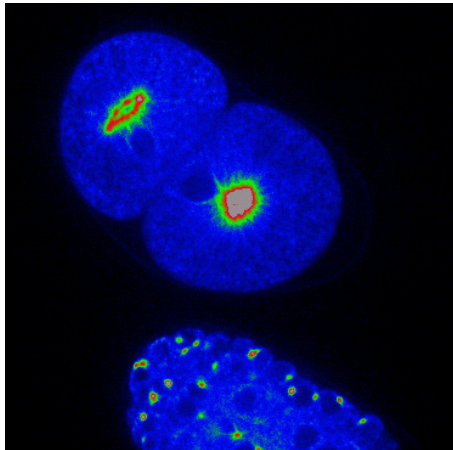


TIME-RESOLVED QUANTIFICATION OF CENTROSOMES
BY AUTOMATED IMAGE ANALYSIS
SUGGESTS LIMITING COMPONENT
TO SET CENTROSOME SIZE IN *C. ELEGANS* EMBRYOS

DISSERTATION



STEFFEN JAENSCH

TIME-RESOLVED QUANTIFICATION OF CENTROSOMES
BY AUTOMATED IMAGE ANALYSIS
SUGGESTS LIMITING COMPONENT
TO SET CENTROSOME SIZE IN *C. ELEGANS* EMBRYOS

DISSERTATION

zur Erlangung des akademischen Grades Doktor rerum naturalium (Dr. rer. nat.)

vorgelegt an der
Technischen Universität Dresden
Fakultät Informatik

eingereicht von
Diplom-Informatiker Steffen Jaensch
geboren am 21. Juni 1980 in Rüdersdorf

Gutachter: Prof. Dr. Michael Schroeder (TU Dresden, Fakultät Informatik)
Dr. Eugene W. Myers (Janelia Farm Research Campus, USA)
Tag der Verteidigung: 02. Dezember 2010

Dresden im September 2010

Steffen Jaensch: *Time-Resolved Quantification of Centrosomes by Automated Image Analysis Suggests Limiting Component to Set Centrosome Size in C. Elegans Embryos*, Dissertation, September 2010

To my family

ABSTRACT

The centrosome is a dynamic organelle found in all animal cells that serves as a microtubule organizing center during cell division. Most of the centrosome components have been identified by genetic screens over the last decade, but little is known about how these components interact with each other to form a functional centrosome. Towards a better understanding of the molecular organization of the centrosome, we investigated the mechanism that regulates the size of the centrosome in the early *C. elegans* embryo.

For this, we monitored fluorescently labeled centrosomes in living embryos and developed a suite of image analysis algorithms to quantify the centrosomes in the resulting 3D time-lapse images. In particular, we developed a novel algorithm involving a two-stage linking process for tracking centrosomes, which is a multi-object tracking task. This fully automated analysis pipeline enabled us to acquire time-resolved data of centrosome growth in a large number of embryos and could detect subtle phenotypes that were missed by previous assays based on manual image analysis.

In a first set of experiments, we quantified centrosome size over development in wild-type embryos and made three essential observations. First, centrosome volume scales proportionately with cell volume. Second, beginning at the 4-cell stage, when cells are small, centrosome size plateaus during the cell cycle. Third, the total centrosome volume the embryo gives rise to in any one cell stage is approximately constant. Based on our observations, we propose a ‘limiting component’ model in which centrosome size is limited by the amounts of maternally derived centrosome components.

In a second set of experiments, we tested our hypothesis by varying cell size, centrosome number and microtubule-mediated pulling forces. We then manipulated the amounts of several centrosomal proteins and found that the conserved centriolar and pericentriolar material protein SPD-2 is one such component that determines centrosome size.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publication:

S. Jaensch, M. Decker, A. A. Hyman, and E. W. Myers (2010) *Automated tracking and analysis of centrosomes in early Caenorhabditis elegans embryos*. *Bioinformatics* 26, i13-i20.

ACKNOWLEDGMENTS

This thesis describes the research I carried out as a PhD student at the Max-Planck Institute of Molecular Cell Biology and Genetics in Dresden, Germany and Janelia Farm Research Campus in Ashburn, Virginia, USA from September 2007 to September 2010. Many people have contributed in very different ways to make this period of my life a very exciting, interesting, challenging and pleasant one. I would like to thank...

Professor Dr. Anthony Hyman for giving me the opportunity to do my PhD in his lab and for again and again emphasizing to think in science in terms of specific questions. With only a single biology class (on neuroscience) during my undergrad, the scientific world of biology was completely new for me; working in your lab gave me great insights and many, many possibilities to explore and learn about this exciting field of research.

Dr. Eugene Myers for hosting me as a visiting student researcher at Janelia Farm and for putting the most abstract ideas into illustrative yet precise words. Gene, thanks for your personal advise and the very active group meeting discussions, which both were truly helpful to me. Working and living at the farm was a very special, certainly enjoyable experience, scientifically and personally.

Markus Decker for the close collaboration throughout my entire PhD time. Double the fun, half the frustration and always something to give each other shit about. Markus, I think we made a pretty good team. Thank you for the great time!

My thesis advisory committee members Dr. Pavel Tomancak, Professor Dr. Michael Schroeder and Professor Dr. Frank Juelicher for their constructive comments and guidance through the PhD program.

All members of the Hyman lab. It's great to work in such an active and social lab. Of course, I will never forget our lab retreat to Bad Muskau. Who wants a sunny relaxing day on the river and a deluxe five-course dinner if you instead get water fights and a knight's banquet ;)

All members of Gene Myers' group, especially Fuhui Long, Nathan Clark and Parvez Ahammad, as well as Hanchuan Peng and his group members for helpful discussions.

Horatiu Fantana for bead injections and Stephan Preibisch for collaborating on SPIM imaging of PH-domain::GFP embryos.

All members of the IT department, especially Matt Boes and Edmund Malik, for the excellent IT infrastructure at the MPI. Everything simply worked or you made it work in no time!

All members of the light microscopy facility for assistance with all sorts of technical imaging challenges.

Eric Hoopfer and Martin Oswald for proof reading.

The Alexander-von-Humboldt Foundation and the Max-Planck Society for financial support.

The city of Dresden for the world's greatest commute to work. There is no better way to start your day than with a quick bike ride along the Elbe river.

My friends for all the fun, love and support. You make me happy! Thanks for a special breakfast at the Skybar. Thanks for great kayaking, surfing and camping trips, cooking sessions, bike rides, volleyball matches, partying or just hanging out. The Dresden crowd I would like to especially thank for our eight year's tradition of Sunday card playing nights (Doppelkopf) that helped me keep my sanity even in most turbulent times.

And finally, with all my heart, I would like to thank my parents, grandparents and brothers for being the most wonderful family I could wish to have.

