Bioimage Informatics

Lecture 9, Spring 2012

Bioimage Data Analysis (II):

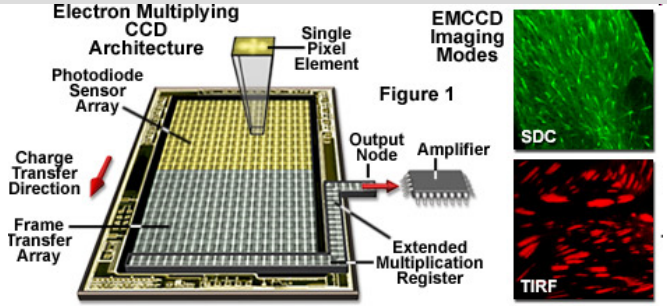
Applications of Point Feature Detection Techniques

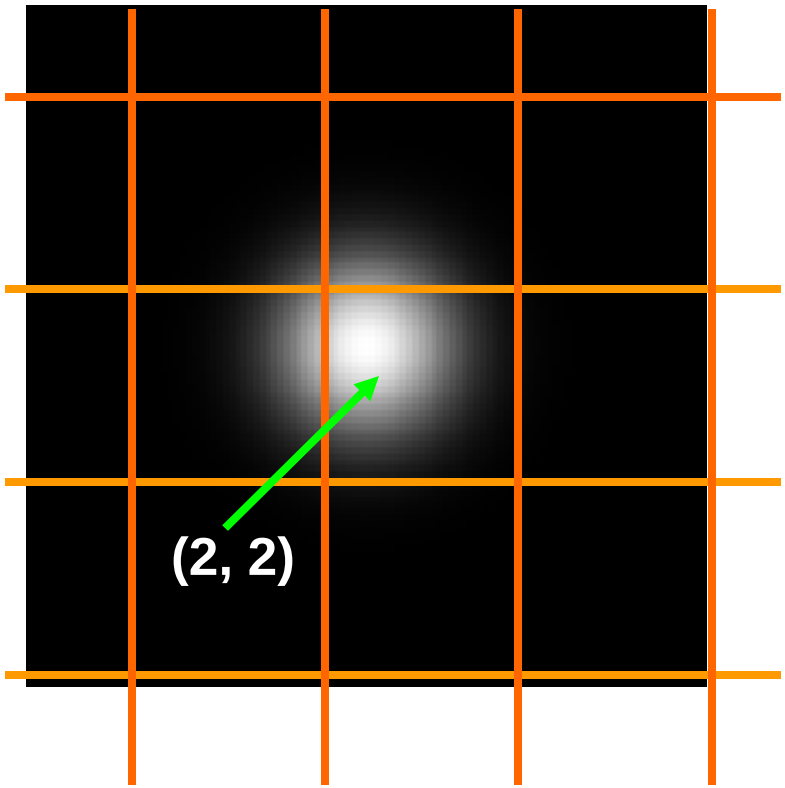
Outline

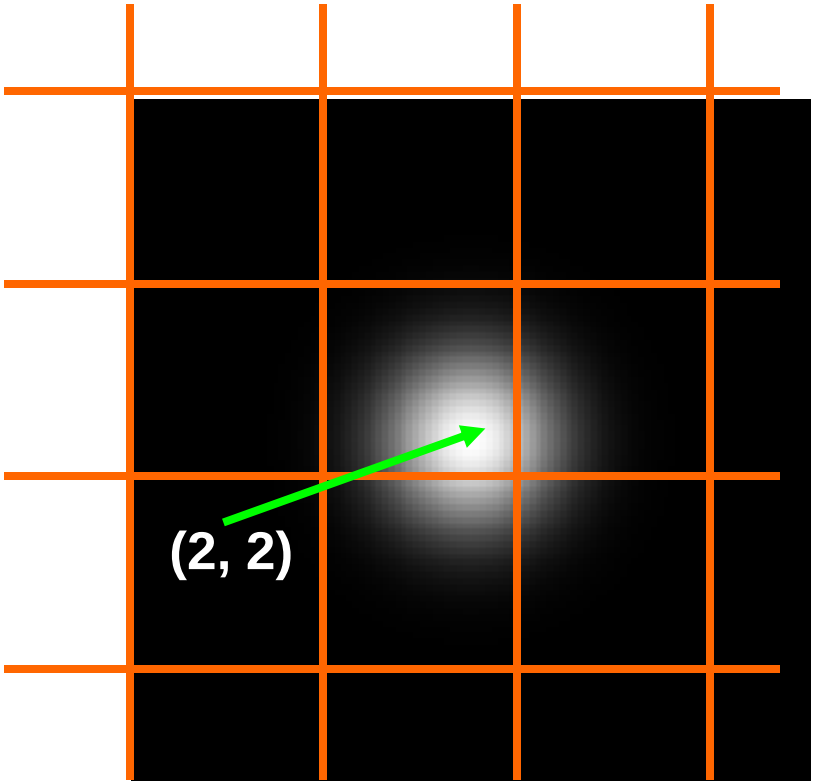
- Review: subpixel resolution point feature detection
- Other point feature detection techniques
- Reproducible research in computational science
- Application I: molecule counting
- Application II: single molecule imaging
- Application III: protein colocalization analysis
- Application IV: super-resolution imaging

- Review: subpixel resolution point feature detection
- Other point feature detection techniques
- Reproducible research in computational science
- Application I: molecule counting
- Application II: single molecule imaging
- Application III: protein colocalization analysis
- Application IV: super-resolution imaging

Sub-Pixel Resolution Point Detection







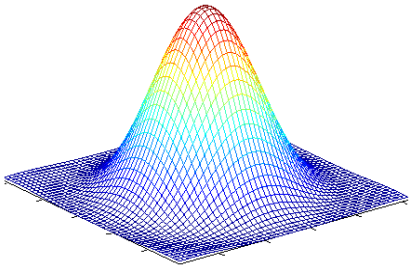
Sub-pixel Detection by Gaussian Fit

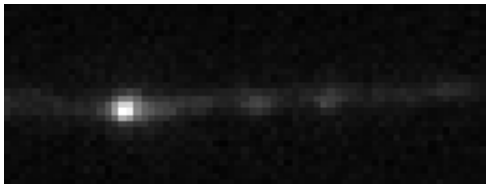
- Fit using a Gaussian kernel, which represents the ideal image of a point

$$K(x,y;x_0,y_0) = K(x-x_0,y-y_0) = A\cdot exp\left[-\frac{(x-x_0)^2+(y-y_0)^2}{B}\right]$$

- Problem formulation: to minimize the difference between the translated kernel and the image

$$\min_{(x_0,y_0)\in R^2} \left| I(x,y) - K(x,y;x_0,y_0) \right|$$





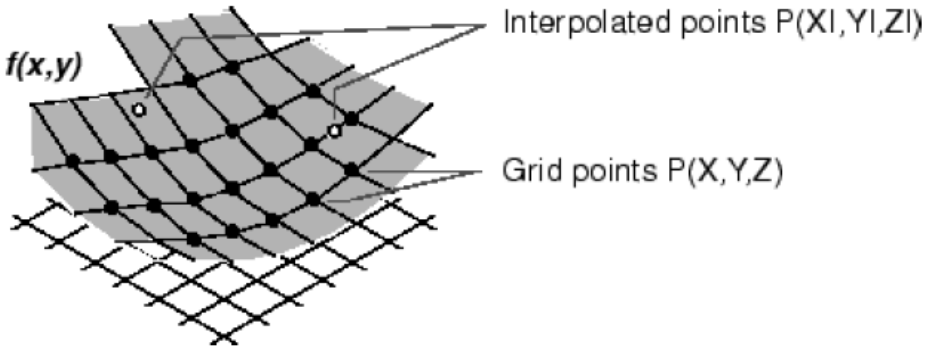
Gaussian Fitting Implementation Details (I)

- Calculation of $\left|I\left(x,y\right)-K\left(x,y;x_0,y_0\right)\right|$

$$E\left(x_0,y_0\right)=\sum_{i=1}^{M}\sum_{j=1}^{N}\left|I\left(i,j\right)-K\left(i-x_0,j-y_0\right)\right|$$

May not be integer coordinates

- Intensity interpolation using interp2



Gaussian Fitting Implementation Details (II)

- How to set the Gaussian kernel

  - By fitting the Airy Disk using a Gaussian

$$\sigma = \frac{0.61 \cdot \lambda}{NA} / 3$$

  - By measuring PSF (often using beads) then fitting with a Gaussian

- Spatial sampling: three-times oversampling of Airy disk

$$\frac{\text{Airy disk radius}}{\text{pixel size}} \geq 3$$

- Under high SNR, spatial sampling may be relaxed to 2~2.5.

