular movements of the XY translation stage resulted in a slightly irregular displacements of the bead. Stage drift along the optical axis over the time of acquisition was minimized through the use of Perfect Focus system (Nikon) to maintain the bead

in focus. The image acquisition time for each image ranged from 600-1800 ms. Acquisition time was selected so as to maintain a signal to noise ratio  $>$  30 in both imaging channels. The top panel in this figure shows the maximum intensity superposition of 240 image planes obtained in this manner. The EGFP and tdTomato images are high-lighted with green and red squares respectively. The high signal to noise ratio ensures centroid localization accuracy approaching 1 nm for each bead image along both X and Y axes using 2-D Gaussian curve fitting. The image transform for registration between the EGFP and tdTomato fields was computed using the local weighted mean algorithm implemented in MatLAB (MathWorks, Natick, Md). The use of Dual-View attachment to acquire the EGFP and tdTomato channels simultaneously introduces spatially varying offsets between the centroids of the EGFP image of a bead and the corresponding tdTomato image (displayed as vectors), which are apparent in the left middle panel. A histogram of the absolute offset between the centroids of an EGFP spot and the corresponding tdTomato spot prior to registration is displayed in the bottom left panel. This variation along with chromatic aberrations is corrected by the image registration procedure (vector map in the right middle panel, and a histogram of corresponding offset magnitude). The average offset magnitude after image registration was found to range from 3-6 nm (mean offset magnitude of  $3.7 \pm 1.9$  nm in this particular example). This procedure was carried prior to or just after each experiment. Each experiment lasted for 30-40 minutes to minimize effects of drift during the experiment.

The distance measurements were conducted with a custom interactive graphical user interface written in MatLab. In each experiment, the tdTomato image stack was first registered with respect to the corresponding EGFP stack based on the image transform obtained as discussed above. Two kinetochore clusters (distinguished as metaphase and anaphase cells) in each cell were manually selected using the EGFP image. The region of interest was then automatically defined for the selected kinetochore cluster as described in the methods section. The same region was then transferred to the corresponding tdTomato image. The centroids of these regions were then determined by 2-D Gaussian fitting, and the data was appropriately stored.

6

### **Supplemental Data 4: Residual chromatic aberrations & spindle orientation**

Chromatic aberrations can introduce a systematic bias in the distance measurements. Although the image registration procedure described in the previous section will also correct for chromatic aberrations. Generally, however, the bead chosen for image registration procedure is not at the same depth as the average depth at which kinetochore clusters are imaged. A small bias due to residual chromatic aberrations may still be expected. We offer two data that demonstrate that the bias due to chromatic aberrations is negligibly small in our data. The first evidence is the measurement of the distance between to SPB proteins carried out in two strains that have the color of the fluorescent label switched. The random orientation of the spindle axis within the image plane in any data set will also average out the directional offset between the red and green PSFs resulting from chromatic aberrations. The average value of the angle subtended by the spindle axes in the set for a given distance measurement is  $0 \pm 40^{\circ}$ .

#### **Supplemental Data 5: Effects if tilting of kinetochore axis along the optical axis**

The separation between the EGFP and tdTomato centroid determined in the "in focus" plane is projection of the true distance on the image plane. Since the long axes of all sixteen kinetochores are closely aligned with the spindle axis, the orientation of the spindle axis can be used to deduce the average tilt in the kinetochore axes. In this study, we analyzed spindles in which, the centroids of the two kinetochore clusters were separated by less than 600 nm along the optical axis. We compensate for the spindle axis tilt along the optical axis by dividing the measured separation with the cosine of the angle that the spindle axis (as determined by the centroids of sister kinetochore clusters) makes with the image plane.

$$
\theta_{\text{tilt}} = a \tan\left(\frac{L_{K-K}}{z}\right)
$$

 $z = 100 +$  separation between the two *in-focus* planes along the optical axis (nm)  $L_{K-K}$  = Separation between two EGFP centroids in the image plane (nm) , *where*

and,

$$
D_{\text{true}} = D_{\text{measured}} / \frac{D_{\text{true}}}{\cos \phi_{\text{tilt}}} \text{ where,}
$$