CONTENTS

1	INTRODUCTION	1
1.1	Automated Image Analysis in Biology	1
1.2	The Centrosome in <i>C. Elegans</i>	4
1.2.1	Functions and Roles	4
1.2.2	Structure	5
1.2.3	The Centrosome Cycle	8
1.3	Biological Question: What Sets the Size of a Centrosome?	9
1.4	Aims of this Thesis	11
2	AUTOMATED IMAGE ANALYSIS FOR AN ASSAY OF CENTROSOME SIZE	13
2.1	Imaging Centrosomes in Living Embryos	13
2.2	Centrosome Tracking	16
2.2.1	Tracking Challenges and Simplifying Properties	16
2.2.2	Related Work	18
2.2.3	Overview of the Tracking Algorithm	23
2.2.4	Object (Over-)Detection	25
2.2.5	The Bipartite Graph Matching Paradigm	28
2.2.6	Finding Reliable Core Trajectories	31
2.2.7	Extending Core Trajectories to Full Trajectories	32
2.2.8	Trajectory Stitching	34
2.2.9	Tracking Performance	35
2.3	Measuring Centrosome Size	37
2.3.1	Definition of Centrosome Size and Measuring Algorithm	37
2.3.2	Validation of the Measuring Accuracy	39
2.4	Biological Identities of the Tracked Centrosomes	42
2.5	Detection of the Nuclear Envelope Breakdown	43
2.5.1	Accuracy of NEBD detection	45
3	3D-SEGMENTATION OF PH-DOMAIN::GFP IMAGES	47

3.1	Spinning Disk Confocal Microscopy Images	47
3.1.1	Automatic initialization	51
3.2	Single Plane Illumination Microscopy Images	51
3.3	Cell Volumes in <i>C. Elegans</i> Embryos	53
4	WHAT SETS THE SIZE OF A CENTROSOME?	57
4.1	Centrosome Disintegration Begins at the Metaphase-Anaphase Transition	57
4.2	The Growth Kinetics of Centrosomes in Wild-Type Embryos Change through Development	59
4.3	Centrosomes become smaller in Each Round of Cell Division	60
4.4	Centrosome Volume Scales Proportionately with Cell Volume	62
4.5	Centrosome Size is Independent of Lineage	63
4.6	Total Centrosome Volume is Constant Across Cell Stages	65
4.7	Total Centrosome Volume is Independent of Centrosome Number	66
4.8	Several Components are Required for a Full Size Centrosome	68
4.9	Over-Expression of SPD-2 Leads to Bigger Centrosomes	70
4.10	Cytoplasmic Intensity Analysis	73
4.11	Fluorescence Recovery After Photo-Bleaching Experiments	75
5	SUMMARY AND DISCUSSION	79
5.1	Tracking Algorithm	80
5.2	Detection of Subtle Phenotypes Invalidates Hypotheses in Previous Studies	82
5.3	Limiting Component Hypothesis for Centrosome Size	83
6	DIRECTIONS FOR FUTURE WORK	87
6.1	Tracking Algorithm	87
6.2	Automated Image Acquisition and Genome-Wide Screen	87
6.3	Imaging with Single Plane Illumination Microscopy	88
6.4	Experiments in the Context of the Limiting Component Hypothesis	88
7	MATERIALS AND METHODS	91
	BIBLIOGRAPHY	95

LIST OF FIGURES

Figure 1.1	A typical example of a wild-type <i>C. elegans</i> embryo.	4
Figure 1.2	Schematic drawing of a centrosome.	6
Figure 1.3	Assembly pathway of the centrosome.	7
Figure 1.4	Life cycle of a centrosome.	8
Figure 2.1	Examples of <i>C. elegans</i> embryos with different centrosome markers.	14
Figure 2.2	A typical centrosome trajectory	23
Figure 2.3	Workflow of the centrosome tracking algorithm.	24
Figure 2.4	Response of Mexican hat filters of various sizes to centrosome and background image patches.	26
Figure 2.5	Measuring radius, total intensity and background level by Gaussian fitting.	26
Figure 2.6	Illustration of a radius hysteresis situation.	28
Figure 2.7	Illustration of an exchange during the backward extension phase.	34
Figure 2.8	The intensity of the centrosomes strongly depends on the distance to the embryo surface.	38
Figure 2.9	Refining centrosome radius.	39
Figure 2.10	Imaging beads inside <i>C. elegans</i> embryos.	40
Figure 2.11	The measured radius of injected beads increases slightly with increasing z-position.	41
Figure 2.12	<i>C. elegans'</i> invariant cell lineage.	43
Figure 2.13	Typical time vs. nuclear intensity curve computed by the nuclear envelope breakdown detection algorithm.	44
Figure 3.1	Spinning disk confocal stack of an embryo with PH-domain::GFP labeled cell membranes.	49

Figure 3.2	Workflow of the segmentation algorithm for PH-domain::GFP images acquired with spinning disk confocal microscope.	50
Figure 3.3	Computation of an initial watershed marker for each cell in the center plane.	52
Figure 3.4	Single plane illumination stack of an embryo with PH-domain::GFP labeled cell membranes.	53
Figure 3.5	Workflow of the segmentation algorithm for PH-domain::GFP images acquired with single plane illumination microscope.	54
Figure 3.6	3D view of the segmentation results for SPIM images.	54
Figure 3.7	Cell volumes in the 1- to 4-cell stage.	55
Figure 4.1	Centrosome growth in wild type and in the absence of cortical force generation.	58
Figure 4.2	Centrosome growth through development.	60
Figure 4.3	Centrosomes become smaller throughout embryonic development.	61
Figure 4.4	Centrosome volume versus respective cell volume for three different centrosome markers.	62
Figure 4.5	Centrosome size plateaus in AB cells of <i>ani-2</i> (RNAi) embryos.	63
Figure 4.6	Centrosomes in AB and P ₁ are approximately equal in size when P ₀ divides symmetrically.	64
Figure 4.7	Total centrosome volume is conserved through development.	65
Figure 4.8	Total centrosome volume is independent of the number of centrosomes per cell (<i>zyg1(b1)</i> and <i>zyg-1(it29)</i>).	67
Figure 4.9	Total centrosome volume is independent of centrosome number per cell (<i>sas-4</i> (RNAi)).	68
Figure 4.10	Three potential candidates may act limiting on centrosome size.	69
Figure 4.11	Centrosome growth in P ₀ in the absence of long microtubules is identical to wild-type.	69

Figure 4.12	Inhibiting transcription does not affect centrosome growth in the one- to four-cell stage.	70
Figure 4.13	Centrosome size is sensitive to SPD-2::GFP.	71
Figure 4.14	SPD-2 amounts are tightly regulated in the embryo.	73
Figure 4.15	The size of centrosomes is sensitive to total SPD-2 amounts.	74
Figure 4.16	Cytoplasmic intensity in N2 and SPD-2::GFP one-cell embryos.	76
Figure 4.17	Fluorescence recovers evenly over the centrosome after photo-bleaching.	78