Gaussian Fitting Implementation Details (III)

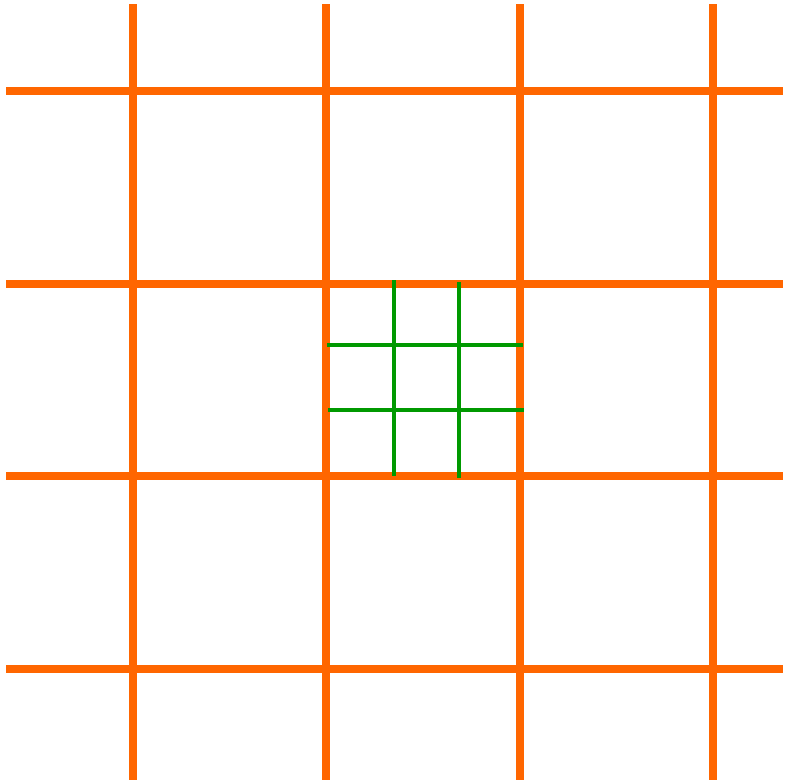
- Optimization strategy I: exhaustive search

- Implementation of exhaustive search

  - Oversampling the kernel and images

  → use a small pixel size: e.g. 10nm

  - If multiple minima were identified, use their average position



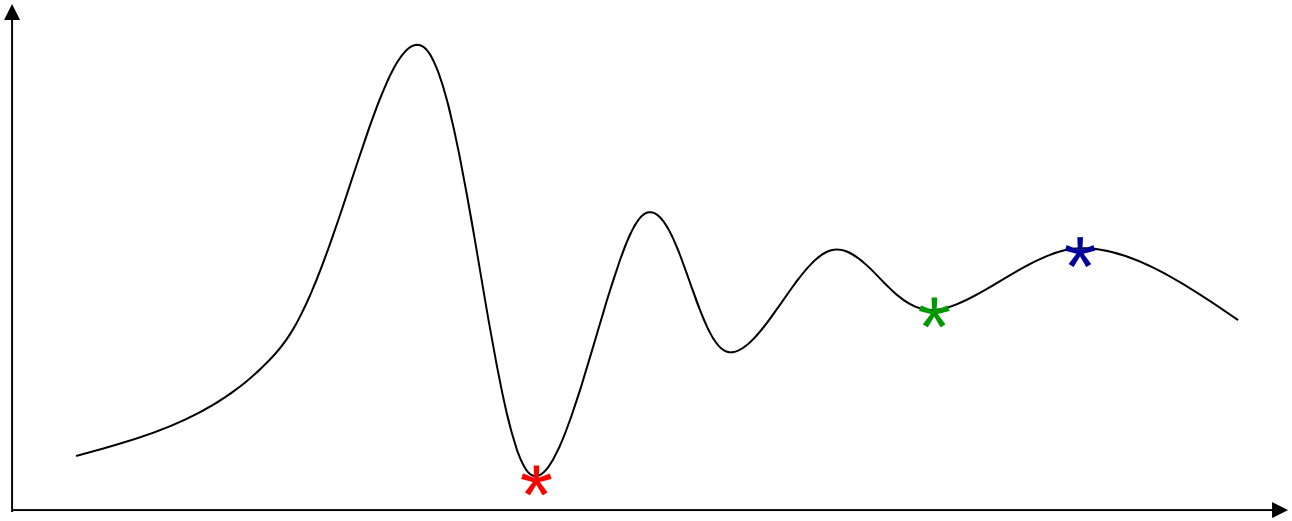
Gaussian Fitting Implementation Details (IV)

- Initialization:
  - Use detected local maxima to localize the search

- Strengths and weaknesses
  - Good (not necessarily optimal) solution guaranteed
  - Computationally expensive, slow

Gaussian Fitting Implementation Details (V)

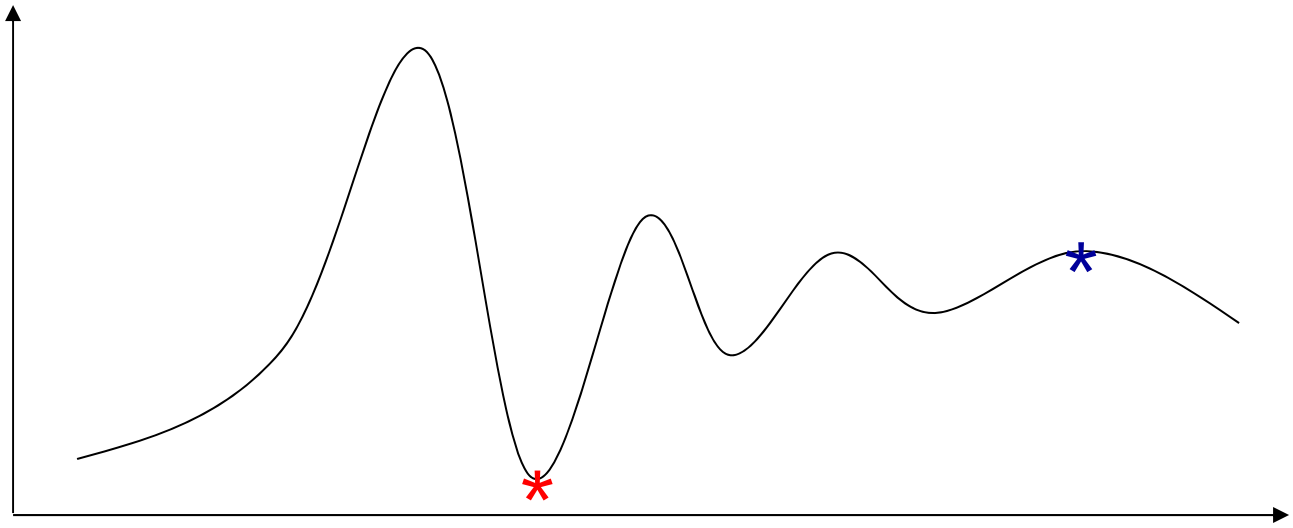
- Optimization strategy II: optimization search
  - Many optimization techniques can be applied



- Initialization:
  - Use detected local maxima
  - Use multiple randomly selected initiation points

Gaussian Fitting Implementation Details (VI)

- Strengths and weaknesses
  - Not limited by oversampling rate
  - May be trapped in local minimum



Sub-pixel Detection by Correlation

- Detection by maximization correlation

$$C_{x_0,y_0} = \sum_{i=1}^{M}\sum_{j=1}^{N} I_{i,j} K\left(i - x_0, j - y_0\right)$$

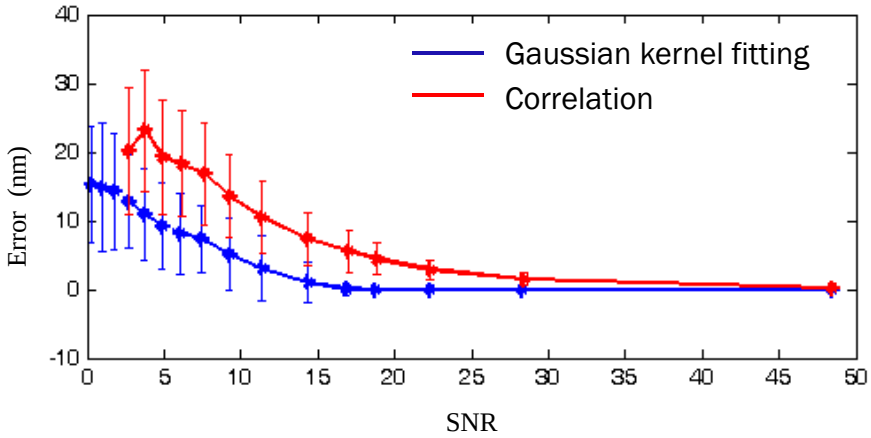
- Often the correlation function is normalized

$$C_{x_0,y_0} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I_{i,j} K\left(i - x_0, j - y_0\right)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} I_{i,j}^2} \cdot \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} K^2\left(i - x_0, j - y_0\right)}}$$