 $D<sub>true</sub>$  = the true separation between the EGFP and tdTomato centroids

 $D_{measured}$  = the measured separation between the EGFP and tdTomato centroids

The kinetochore axes have yet another degree of freedom, which is the curling of the kinetochore axis along the optical axis, and it can't be detected. If we assume that the curling of the kinetochore axis within the image plane and along the optical axis are the same, we can estimate its effect from the measured curling of the kinetochore axis within the image plane. This angle is given by the angle of the kinetochore axis with the spindle axis in the image plane. We chose the full-length measurements for the

NDC80 complex for two reasons – the complex is exists as a single macro-molecule, and the distance is larger than the uncertainty in centroid determination. The average value of this angle for the entire dataset was  $23 \pm 16^{\circ}$  (cosine of this angle is 0.92). If the kinetochores have a similar tilt along the optical axis, all of our measurements would have to be increased by  $\sim 8\%$ . E.g. the full length of the NDC80 complex with the correction would be ~ 57 nm instead of the measured length of 55 nm. However, because of the absence of direct measurements, and because of the small magnitude of the underestimation involved, we did not apply this correction to our measurements.

#### **Supplemental Data 6: Maximum likelihood estimation of the mean centroid separation**

Measurements of the distance separating two centroids that are defined by Gaussian probability distributions have been shown to be distributed according to a 2-D non-Gaussian probability density function given by [2]:

$$
P_{2D}(r) = \left(\frac{r}{\sigma^2}\right)^* \exp\left(\frac{\mu^2 + r^2}{2\sigma^2}\right)^* I_0\left(\frac{r\mu}{\sigma^2}\right) \tag{1}
$$

Where,

 $r =$  measured magnitude of distance separating the two centroids

 $\sigma$  = expected standard deviation for each centroid location measurement (experimental errors)

 $\mu$  = true distance separating the two centroids

*I<sup>0</sup>* = modified Bessel function of order zero

This is an asymmetric probability distribution skewed towards larger values of *r*. A maximum likelihood method was devised for estimating the true distance that is based on fitting the observed probability distribution of distance measurements with the theoretical probability distribution given by equation (1). The ability to accurately measure the distance separating the two centroids depends on the relative magnitudes of the true distance being measured  $(\mu)$  and the expected uncertainty in centroid determination (σ). It follows that for correct determination of using this method, the measurements of the two centroids must be normally distributed. We have implemented this method for our measurements of the centroids that reflect the centroid of a cluster of fluorophores rather than single molecules. The implicit assumption is that the centroid localization errors will be normally distributed as in the single molecule case.

#### **Supplemental Data 7: Imbalanced fluorophore maturation and the limit of resolution**

Centroid separation measurements for kinetochore clusters contain an additional source of variance besides the uncertainty in centroid determination, residual errors from image registration, and biological variation in the kinetochore structure. Each diffraction-limited spot represents a cluster of 16 kinetochores that are spread over an area, and kinetochore each carrying multiple copies of EGFP and tdTomato. Furthermore, the maturation efficiency for fluorescent proteins is less than 100%. Maturation efficiency has been estimated [3] at 90% for EGFP. Since the maturation half-life for tdTomato is  $\sim$  3fold higher than that for EGFP, its maturation efficiency can be expected to be lower [4]. These two factors combine to introduce a significant error in individual distance measurements. The effective error in each measurement is:

$$
\sigma_{\text{Total}}^2 = \sigma_{\text{centroid}}^2 + \sigma_{\text{resistration}}^2 + \sigma_{\text{maturation}}^2 \tag{2}
$$

 $\sigma_{\textit{centroid}}$  $=$  uncertainty in centroid determination ( $<$  10 nm, Supplemental Data 3)  $\sigma_{\textit{reeistration}}$  $=$  average residual error after image registration ( $\sim$  6 nm, Supplemental Data 4) *maturation* = error due to differential fluorophore maturation within a kinetochore cluster

Centroid determination error and error due to registration account for only 50% of the measurement variance. Exact values for fluorophore maturation are unknown. The distribution of kinetochores within the cluster is less than 150 nm.