LIST OF TABLES

Table 1	Minor and major tracking error rates.	36
---------	---------------------------------------	----

ABBREVIATIONS

CAI	Codon Adaptation Index
CCD	Charge-Coupled Device
EMCCD	Electron-Multiplying Charge-Coupled Device
FRAP	Fluorescence Recovery After Photo-Bleaching
γ TUB	γ -Tubulin
GFP	Green Fluorescent Protein
mRNA	Messenger RNA
NEBD	Nuclear Envelope Breakdown
MTOC	Microtubule Organizing Center
PCM	Pericentriolar Material
PH (Domain)	Pleckstrin Homology (Domain)
PSF	Point Spread Function
RNAi	RNA interference
SD	Spinning Disk (Confocal Microscopy)
s.d.	standard deviation
SPIM	Single Plane Illumination Microscopy (also: Selective Plane Illumination Microscopy)
YFP	Yellow Fluorescent Protein

INTRODUCTION

1.1 AUTOMATED IMAGE ANALYSIS IN BIOLOGY

Recent advances in imaging and fluorescent labeling techniques have enabled the visualization of structures and processes on all biological scales ranging from the whole organism level to the intra-cellular level. Often combined with systematic perturbations, for example by decreasing or increasing gene activity, large amounts of image data containing rich biological information are being produced in various research projects. The common goal in all of them is to convert the raw images into biological knowledge. To do so automated computational methods are needed because manual image analysis introduces subjectivity, often misses subtle phenotypes and/or the sheer amount of image data renders manual analysis impractical. Research and development of these methods are driven by the specific applications [1] which may belong to one of the following categories.

(1) In studies of animal behavior the goal is to understand the genetic and neuronal circuit bases that generate these behaviors. For example, Dankert et al. [2] have developed an automated assay that recognizes aggression and courtship behavior such as lunging and wing extension in pairs of fruit flies (*Drosophila melanogaster*) filmed from above at 30 frames per second. Thousands of fly lines in which different sets of neurons are activated (and fluorescently labeled) are screened with the goal to discover which neurons trigger or inhibit specific behaviors. Similarly, Schafer and collaborators have developed

assays for the widely studied worm *Caenorhabditis elegans* for automatically scoring egg-laying [3] and foraging [4] behaviors.

(2) Much effort is currently being put into tracing neurons [5, 6, 7] in 3-dimensional light and electron microscopic images in order to reconstruct the wiring diagram of the fly and mouse brain on the neuronal and the synaptic level, respectively. One of the major computational challenges in understanding the brain anatomy is the development of registration methods. Registration spatially transforms images of individual brains that differ in size and exact morphology in order to match them to a common reference brain.

(3) Studies of embryonic development aim at understanding how the information encoded in the genome leads to the development of a single cell into a complex organism. New imaging and image processing methods are jointly being developed to allow in toto imaging [8, 9], an approach to monitor and digitally reconstruct the development of an entire organism with non-stereotypic development such as *Drosophila* [10, 11, 12] and *Zebrafish* [13, 14]. Besides giving insights into the cell positioning and migration patterns, such reconstructions are used to monitor gene expression on the single cell level in developing embryos [15, 16] as well as in adults [17, 18].

(4) To identify genes or chemical reagents that cause a specific cellular phenotype, large-scale, often genome-wide screens are being carried out. These screens are used to determine gene function [19], protein-protein interaction [20] and subcellular protein localization [21, 22] as well as for drug discovery [23, 24]. For example, one goal of the Mitocheck Project [25] is to identify all human genes involved in mitosis [26]. To this end, each of the about 21,000 protein-coding genes has been silenced by RNA-interference [27] in cells with fluorescently labeled chromosomes. The resulting images are then classified automatically into one of 16 predefined morphological classes using support vector machines [28]. By clustering genes according to their phenotypic similarity, gene function can be predicted based on the genes in a cluster that have previously been characterized.

(5) Studies of intracellular dynamics try to decipher the molecular mechanisms that generate complex, fundamental processes within the cell such as

the assembly and positioning of the mitotic spindle, transport of organelles and vesicles or establishment of polarity axes. Such studies are often hypothesis-driven and the goal is to describe the time-dependent behavior, interactions and biophysical properties of the involved proteins. Fluorescently tagged proteins are monitored in single living cells to visualize the process under study at relatively high temporal resolution. The purpose of image analysis here is to achieve an accurate phenotypic description under various genetic and environmental conditions in terms of quantitative physical measurements such as velocity and motion direction, growth and diffusion rates, size, life-time, protein amount or force.

For example, Brangwynne et al. [29] studied the mechanism by which P granules localize preferably to the posterior half in the *C. elegans* one-cell embryo. Tracking the motion paths and measuring the size of fluorescently labeled P granules in time-lapse images showed that the mechanism is independent of cytoplasmic flow and is instead regulated by dissolution and condensation of the P granules in a liquid-like fashion. Another intensively studied intracellular structure is the microtubule, a constantly growing and shrinking tubular component of the cytoskeleton that is involved in chromosome segregation and cytokinesis during cell division (mitosis) and transport events throughout the cell cycle. To quantify the growth and shrinkage rates, the length, or the number of microtubules in fluorescence images, the computational problem of tracing entire microtubules or tracking fluorescently labeled microtubules tips has been addressed in [30, 31, 32, 33].

Automated image analysis in the context of biological studies is rarely just a matter of choosing an appropriate existing method or ready-to-use tool. Rather, active research to develop these computational methods is being carried out in the context of the specific applications and their associated image data. Therein, the two most fundamental challenges are (1) image segmentation, the partitioning of an image into semantically meaningful regions, and (2) in the case of time-lapse images, tracking of objects.