- Same strategy as in Gaussian fitting as the only difference is the cost function
  - Strategy I: exhaustive search
  - Strategy II: optimization search

Gaussian Fitting vs Correlation

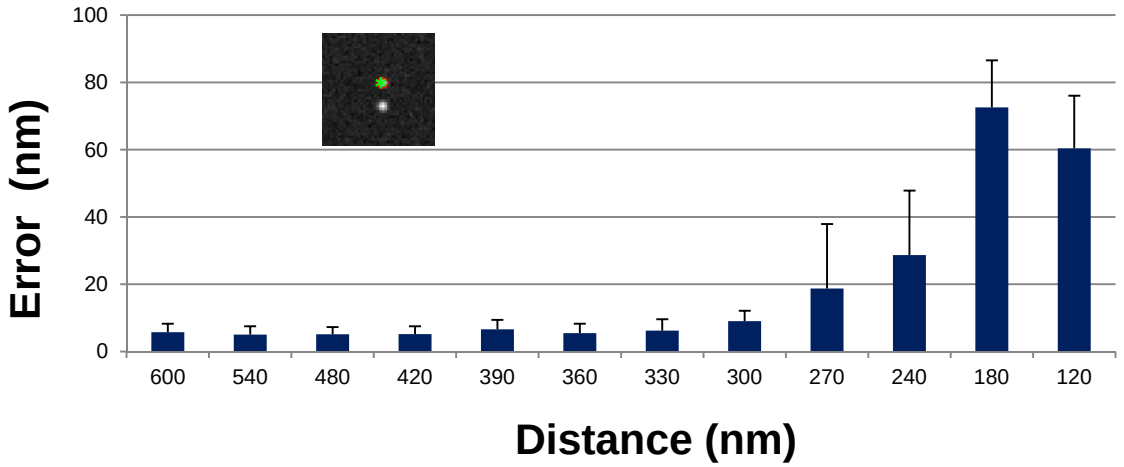
- For point features, Gaussian fitting is the best method overall.



- For larger non-diffraction limited features, correlation gives better resolution.

Limitations of Sub-Pixel Detection

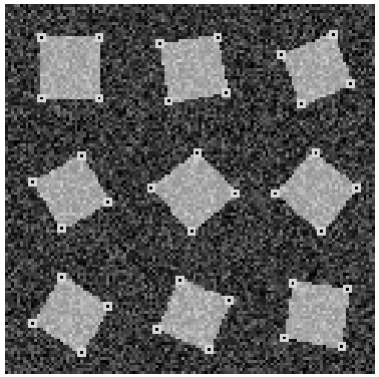
- When the distance between the two point features goes below the Rayleigh limit, they can no longer be resolved reliably unless under very high SNR.



- Review: subpixel resolution point feature detection
- Other point feature detection techniques
- Reproducible research in computational science
- Application I: molecule counting
- Application II: single molecule imaging
- Application III: protein colocalization analysis
- Application IV: super-resolution imaging

Other Point Feature Detection Techniques

- Corners rather than individual points are used more often in computer vision.



- A variety of techniques are also available from medical image analysis.

- J. Shi and C. Tomasi. "*Good Features to Track*." IEEE Conf. Computer Vision & Pattern Recognition, pp. 593-600, 1994.
- S. M. Smith and J. M. Brady. "*SUSAN - a new approach to low level image processing*." *International Journal of Computer Vision, vol. 23,* 45–78, 1997

- Review: subpixel resolution point feature detection
- Other point feature detection techniques
- Reproducible research in computational science
- Application I: molecule counting
- Application II: single molecule imaging
- Application III: protein colocalization analysis
- Application IV: super-resolution imaging

Open Source & Reproducible Research

An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.

—D. Donoho (http://www-stat.stanford.edu/~donoho/)

- Jon Claerbout is often credited as the first who proposed reproducible computational research.
- There are challenges. But these challenges can be overcome.
- Methods for public-funded biological studies should be open-source.

http://reproducibleresearch.net/index.php/Main_Page

http://sepwww.stanford.edu/data/media/public/sep/jon/

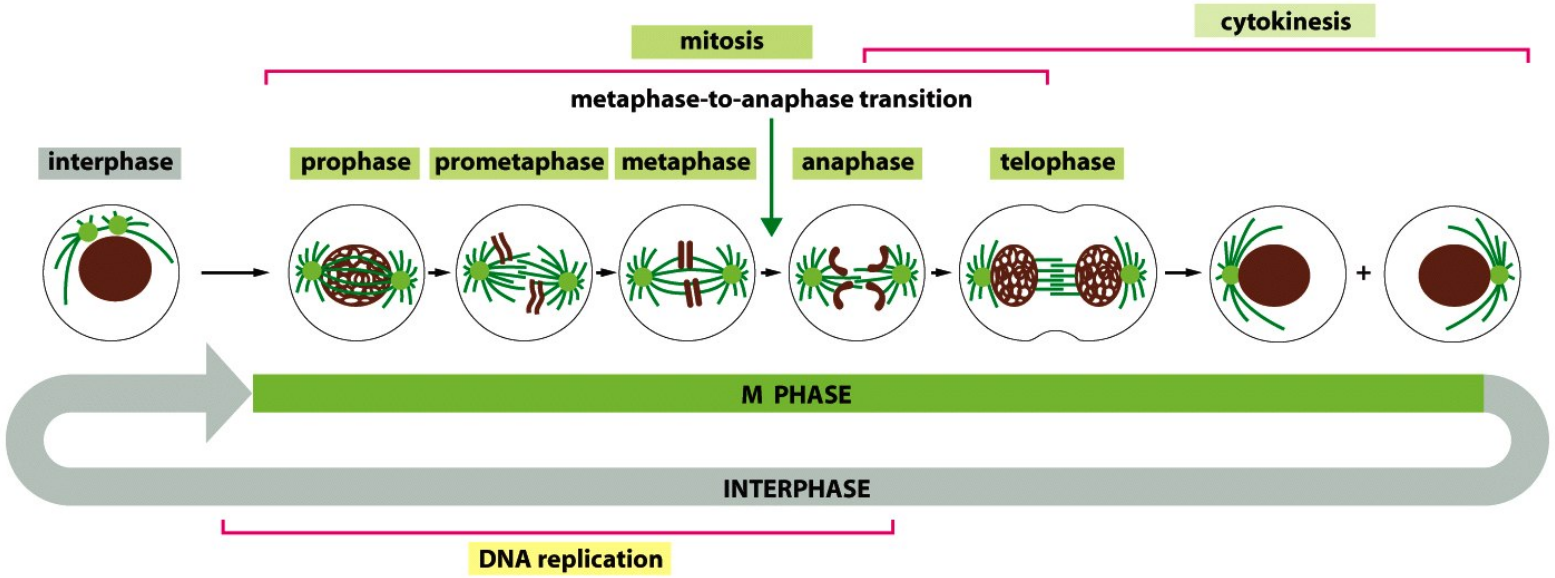
Open Source & Reproducible Research (II)

- Current literatures of image processing and computer vision often are formulated mathematically and do not provide source code.