The magnitude of the effective measurement error is also the smallest distance that can be measured using the maximum likelihood method described in the previous note. The maximum likelihood method only considers the probability distribution of EGFP-tdTomato centroid vector magnitudes for a

given pair of labeled kinetochore proteins. In cases where the actual separation between two protein domains is smaller than the measurement error, the maximum likelihood method will therefore estimate the mean value of the measurement error rather than the expected distance. E.g. the C-termini of Ndc80p and Nuf2p are separated by an 80 amino acid tail on Ndc80p (ref. [5]). We localized the C-termini of Ndc80p and Nuf2p with respect to the C-terminus of Ctf19p that is far removed from these protein domains, and hence can be accurately determined by the maximum likelihood estimation method. We found that in anaphase, the distance separating the C-termini of both Ndc80p and Nuf2p from the Ctf19p C-terminus is 24 nm and 25 nm respectively, whereas in metaphase, these distances were 37 nm and 33 nm respectively. The difference between these separations provides the separation between the C-termini of Ndc80p and Nuf2p: -1 nm in anaphase and 4 nm in metaphase. These values compare well with the separation of Ndc80p and Nuf2p C-termini projected along the spindle axis ( $1 \pm 5$  nm; N=160 in anaphase and  $6 \pm 11$  nm; N=98 in metaphase). In contrast, the maximum likelihood estimation method predicts an anaphase separation of 10 nm between these two protein domains (metaphase data does not generate a satisfactory fit), which can be used as an estimate of the measurement error for our technique. We thus estimate the resolution limit for our technique to be approximately 10 nm.



# **Supplemental Data 8: Comparison of raw means and maximum likelihood estimates**

\* - Metaphase localizations for these proteins could not be determined reliably using the non-Gaussian probability distribution fits due to high variation in the data relative to the distance being measured. The value cited in these cases is the separation *along the spindle axis.*

Negative signs for distances were assigned to indicate that the distance inside (towards the centromere) of the C-terminus of Ndc80p or Nuf2p. Positive sign implies that the protein localizes on the outside (away from the centromere) of this reference point.

Confidence limits obtained from maximum likelihood estimation are typically asymmetric [6]. The 68% confidence interval cited above is the larger limit of the two obtained from maximum likelihood estimation.



# **Supplemental Data 9: Protein localizations in metaphase and anaphase cells**

\* - Metaphase localizations for these proteins could not be determined reliably using the maximum likelihood method due to high variation in the data relative to the distance being measured.

68% confidence limit cited is the larger limit obtained from maximum likelihood estimation.

Negative signs for distances were assigned to indicate that the distance inside (towards the centromere) of the C-terminus of Ndc80p or Nuf2p. Positive sign implies that the protein localizes on the outside (away from the centromere) of this reference point.

**† -** Although the distances were measured with respect to Ndc80p-C, Fig. 3 was constructed by using Spc24p-C as the point of reference. These columns display each distance from Spc24p-C. The metaphase values was carried out by subtracting 17 nm from the original measurement, while the anaphase values were obtained by subtracting 12 nm.

#### **Supplemental Data 10: Estimation of the systematic underestimation in anaphase measurements**

The release of tension in anaphase can be expected to provide more freedom to the orientation of the two SPBs that nucleate spindle microtubules. Kinetochore-microtubules in anaphase cells are only 60 nm in length on average [1], and can therefore be assumed to be perpendicular to the face of the SPBs. The increased rotational freedom in SPB orientation along the Z-axis will introduce a systematic underestimation in distance measurements for the kinetochore. To estimate this distance, we carried out distance measurements in a strain that had two SPB proteins: Spc72p and Spc110p labeled at their Ctermini with EGFP and tdTomato respectively. We found that the metaphase distance separating these two ends was 75 nm, whereas the anaphase distance was found to 70 nm. If we ascribe this decrease in SPB dimension to the greater orientations freedom for the SPB face along the optical axis, we can expect an underestimation of no more than 7% in anaphase measurements.