In this thesis we will study the process of centrosome formation and maturation in *C. elegans* embryos, which is an instance of the fifth category, intracellular dynamics. For this we will develop an analysis pipeline to track

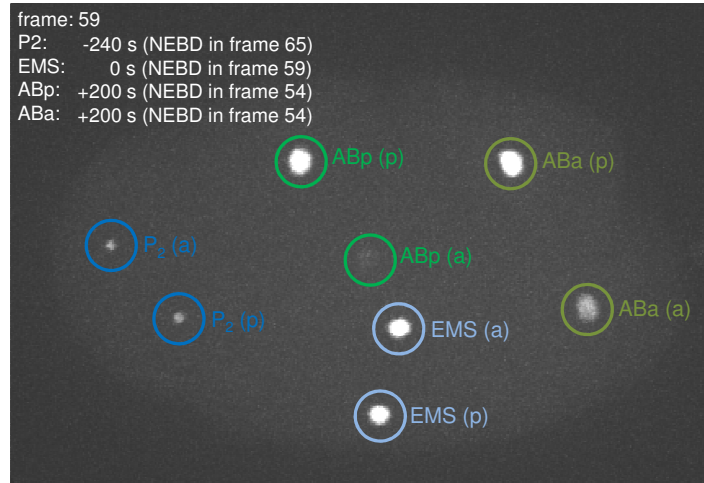


Figure 1.1: A typical example of a wild-type *C. elegans* embryo with γ -Tubulin::GFP labeled centrosomes in the 4-cell stage (maximum intensity projection). The centrosomes are tracked in 3D over time, annotated in terms of cell names and anterior/-posterior position and registered in time by the nuclear envelope breakdown (NEBD) as shown in the upper left corner with the algorithms developed in this thesis.

centrosomes in, and extract quantitative measurements from, 3-dimensional fluorescence time-lapse images (see Figure 1.1) in a fully automated manner. The upcoming sections will discuss the biological background of centrosomes and our specific question.

1.2 THE CENTROSOME IN *c. elegans*

The centrosome, discovered by Edouard Van Beneden in 1883 [34] and named in 1888 by Theodor Boveri [35], is a dynamic organelle found in all animal cells. It plays important roles during mitosis and duplicates exactly once per cell cycle. This section reviews our current understanding of the function and structure of the *C. elegans* centrosome and describes its life-cycle with respect to the images to be analyzed in this thesis.

1.2.1 Functions and Roles

Often referred to as microtubule organizing center (MTOC), the centrosome's primary function is the organization of microtubules by controlling their num-

ber, length, orientation and dynamics [36]. Anatomically, centrosomes constitute the poles of the mitotic spindle and are involved in the assembly [37] and positioning [38] of the spindle as well as in segregation of the chromosomes [39]. In *C. elegans* the centrosome is essential for spindle assembly as the depletion of many centrosome components leads to severe spindle defects [40, 41, 42, 43, 44, 45] and thus to embryonic lethality. In other systems such as *Xenopus* and *Drosophila*, bipolar spindles can form in some cells in the absence of centrosomes [46, 47, 48], but not in all [49].

The centrosome is further discussed to play an active role in cytokinesis [39, 50] and also has functions independent of its microtubule organizing capacity. The location of the centrosome defines a symmetry breaking cue that initiates polarity establishment in one-cell *C. elegans* embryo [51, 39]. Moreover, signaling proteins have been found to localize to centrosomes and promote timely entry into mitosis [52]. The general role of the centrosome in cell cycle progression is however still subject to debate [53].

Given its diverse cellular functions, it is not surprising that human diseases have been linked to centrosome abnormalities. For instance, in many cancer cells, an increased number of centrosomes is observed. Extra centrosomes result in the assembly of multi-polar spindles, which leads to chromosome segregation errors [39]. Whether this phenotype is the cause of genomic instability and therefore cancer is an open question [54]. Other studies argue that it is a failure in the centrosome function during asymmetric cell division, rather than genomic instability, that causes cancer [55].

1.2.2 Structure

The centrosome is a non-membrane bound, approximately spherical organelle composed of two domains. At the center it harbors a pair of orthogonally arranged centrioles which are surrounded by a mass of pericentriolar material (PCM) [39] (Figure 1.2). The centriole is a tubular structure with a nine singlet¹ array of microtubules [56] and recruits an initial small pool of PCM

¹ The centrioles in many other organisms are composed of nine doublet or triplet arrays of microtubules.

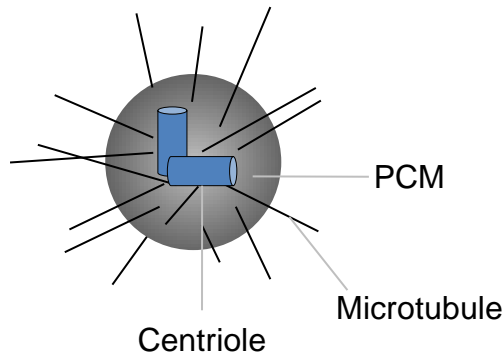


Figure 1.2: Schematic drawing of a centrosome.

components [57]. In this way, centrioles define the site of centrosome assembly [56] and thus the number of centrioles controls the number of centrosomes [58, 59]. The PCM is an electron-dense proteinaceous matrix [60], and is often described as an amorphous cloud [39], owing to the fact that its precise structure and physical properties are still poorly understood. Upon entry into mitosis, the PCM dramatically increases in size by the recruitment of additional components, a process termed maturation [61] that depends in parts on the aurora-A kinase AIR-1 [41]. Accompanied with the increase in size of the PCM is an increase of the number of microtubules that nucleate from the PCM in a γ -TUB dependent manner [62]. The review in [39] argues that the PCM can be further divided into two sub-domains, a core PCM and a peripheral PCM; fluorescence imaging showed that some of the PCM components localize in a more outer region of the PCM, giving them a donut-shaped appearance.

As stated above, centrosome assembly depends on the presence of centrioles. Conversely, work by Dammerman et al. [57] suggests that the PCM is in turn essential for centriole duplication because depletion of PCM components such as γ -TUB and SPD-5 impaired daughter centriole formation. However, the initial level of PCM present before maturation occurs, was sufficient for centriole duplication.

Thanks to fluorescent labeling, RNA interference (RNAi), as well as yeast two hybrid techniques, many centrosome components in *C. elegans* could be identified and their localization dependencies and interactions studied. A rough centrosome assembly pathway mainly based on the interactome in [63] is depicted in Figure 1.3. This pathway is essentially a collection of the results