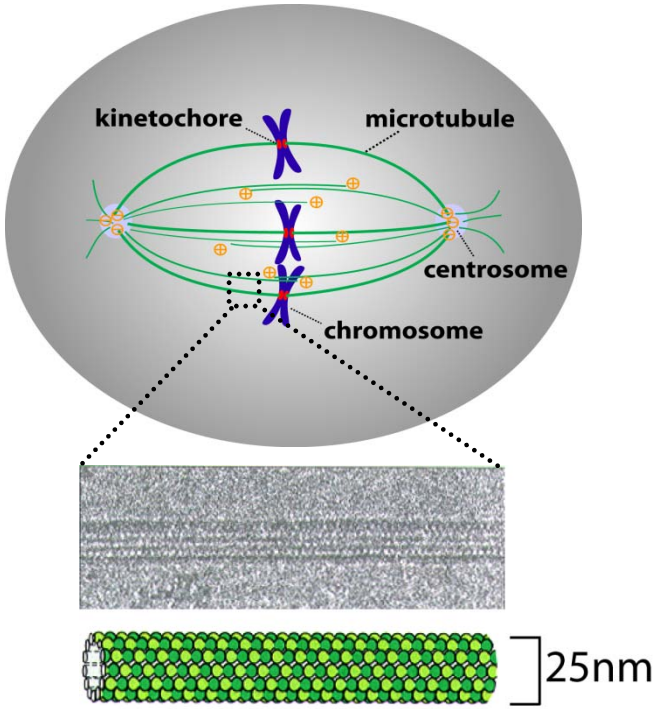
- Challenges
  - implementation (numerical issues)
  - parameter tuning
  - robustness a major performance issue

- Review: subpixel resolution point feature detection
- Other point feature detection techniques
- Reproducible research in computational science
- Application I: molecule counting
- Application II: single molecule imaging
- Application III: protein colocalization analysis
- Application IV: super-resolution imaging

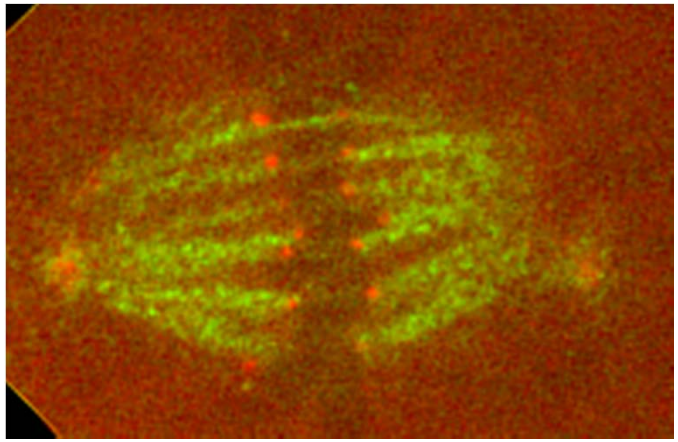
Overview of Cell Cycle



Dynamic Microtubules in the Mitotic Spindle

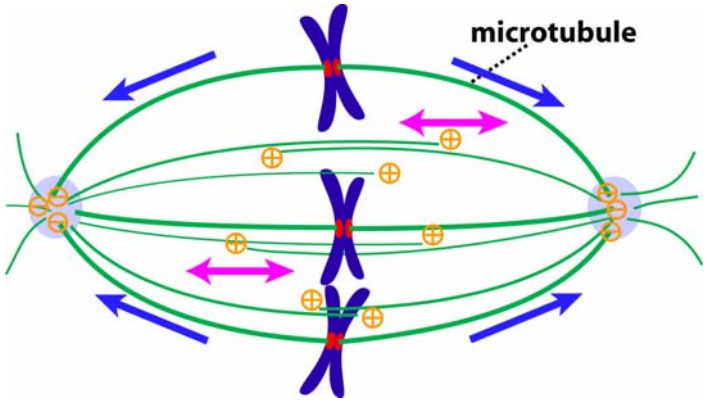


Alberts et al., MBoC5

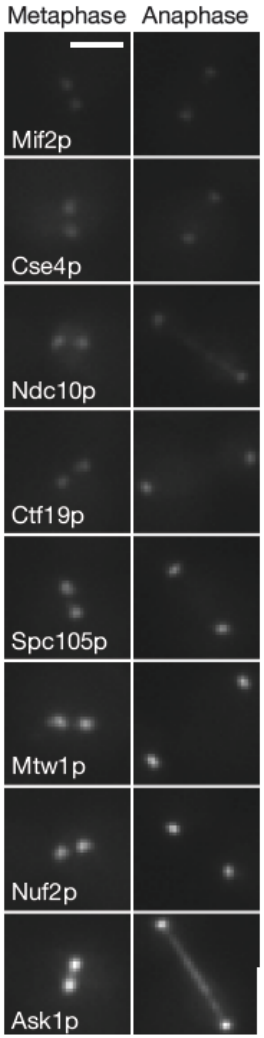


Green: microtubule
Red: kinetochore

5 μm



Molecule Counting by Ratio Imaging



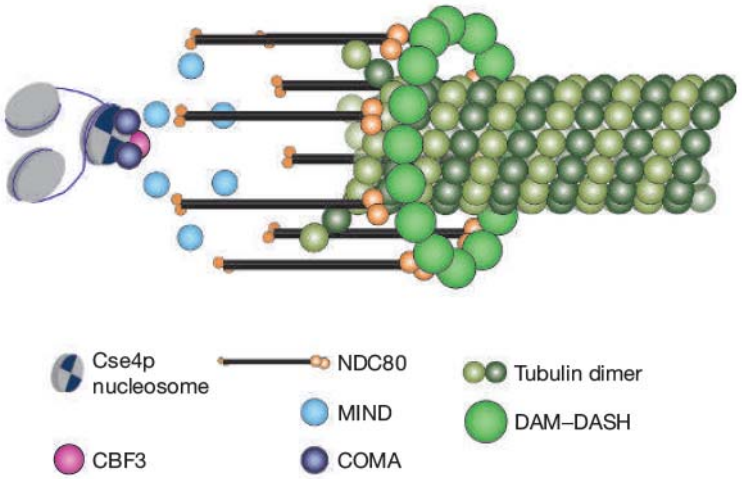


Table 1 Metaphase and anaphase–telophase ratio measurements

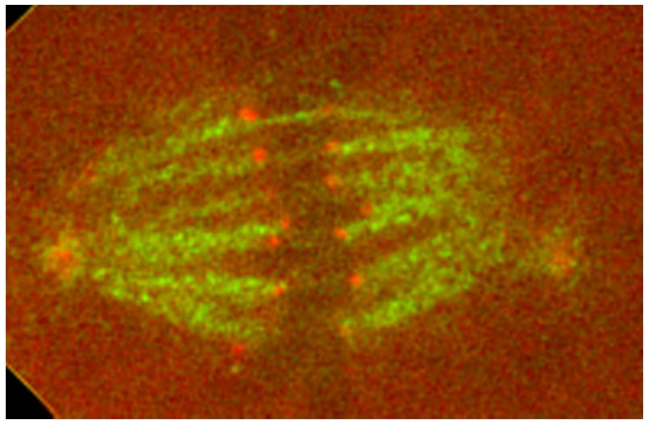
Complex	Protein	Vertebrate homologue	Metaphase ratio	Anaphase ratio	Metaphase number	Anaphase number
Nucleosome	Cse4p	hsCENP-A	1	1	2	2
CBF3	Ndc10p	–	1.9 ± 0.2	1.3 ± 0.01	4	2–3
CBF3	Cep3p	–	0.9 ± 0.2	0.6 ± 0.01	2	1–2
–	Mif2p	hsCENP-C	5.4 ± 0.4*	5.5 ± 0.10*	1–2	1–2
COMA	Ctf19p	hsCENP-F	3.4 ± 0.3*	3.4 ± 0.20*	3	2
–	Spc105	CeKNL-1	2.4 ± 0.01	2.4 ± 0.01	5	5
MIND	Mtw1p	hsMis12	3.3 ± 0.2	2.4 ± 0.10	6–7	4–5
NDC80	Nuf2p	hsNuf2	4.0 ± 0.2	3.6 ± 0.20	8	7
DAM–DASH	Ask1p	–	9.0 ± 1	5.3 ± 0.30	16–20	10–11
CTF3	Ctf3p	hsCENP-I	–	0.5 ± 0.01	–	1
CHL4–IML3	Chl4p	–	–	0.26 ± 0.01	–	<1
NKP1–NKP2	Nkp2p	–	–	6.1 ± 0.05*	–	1

The ratios shown are the average ratios obtained from three experiments with at least 20 measurements for metaphase cells, and at least two experiments with up to 80 measurements for late anaphase–telophase cells. The coefficient of variation (s.d. / mean) was better than 0.26 in all the measurements with the exception of Cep3p, for which the coefficient is 0.5. The asterisks indicate that the reported ratio is (Nuf2p – signal) : (protein signal). Mif2p–GFP, Ctf19p–GFP and Nkp2–GFP measurements were carried out with Nuf2p–GFP as the reference signal (see Methods).