#### **Supplemental Data 11: A case for centromeric nucleosome vs. hemisomes in budding yeast**

A recent study of the conformation of chromatin using chromatin immunoprecipitation against CENP-A in *Drosophila* cells demonstrated that the CENP-A bound DNA is packed as hemisomes containing one molecule each of CENP-A, H2A, H2B and H4 (ref. [7]). Several lines of evidence suggest that the budding yeast centromere may not follow this pattern. (1) Histone depletion experiments [8] show that the depletion of histone H3 leaves the nucleosomal compaction within centromeric chromatin unaltered, whereas depletion of the histone H4 or Cse4p disrupt the wild-type pattern of nucleosomal compaction. (2) Fitzgerald-Hays et al. created two mutant alleles of Cse4p – one with a mutation within the histone-fold-domain, and the other with a mutation in the N-terminal tail of the protein [9]. Both these alleles were non-functional, and hence lethal to the cells. The authors found that when co-expressed in the same strain, the two mutant alleles could function together to restore wild-type Cse4p function. This synthetic restoration of protein function is possible only if the two mutant alleles dimerize, so that one could provide the essential function of the HFD, whereas the other monomer could provide the function of the N-terminal tail. (3) The length of nuclease resistant centromeric DNA (220 base pairs long) and the recent finding that Cse4p localization is limited only to the centromeric nucleosome indicate that the kinetochore is likely supported by a single centromeric nucleosome containing Cse4p (or two hemisomes, ref. [10],[11]). Based on these findings, we use a representation of a nucleosome in the model 3-D drawing of the yeast kinetochore-microtubule attachment.



**Supplemental Data 12: List of strains & plasmids used in this study**



## **REFERENCES**

- 1. Winey, M., Mamay, C.L., O'Toole, E.T., Mastronarde, D.N., Giddings, T.H., Jr., McDonald, K.L., and McIntosh, J.R. (1995). Three-dimensional ultrastructural analysis of the Saccharomyces cerevisiae mitotic spindle. J Cell Biol *129*, 1601-1615.
- 2. Churchman, L.S., Flyvbjerg, H., and Spudich, J.A. (2006). A non-gaussian distribution quantifies distances measured with fluorescence localization techniques. Biophys J *90*, 668-671.
- 3. Patterson, G.H., Knobel, S.M., Sharif, W.D., Kain, S.R., and Piston, D.W. (1997). Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. Biophys J *73*, 2782- 2790.
- 4. Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nature biotechnology *22*, 1567-1572.
- 5. Wei, R.R., Sorger, P.K., and Harrison, S.C. (2005). Molecular organization of the Ndc80 complex, an essential kinetochore component. Proc Natl Acad Sci U S A *102*, 5363-5367.
- 6. Barlow, R.J. (1989). Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences, (New York: John Wiley and Sons).
- 7. Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric Structure of Centromeric Nucleosomes in Interphase Drosophila Cells. PLoS Biol *5*, e218.
- 8. Saunders, M.J., Yeh, E., Grunstein, M., and Bloom, K. (1990). Nucleosome depletion alters the chromatin structure of Saccharomyces cerevisiae centromeres. Mol Cell Biol *10*, 5721-5727.
- 9. Morey, L., Barnes, K., Chen, Y., Fitzgerald-Hayes, M., and Baker, R.E. (2004). The histone fold domain of Cse4 is sufficient for CEN targeting and propagation of active centromeres in budding yeast. Eukaryot Cell *3*, 1533-1543.
- 10. Bloom, K.S., and Carbon, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. Cell *29*, 305-317.
- 11. Furuyama, S., and Biggins, S. (2007). Centromere identity is specified by a single centromeric nucleosome in budding yeast. Volume 104. pp. 14706-14711.