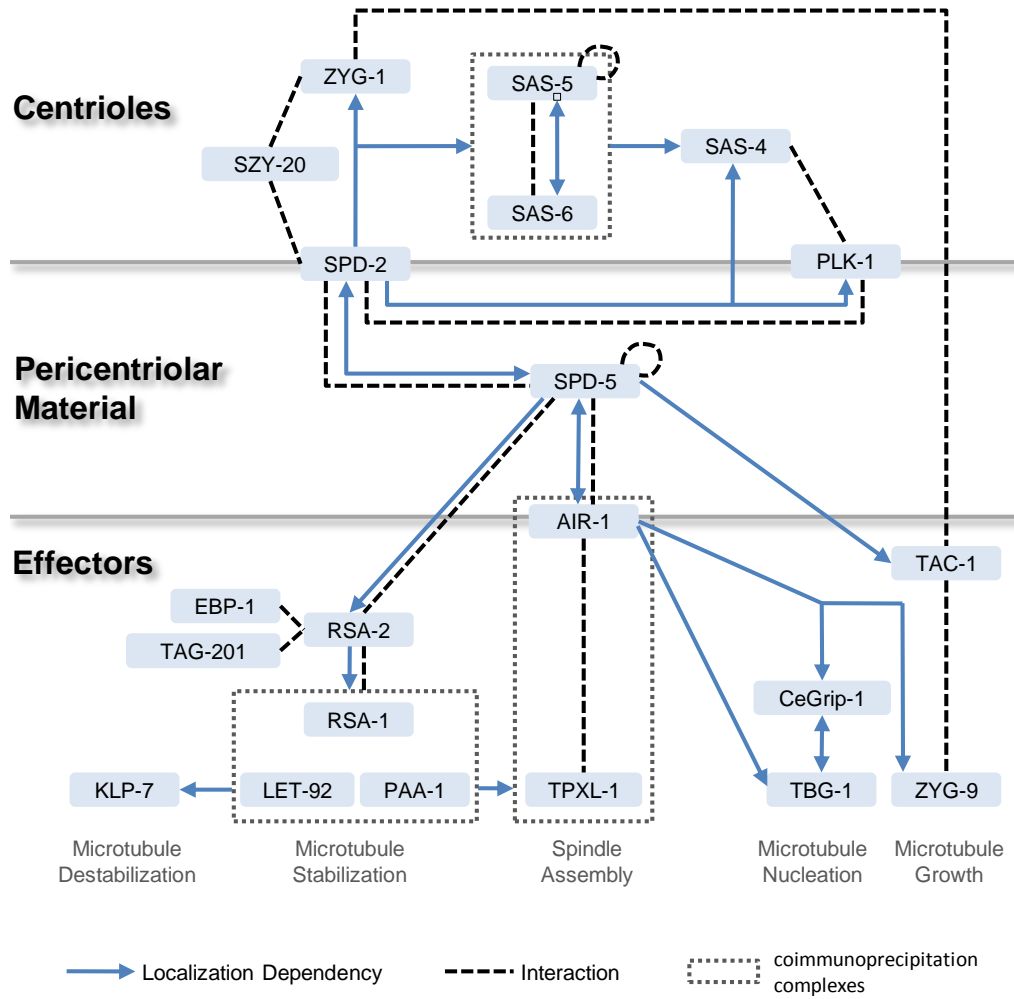


Figure 1.3: Assembly pathway of the centrosome. The hierarchy is based on localization and function of the known centrosome components. Adapted from the diagram in [63].

found by the work of several group in *C. elegans* over the last decade. The PCM is in parts composed of large coiled-coil proteins that are thought to form a scaffold for other proteins, so-called effectors, that subsequently dock to this scaffold in order to carry out the actual function of the centrosome [64, 65]. In *C. elegans* the centrosome components SPD-2 [66, 60, 40] and SPD-5 [42], two coiled-coil proteins that take a central position in Figure 1.3, could play this scaffolding role. It should be noted, however, that with the current pieces of information it remains difficult to get a complete picture of the mechanisms underlying centrosome assembly and maturation. Due to the complex interactions and mutual dependencies of the centrosome components, it is often not

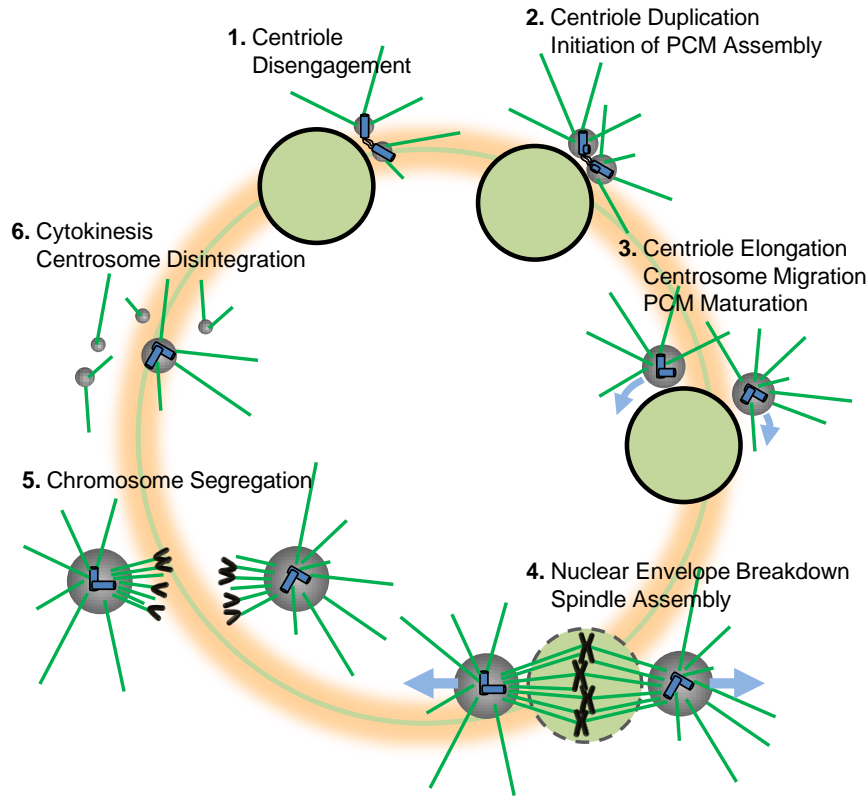


Figure 1.4: Life cycle of a centrosome.

possible to distinguish between direct and indirect functions of the individual components. For example, upon depletion of SPD-5 not only does centrosome maturation fail, but spindle assembly and polarization defects also occur [42]. However, the direct role of SPD-5 might be restricted to centrosome maturation, with spindle assembly and polarization defects being a secondary consequence of the absence of structurally intact centrosomes.

1.2.3 The Centrosome Cycle

The highly dynamic nature of the centrosome underlies a life cycle [67, 54, 49] that is coupled to the cell cycle in a mutually dependent manner [68, 49]. This section describes the centrosome cycle (see Figure 1.4) in conjunction with the events that are important for the images to be analyzed in this thesis.

Centrosomes are formed temporarily to orchestrate the organization of the microtubule fibers that segregate two copies of the DNA complement of the cell. The process of centrosome formation and eventual dissolution is broadly

outlined as follows. At the beginning of each round of the cell cycle the cell starts out with a pair of disengaged centrioles inherited from the mother cell. Each centriole begins to duplicate with a newly born, tightly engaged procentriole [54]. As the procentrioles elongate to form full length centrioles, more PCM begins to assemble around each centriole pair. The immature centrosomes, located next to the nucleus, then separate from one another and move to opposite poles of the nucleus while now continuously increasing in size. At the same time, the centrosome-nucleus complex positions itself at the center of the cell by means of the microtubules that grow outwards from the centrosome and attach to the cell cortex. As the nuclear envelope breaks down, additional microtubules grow towards and attach to the condensing chromosomes, assembling the mitotic spindle. Thereafter the centrosomes are pulled apart, toward opposite sides of the cell, by means of these fibers. As the centrosomes do so, they pull copies of each chromosome toward them to form the DNA content of each emerging daughter cell. Finally, the centrosomes come to a stop and rapidly disintegrate by releasing the PCM from the centrioles. The disintegration is in part due to cortical microtubule-mediated pulling forces that tear the PCM apart. This appears in the images to be analyzed as either an explosion into several pieces that then rapidly dissolve or as a dissolving of the entire centrosome. After completion of cytokinesis, a new nucleus forms in each daughter cell and another round of the cycle can begin.