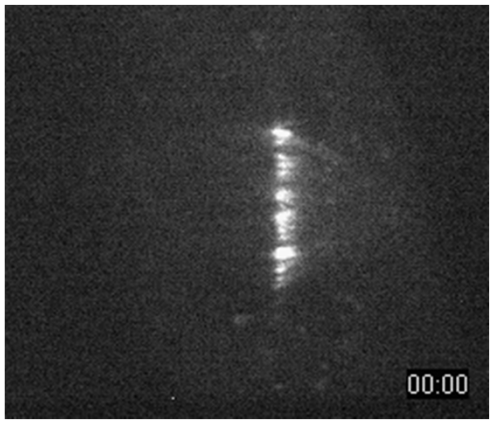
A. P. Joglekar et al, Molecular architecture of a kinetochore-microtubule attachment site, *Nat. Cell Biol.*, 8:581-585, 2006.

- Review: subpixel resolution point feature detection
- Other point feature detection techniques
- Reproducible research in computational science
- Application I: molecule counting
- Application II: single molecule imaging
- Application III: protein colocalization analysis
- Application IV: super-resolution imaging

Confirmation of Poleward Flow of Spindle Microtubules

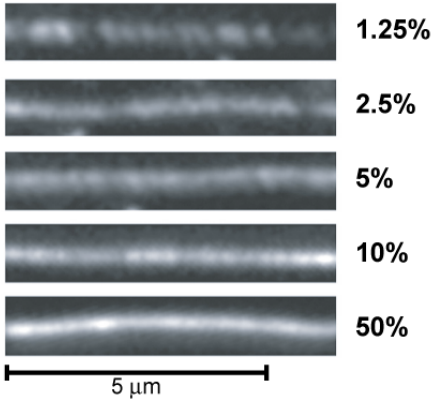


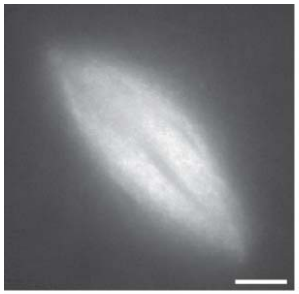
5 μm

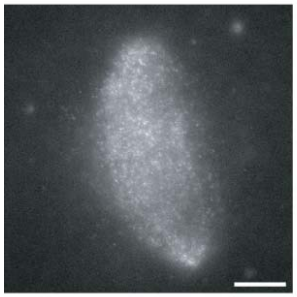


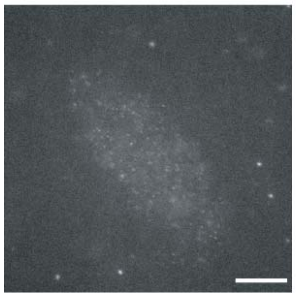
Cameron et al, *JCB*, 173:173-179,2006

Fluorescent Speckle Microscopy (FSM)