1.3 BIOLOGICAL QUESTION: WHAT SETS THE SIZE OF A CENTROSOME?

The previous section showed that functional centrosomes are essential for the cells to properly divide and ultimately for the organism's survival. Yet their precise structure and the molecular mechanisms of centrosome assembly, maturation and disintegration are poorly understood. One aspect of these open problems is the size of the centrosome. As the *C. elegans* one-cell embryo, constrained by the eggshell, goes through rapid rounds of cell divisions, its cells become smaller with each division. By the 8-cell stage, for instance, the cells are only between 10 and 20 per cent of the original volume. As a consequence, the embryo also scales down the size of its intracellular structures during de-

velopment including the mitotic spindle and the centrosome. For the spindle this has the simple reason that a large one-cell spindle is too big to fit into a small cell later in development. The length of the mitotic spindle is in part set by the size of the centrosome [69]. A reduction in centrosome size leads to a consequent reduction in spindle length. However, how centrosome size is regulated is only vaguely known.

In this thesis we will investigate the molecular mechanism that sets the size of the centrosome in early *C. elegans* embryos. This is an important problem for two reasons. First, the general problem of how a cell adjusts the size of its dynamic structures is largely unsolved [70]. For instance, understanding how centrosome size is controlled might reveal a general size regulation mechanism. Second, a detailed study of centrosome size might give new insights into its structure and the assembly and maturation process.

Recent work has identified several components that influence centrosome size. While characterizing the centriole duplication factor SAS-4, Kirkham et al. [44] find that partial depletion of SAS-4 leads to incomplete centriole assembly and that these incomplete centrioles organize less PCM. Song et al. [45] identify the putative RNA-binding protein SZY-20 as a negative regulator of centrosome size. They show that in the absence of SZY-20, centrosomes are bigger than in wild-type and that the levels of centriolar and pericentriolar components including ZYG-1, SPD-2, SPD-5 and γ -TUB (see Figure 1.3) are increased. They also find that the centriole duplication factor SAS-6 influences centrosome size. In summary, from previous work it appears that the centriole plays a role in determining the size of the PCM but what this role is remains unclear. While these studies describe aspects of the molecular pathway that is important for centrosome size regulation, its underlying physical principle is not discussed.

In these studies centrosome size was assayed only in one- and two-cell embryos. For this, a single time point (e.g. metaphase) was chosen from live-cell recordings or the embryos were fixed. In either case, temporal information on how centrosomes grow during the cell cycle was ignored. Moreover, centrosome size is often described only qualitatively or, if quantitatively, coarse manual methods have been used. In this thesis we will overcome these short-

comings by developing an assay for computationally quantifying the kinetics of centrosome size in multi-cellular *C. elegans* embryos. Supported by the so-acquired time-resolved data in wild-type and several mutant embryos, we propose a limiting component hypothesis. In this simple model, the size of the centrosome is limited through the available amount of centrosomal protein in the cytoplasm. Proportionately distributing the cytoplasm, and therefore the centrosome components, to the daughter cells ensures that centrosome size scales down with cell size through embryonic development. We provide evidence that one component that limits centrosome size is the conserved centriolar and pericentriolar protein SPD-2 [66, 60, 40].

1.4 AIMS OF THIS THESIS

The overall goal of this work is to develop an assay for quantifying the dynamics of centrosomes in *C. elegans* embryos and apply it to the question of how the size of a centrosome is regulated. This is an interdisciplinary area of research in computer vision and molecular biology.

Towards the biological question the goal is to quantitatively characterize how centrosomes grow throughout embryonic development, how this is affected under various genetic manipulations and how centrosome size correlates with cell size and the amount of protein in the cytoplasm.

Computationally, the goal is to develop and implement a suite of algorithms to analyze 3D time-lapse images of fluorescently labeled centrosomes and cell membranes in *C. elegans* embryos. This includes

- detecting and tracking centrosome positions
- a definition and measuring method for centrosome size that is robust to imaging artifacts
- determining cell identities within *C. elegans*' invariant cell lineage
- registering the tracked centrosomes in time
- a segmentation algorithm for fluorescently labeled cell membranes to quantify cell volumes.

Estimating the size of centrosomes, which appear as blurred spots in the images, is a task at which humans are prone to misjudgements. As a result, manual measurements are biased towards a desired outcome of the respective experiment or miss subtle but important phenotypes. Although no automatic method can compute the perfectly error-free measurement in the presence of imaging noise and artifacts, the overall system that we develop should alleviate both of these problems. Additionally, our goal is to increase the low throughput inherent to manual or semi-automatic methods by fully automating all analysis steps.

AUTOMATED IMAGE ANALYSIS FOR AN ASSAY OF CENTROSOME SIZE

This chapter describes the assay we have developed for quantifying centrosome size and other parameters over time in wild-type and mutant *C. elegans* embryos. We first explain the imaging procedure (Section 2.1) and then the algorithms to analyze the resulting 3D time-lapse images. Figures 1.1 and 2.1 show several example images. The primary steps are centrosome tracking (Section 2.2) and measurement (Section 2.3). Based on the established tracks, our system also determines the biological identities of the centrosomes within *C. elegans*' invariant cell lineage (Section 2.4) and recognizes the nuclear envelope breakdown [71] in each cell division (Section 2.5). The idea is to convert a set of centrosome movies into aligned time series of centrosome statistics ready for biological interpretation.