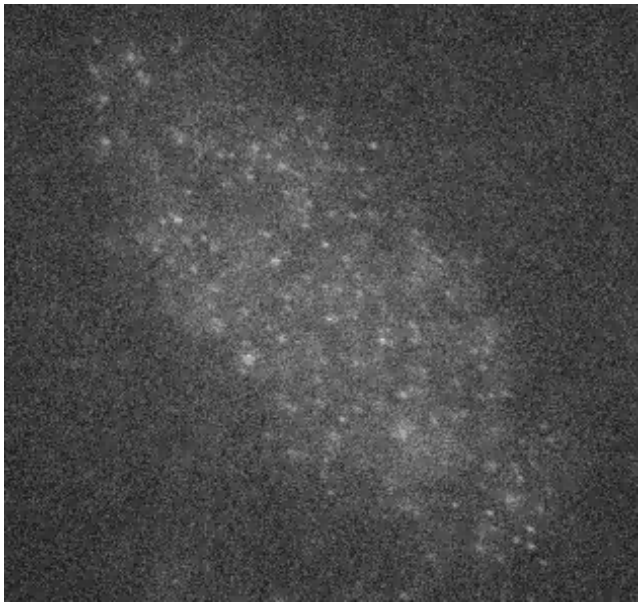




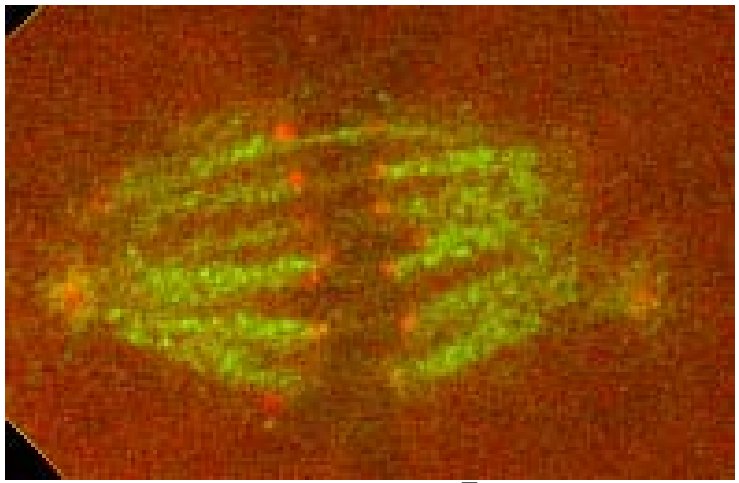
1.32x10^{-4}

1.32x10^{-5}

1.32x10^{-6}

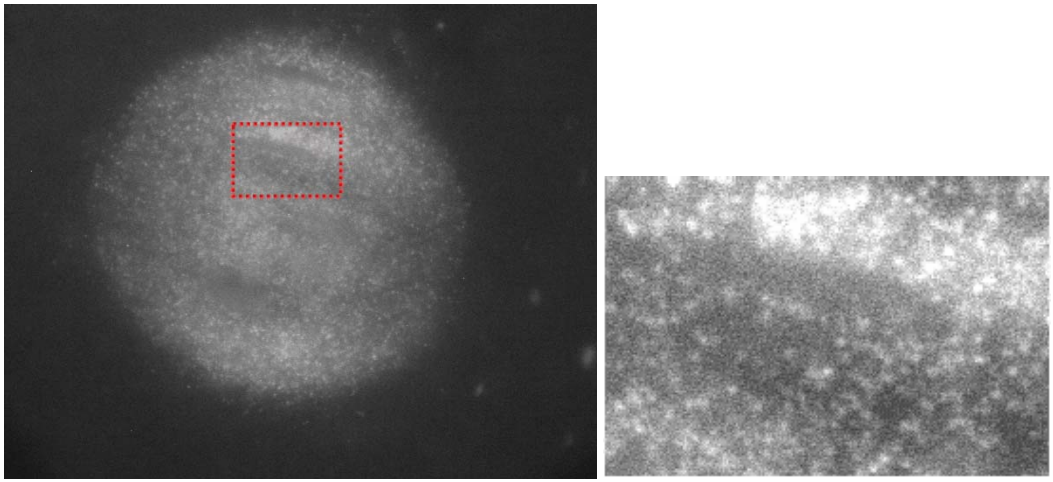


FSM of Dynamic Spindle Architecture



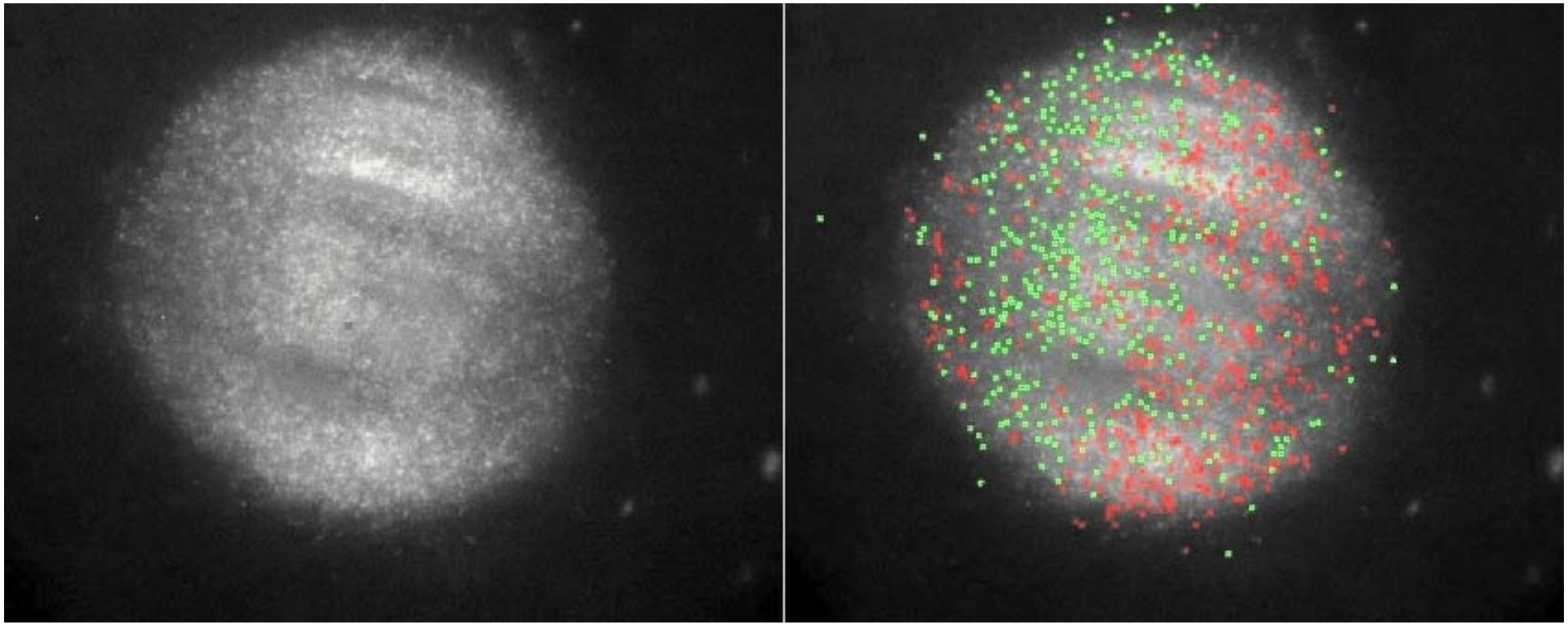
5 μm

Fluorescent speckle microscopy



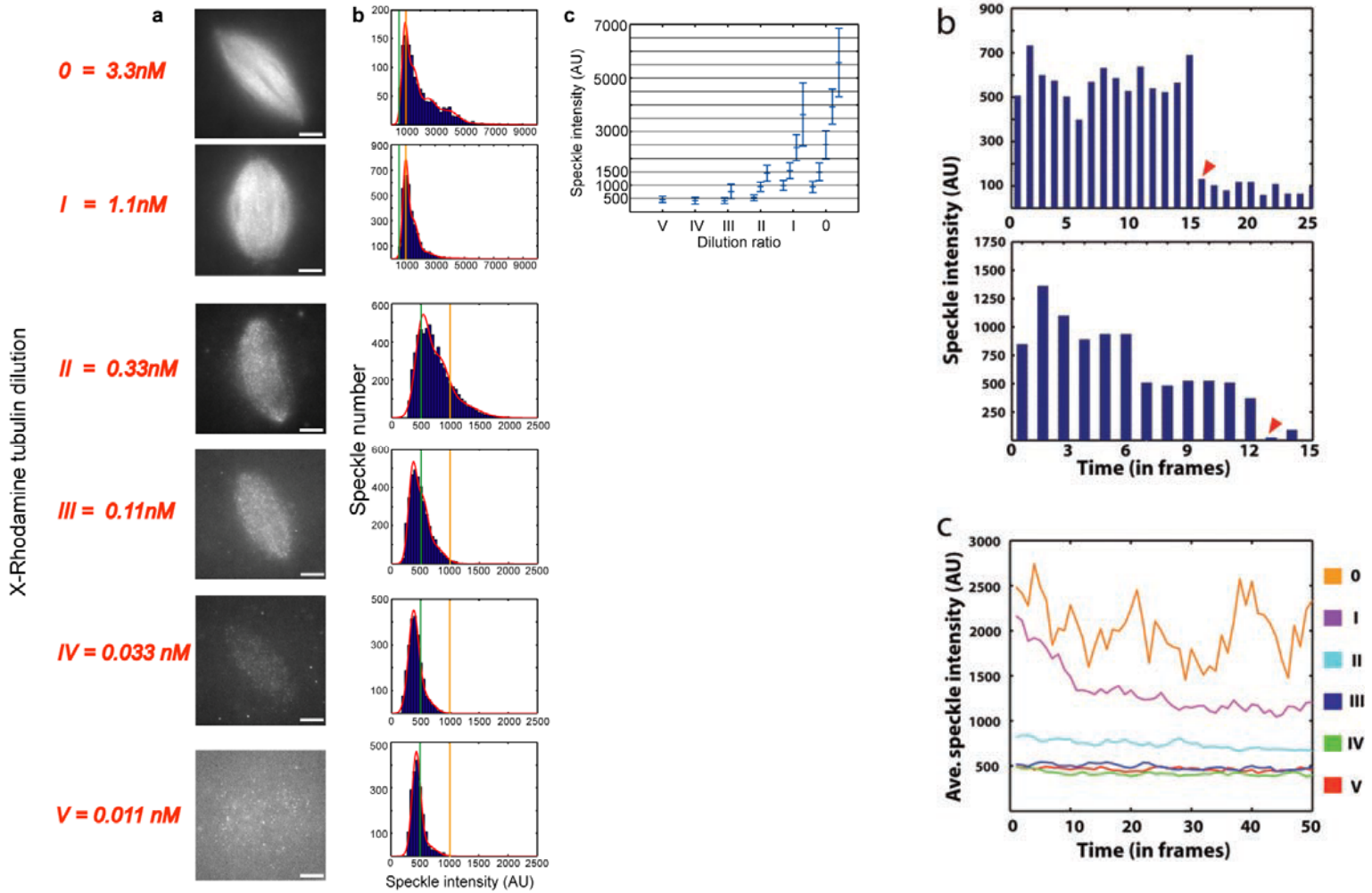
10 μm

Quantitative Mapping of Spatial-Temporal Spindle Dynamics



Yang et al., *J. Cell Biol.,* 182:631-639, 2008

Analyzing Dynamics of Individual Microtubules by Single Fluorophore FSM



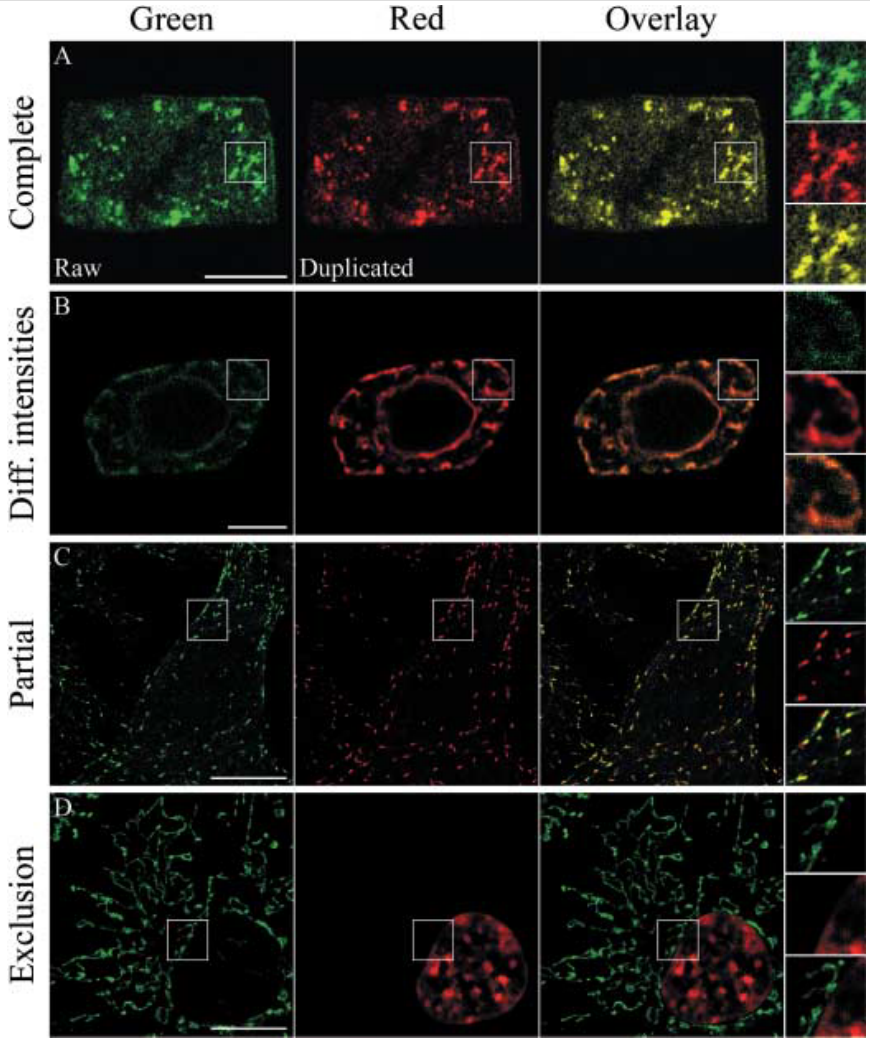
Yang et al., Architectural dynamics of the meiotic spindle revealed by single-fluorophore imaging, *Nat. Cell Biol., 9:1233-1242,* 2007

- Review: subpixel resolution point feature detection
- Other point feature detection techniques
- Reproducible research in computational science
- Application I: molecule counting
- Application II: single molecule imaging
- Application III: protein colocalization analysis
- Application IV: super-resolution imaging

Protein Colocalization Analysis

- To determine whether protein molecules colocalize is essential to many biological studies.

- Protein colocalization analysis can be performed in both fixed and live cells.



Bolte & Cordelieres, A guided tour into subcellular colocalization analysis in light microscopy, J. Microscopy, 224:213-232, 2006.

Protein Colocalization Analysis (II)

- A commonly used analysis protocol
  - Step 1: Detect point features in a selected channel.
  - Step 2: Identify the intensities at another channel at detected point feature locations.
  - Step 3: Calculate correlation coefficient

- Commonly used definitions of colocalization coefficients
  - Overlap coefficient

$$r_o = \frac{\sum_i A_i \times B_i}{\sqrt{\left[\sum_i A_i^2 \times \sum_i B_i^2\right]}}$$

  - Pearson's coefficient

$$r_p = \frac{\sum_i (A_i - a) \times (B_i - b)}{\sqrt{\left[\sum_i (A_i - a)^2 \times \sum_i (B_i - b)^2\right]}}$$

Some General Comments

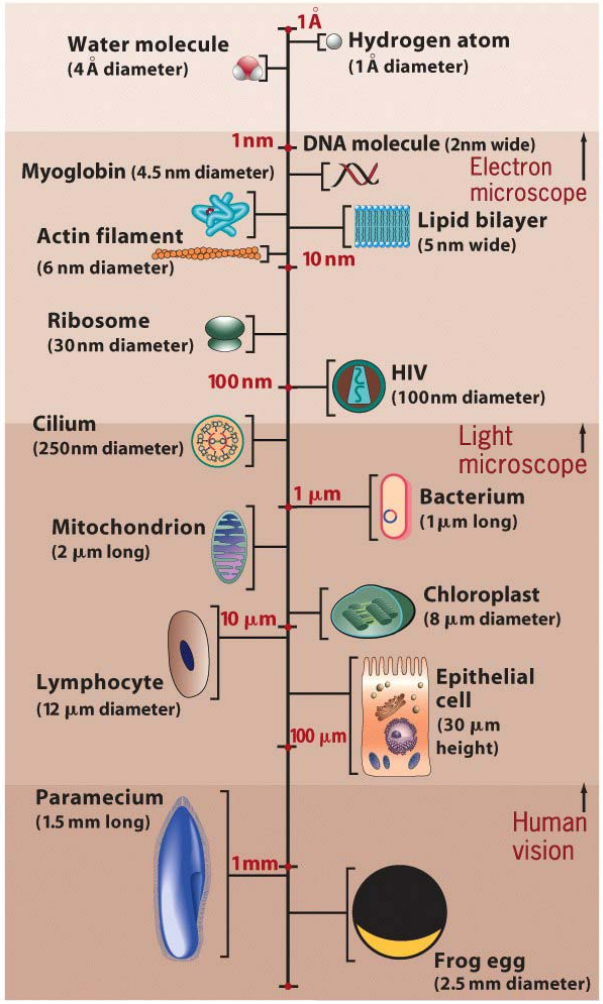
- It is possible but limiting to consider bioimage analysis as just another application.

- Excellent research opportunities in bioimage informatics.

- Challenges
    - Solid training in image processing and computer vision
    - Interdisciplinary background and thinking
        - For identifying and solving problems
        - For collaboration

- Review: subpixel resolution point feature detection
- Other point feature detection techniques
- Reproducible research in computational science
- Application I: molecule counting
- Application II: single molecule imaging
- Application III: protein colocalization analysis
- Application IV: super-resolution imaging

Imaging as a Tool for Studying Cellular Structure



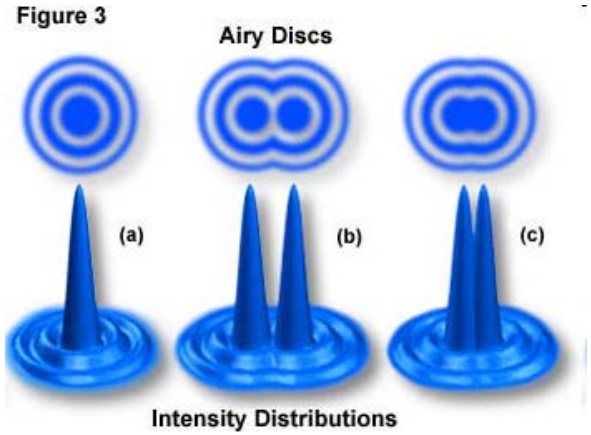
- Crystallography, NMR, spectroscopy
  - Resolution: ≤1 nm
  - Live/physiological condition: **NO**
    → Samples must be specially prepared
- Electron microscopy
  - Resolution: between 1nm &100nm
  - Live/physiological condition: **NO**
    → Samples must be fixed
- Light microscopy
  - Resolution: ≥ 100nm
  - Live/physiological condition: **YES**

Performance Metrics of a Fluorescence Microscope

- Resolution:

  - Rayleigh limit: $D = \frac{0.61\lambda}{NA}$

  - Sparrow limit: $D = \frac{0.47\lambda}{NA}$

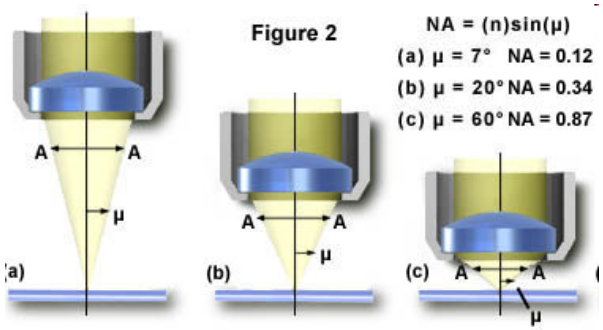


- Numerical aperture (NA)

$$NA = n \cdot \sin \mu$$

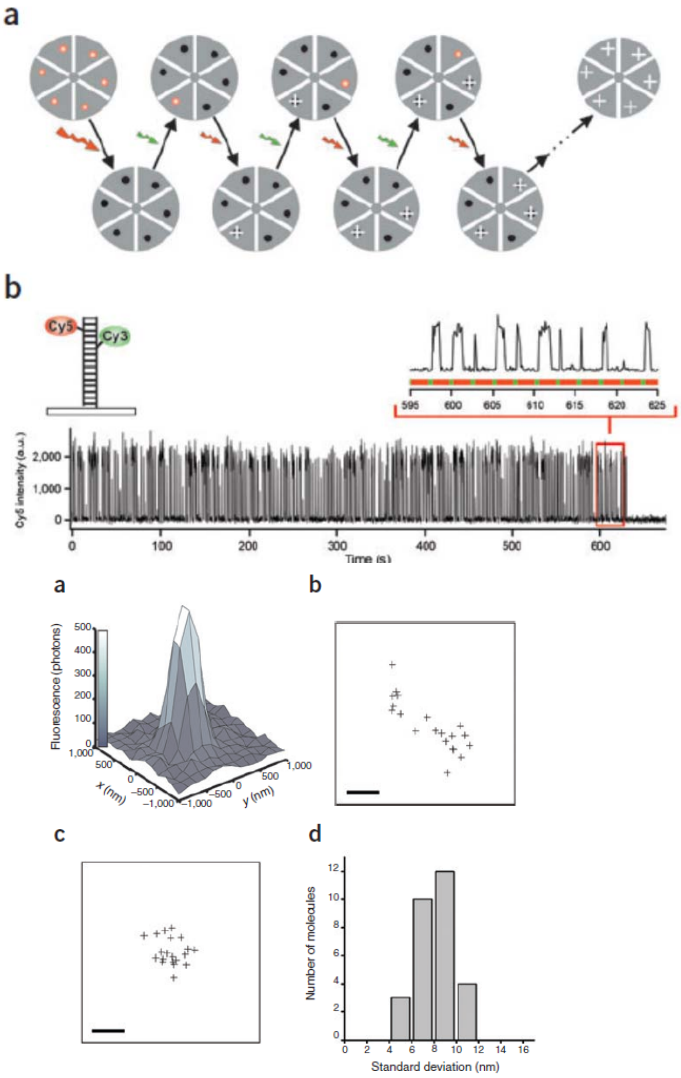
n: refractive index of the medium between the lens and the specimen