2.1 IMAGING CENTROSOMES IN LIVING EMBRYOS

The first step in setting up the centrosome assay is to figure out an optimal imaging protocol. This includes selection of an appropriate microscope and careful experimentation with its parameters in order to obtain the best possible images for *automated* analysis while keeping the embryo alive and viable during the imaging process. To monitor centrosomes in living embryos we chose to use spinning disk confocal microscopy, a fluorescence microscopy technique that provides a good trade-off between acquisition speed and sensitivity for our application. We will briefly review the fundamentals of fluores-

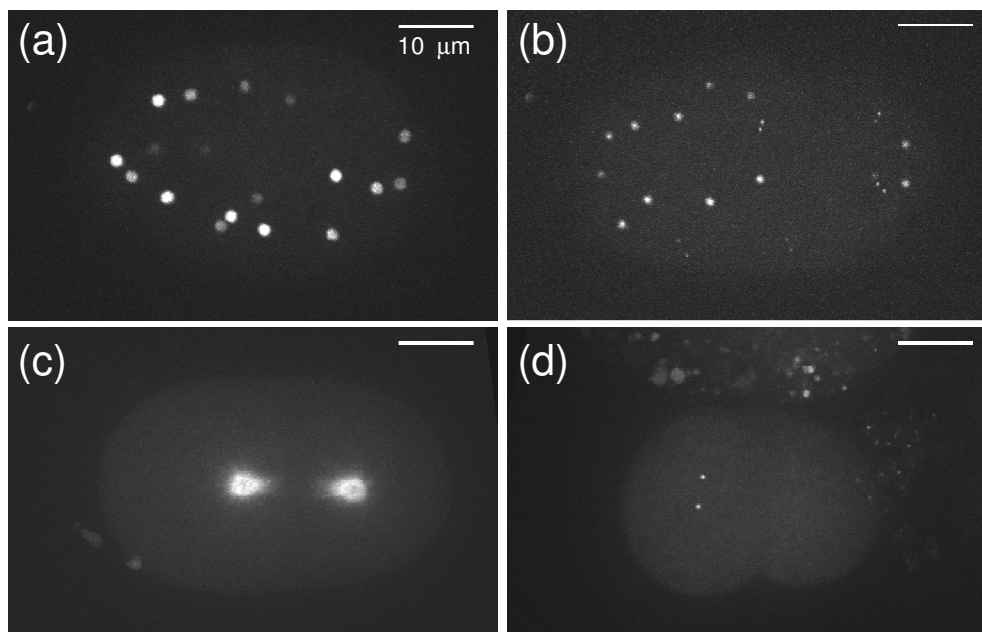


Figure 2.1: Examples of *C. elegans* embryos with different centrosome markers. Images are maximum intensity projections of confocal time-lapse stacks. (a) Wild-type, RSA-1::GFP (b) Wild-type, SPD-2::GFP at approximately the same stage as the embryo in (a). The SPD-2 signal is strong at the centriole (bright tiny dot at the centrosome center) and weak at the PCM. (c) Wild-type, AIR-1::GFP. The AIR-1 signal is also visible on microtubules. (d) 2-Cell mutant embryo (*ani-2*(RNAi)) that is smaller than wild-type with a neighbouring older embryo, γ -TUB::GFP.

cence microscopy, discuss its limitations and describe the protocol for imaging centrosomes.

Fluorescence microscopy is the de-facto standard in biology to visualize structures that are too small to be resolved by the human eye. Fluorescence is the emission of light by certain molecules, called fluorophores. By exciting a fluorophore with light of a characteristic wavelength produced by a laser its energy level rises to an unstable, excited state [72]. As the fluorophore returns to the ground state it emits the absorbed energy in parts as heat and the remaining energy as light of a longer wavelength. By blocking the excitation light with wavelength dependent light filters only the light emitted from the fluorescent structures will contribute to the image. A protein of interest can be visualized using this principle by genetically fusing it to a fluorophore. There are thousands of different fluorophores available for this purpose [73] with green fluorescent protein (GFP) and its variants being the most common ones.

There is a fundamental limit on the spatial resolution in all light microscopy techniques due to diffraction of light by the sample and the microscope lenses. Diffraction causes a light point source to appear as a blurred spot in the image. This response of an imaging system to a light point defines its point spread function (PSF) and therewith its resolution. Two point objects closer to each other than the achievable resolution will appear as a single spot in the image. For confocal microscopy the best achievable resolution in the x/y-direction is about 180 nm and in the z-direction about 500 nm [74]. However, the resolution limit caused by diffraction can be overcome. One reason for this is that the location of a single point object can be determined with much higher precision than the optical resolution. This principle is used in super-resolution imaging approaches such as photoactivated localization microscopy (PALM) [75] and stochastic optical reconstruction microscopy (STORM) [76]. Although the first super-resolution microscopes are now commercially available, it is still an experimental field. Often special fluorophores that can be switched on and off are required, the imaging depth is very limited and time-lapse imaging is not yet done on a standard basis.

The excitation of the fluorophores with a laser beam not only causes the desired emission of light from the tagged protein molecules, but also has negative effects that have to be minimized. High energy doses are lethal or at least toxic for the embryo (photo-toxicity), resulting in abnormal development. Another problem is photo-bleaching. Each fluorophore molecule can fluoresce only a finite number of times, because it will eventually be destroyed by the excitation light [72]. This results in loss of image intensity over time and hence lower image quality. Therefore, the laser power, the exposure time, imaging interval, z-spacing and number of z-planes have to be adjusted so that the embryo develops normally and photo-bleaching is minimized. On the other hand, to increase the signal-to-noise ratio of the images higher laser power or longer exposure time is necessary.

Several variants of fluorescence microscopy techniques exist. They differ in acquisition speed, penetration depth, resolution and detection sensitivity. Laser scanning confocal and two-photon microscopy are too slow for live-cell imaging of centrosomes, while wide-field fluorescence microscopy does

not allow optical sectioning. Spinning disk confocal microscopy provides the best trade-off among the available fluorescence imaging techniques for our application. A spinning disk confocal microscope consists of a rotating disk with multiple pinholes, which makes it possible to scan many points of the sample at the same time. The photon counts are recorded by a charge-coupled device (CCD) camera to produce an image of a 2-dimensional plane and a piezo stage moves the sample with nanometer precision in the z-direction for optical sectioning. To image centrosomes in the developing embryo we use worm strains in which one of the centrosomal proteins, such as γ -TUB, SPD-2 or SPD-5, is fused to GFP or YFP. Imaging takes up to about two hours to capture embryonic development from the one- to 16-cell stage. In most experiments, however, we image only to the end of the one- or four-cell stage. An often encountered practical problem is to find one-cell embryos at the beginning of the centrosome cycle. Although it takes only about 6 seconds to acquire a 3D stack, we image only at time intervals between 20 and 50 seconds, depending on the total number of stacks and the number of planes per stack, to minimize laser induced damage of the embryo as well as said photo-bleaching. Typically, a 3D stack consists of about 50 z-planes with 512×512 pixels and 100 time points are sampled resulting in roughly 2 gigabytes of data per experiment. The pixel size in x and y is 133 nm and the distance between two z-planes is set to 500 or 600 nm. All these settings are chosen so that the embryo appears to develop normally; however a relatively small increase in total laser exposure would be harmful for the embryo. The embryo is usually stationary, but small movements cannot be excluded and manual adjustments of the focus between two time point samplings are sometimes necessary to account for drifts in the z-direction.