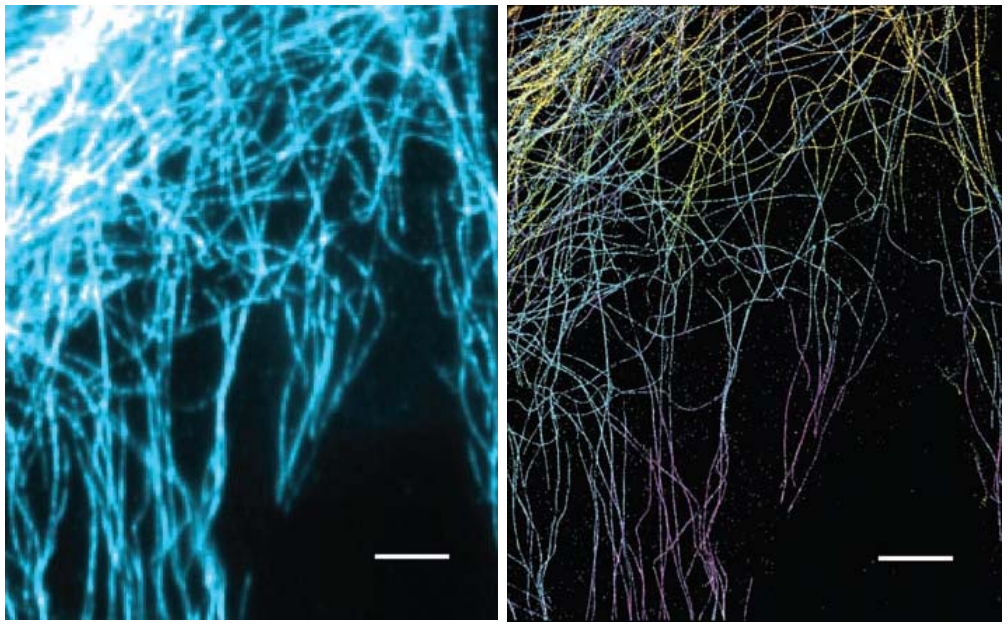
μ: half of the angular aperture



- Water n=1.33
- Immersion oil n=1.51

Stochastic Optical Reconstruction Microscopy (STORM)





Huang et al, Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy, Science, 319:810-813, 2008

M. J. Rust et al, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nature Methods, 10:793-795, 2006.

Photoactivation Localization Microscopy (PALM)

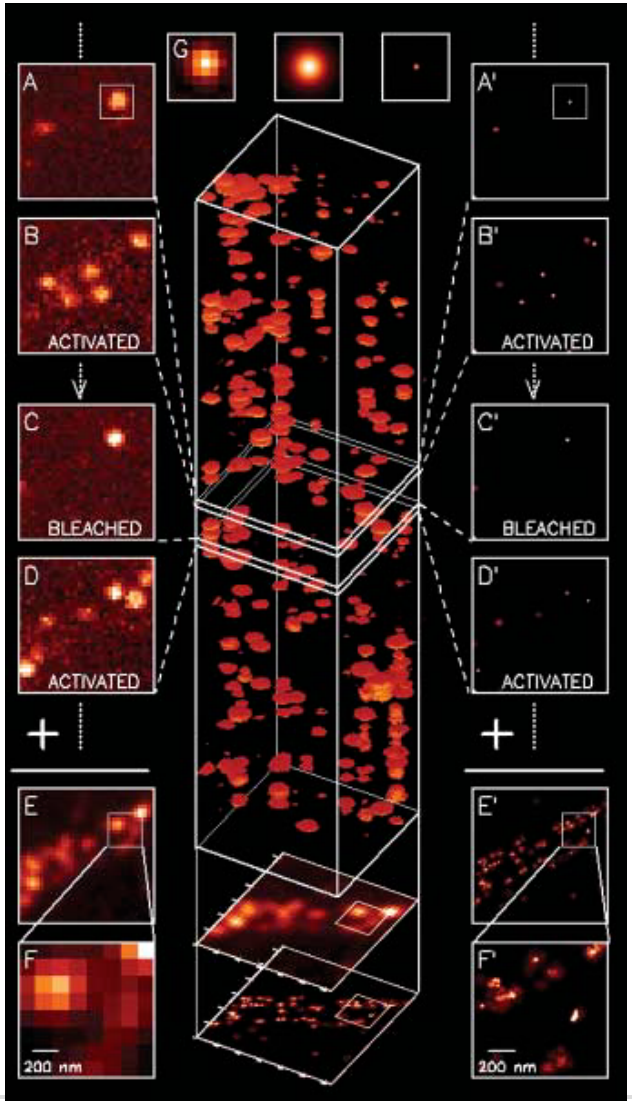
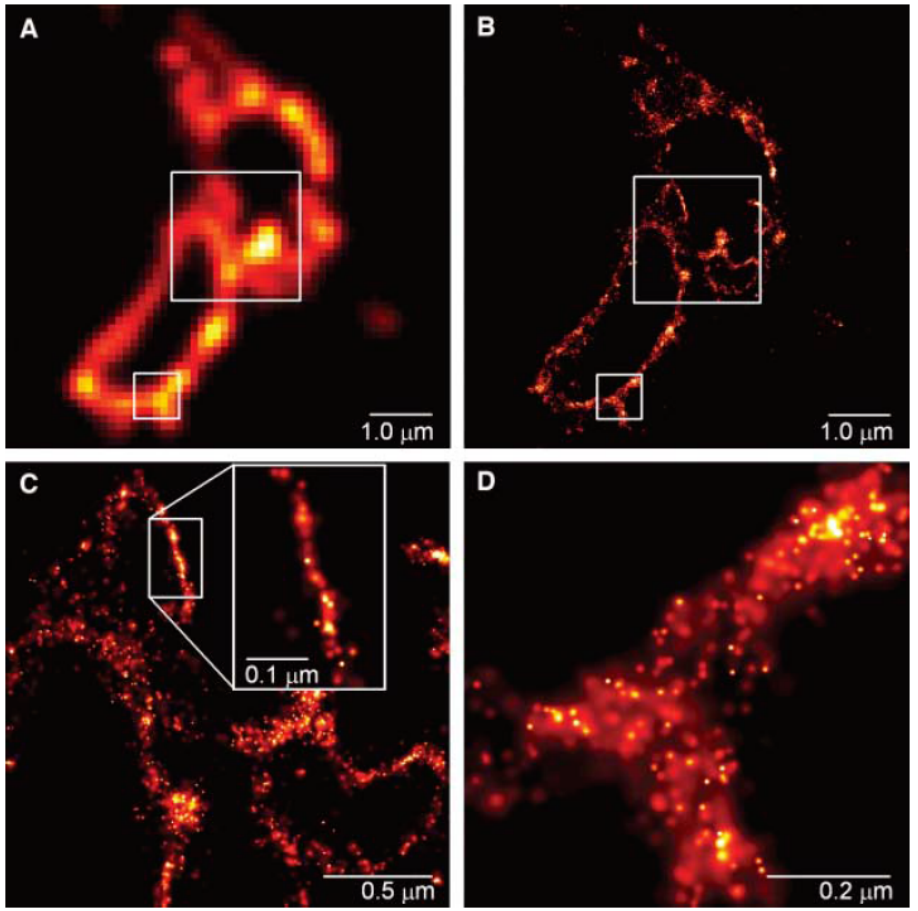


Fig. 1. The principle behind PALM. A sparse subset of PA-FP molecules that are attached to proteins of interest and then fixed within a cell are activated (**A** and **B**) with a brief laser pulse at λ_{act} = 405 mm and then imaged at λ_{exc} = 561 mm until most are bleached (**C**). This process is repeated many times (**C** and **D**) until the population of inactivated, unbleached molecules is depleted. Summing the molecular images across all frames results in a diffraction-limited image (**E** and **F**). However, if the location of each molecule is first determined by fitting the expected molecular image given by the PSF of the microscope [(**G**), center] to the actual molecular image [(G), left], the molecule can be plotted [(G), right] as a Gaussian that has a standard deviation equal to the uncertainty $\sigma_{x,y}$ in the fitted position. Repeating with all molecules across all frames (**A'** through **D'**) and summing the results yields a superresolution image (**E'** and **F'**) in which resolution is dictated by the uncertainties $\sigma_{x,y}$ as well as by the density of localized molecules. Scale: 1 × 1 μm in (F) and (F'), 4 × 4 μm elsewhere.



Questions?