2.2 CENTROSOME TRACKING

2.2.1 *Tracking Challenges and Simplifying Properties*

The task to automatically track fluorescently labeled centrosomes in a developing *C. elegans* embryo is a multiple-object tracking problem with an unknown,

time-varying number of objects. Due to the asynchrony of the cell divisions and the fact that we wish also to track mutational variants, we cannot assume that the number of centrosomes is a power of two at any given time. While many of the difficulties in solving this problem are common to any tracking task in fluorescent imagery, the force driven movement of the centrosomes as well as the relative large thickness of *C. elegans* embryos impose additional challenges.

1. Low laser exposure to minimize photo-toxicity and photo-bleaching results in high noise and low contrast.
2. The signal attenuates significantly towards the bottom of the embryo. This can cause a more than 10-fold contrast difference between two centrosomes with equal amounts of fluorescent protein. Some of the centrosomes close to the bottom of the embryo are therefore extremely dim and hardly detectable by the human eye.
3. Auto- and off-target fluorescence results in disturbing background structures and non-centrosome particles. In particular, the polar body [77], which is extruded from the embryo after meiosis, remains inside the eggshell and is often visible as a fluorescent object.
4. At the beginning of the centrosome cycle (see Section 1.2.3) the distance between the two centrosomes of a cell is below or close to the optical resolution. Especially in the z-direction, this complicates the detection and accurate measurement of two close centrosomes.
5. Microtubule-mediated pulling forces can induce sudden accelerations and changes in motion direction leading to large frame-to-frame displacements. Pulling forces also result in the disruption of the centrosome into many pieces at the end of its life cycle.
6. Structures in the surroundings of the embryo such as remains of the embryo's parent or a neighboring embryo represent additional disturbing objects. This is especially the case for some mutant embryos with a desired phenotype that occurs only with low probability. For these valu-

able mutants it is not always possible to ensure the highest achievable image quality.

7. Small movements of the whole embryo in the x- and y-direction and significant movements in the z-direction occasionally occur.
8. Some of the centrosome markers also label microtubules or the centrioles resulting in additional local maxima in a centrosome's intensity profile.

On the other hand, there are also properties that simplify the tracking problem. Once starting to migrate to opposite poles of the nucleus, the centrosomes are big enough and well separated so that they do not merge or split. The great range in centrosome size and brightness is challenging for the detection, but provides valuable information to resolve object correspondence. Monitoring the centrosome cycle up to the 16-cell stage means that no more than 32 objects need to be tracked simultaneously (in wild-type embryos). Furthermore, a centrosome is roughly ellipsoidal shaped (almost perfectly spherical before onset of disintegration) with approximately Gaussian intensity distribution, its size either increases or stays constant with time up to its disintegration and there is a biophysical limit on its maximum velocity.

2.2.2 *Related Work*

Tracking of objects in image sequences is of great interest for many practical applications. Beyond the analysis of live-cell microscopy images examples include robotics, video analysis of sport games or radar surveillance in air traffic control. The proposed solutions in each of these fields rely on different assumptions. In air traffic control, for instance, well-defined motion models can be used to predict the position of the object at the next time point [78]. For the analysis of sport games, views from multiple cameras can help to resolve ambiguities [79]. The large number of publications indicates that no general solution for object tracking exists and so we will here concentrate on tracking approaches with application to fluorescence microscopy images.

There are two broad classes of tracking methods, probabilistic methods and distance based methods. Most of the proposed techniques divide the tracking

problem into two independent stages. First, in the detection stage the objects are located in each individual frame. In the second stage those detections that correspond to the same physical object are linked into trajectories, referred to in the literature as the correspondence problem, object linking, trajectory establishment or data association. Approaches that *simultaneously* perform object detection and linkage include a method based on dynamic programming [80] that can track one single particle at a time as well as the microtubule tip tracker [81] that can rely on linear motion patterns. Simultaneous detection and linkage directly allows the detection to take temporal information into account, but it is conceptually and computationally complex. Separating detection and linkage simplifies the problem by decoupling two difficult problems. The achievable tracking performance then depends on how much irrelevant information can be filtered and how much useful information can be preserved by the detection method.

Probabilistic methods formulate the tracking task as a state estimation problem [81]. The state x_t of an object is a vector of object properties that includes position and often velocity and intensity. A measurement z_t can be a set of detected object positions, extracted features thereof or the entire image. A measurement model $p(z_t|x_t)$ expresses the likelihood of observing the measurement z_t given the current state x_t and a state evolution model $p(x_t|x_{t-1})$ predicts the state x_t given the previous state x_{t-1} . If the state vector only contains information derived from object position, the state evolution model is a physical motion model of the objects to be tracked. The goal then is to derive the posterior distribution $p(x_t|z_{1:t})$ of the current state given all measurements up to time t [82]. Intuitively speaking, the idea is to directly describe the imaging process and the dynamic behavior of the objects by probabilistic models (i.e., the measurement and the state evolution model, respectively) and then to inverse the model in order to infer the hidden state of the system from the sequence of noisy measurements. Analytically inverting the model is often not possible but approximations can be obtained by numerical simulation methods, in particular sequential Monte Carlo methods [83] (also known as particle filters).

These methods have been applied for tracking of endocytic vesicles labeled with quantum dots [84], microtubule tip tracking [81, 85], virus particles [86], single membrane-associated molecules [87] and centrosomes in *C. elegans* embryos [88]. The concept of probabilistic methods is very appealing because it allows to directly formulate prior knowledge. Concrete implementations, however, suffer from a number of drawbacks. First, numerical simulation is computationally demanding in time and space. As a result, the size of the images and the number of objects that can be handled is limited [87]. For example, in [85] a manually specified region of interest containing 10-20 objects is used for 2D image sequences with 20 time points. Second, the reported tracking performance is not always satisfying. The virus tracking algorithm in [86] achieves an average tracking performance of about 80 per cent based on three movies, and for the centrosome tracking algorithm in [88] a considerable decrease in performance is reported when the embryo has passed the two-cell stage. Finally, while all tracking methods depend on a number of parameters that need to be adjusted, probabilistic methods are especially prone to parameters without physical meaning such as the number of (simulation) particles, the choice of the importance density function [83], probability thresholds for track initialization and termination [81] or the gain parameter in case of the Kalman filter [89].

Distance based methods are based on defining a distance measure (or cost function) that reflects the chance that two detected objects belong to the same physical object and a strategy to assign objects to each other based on the distance measure. While conceptually often simpler than probabilistic methods, the tracking performance critically depends on how well the objects can be detected and how effectively prior knowledge about the objects can be encoded via the distance measure and the assignment strategy.

One widely used technique is the nearest neighbor strategy due to its simplicity and low computational demand. The pair of objects p at time t and q at time $t + 1$ with the smallest spatial distance are linked. Then, excluding all associations involving p or q , the pair with the second smallest distance is linked, and so on. For instance, this approach has been used for tracking lipoprotein receptor molecules [90]. In general, it is applicable when the ob-

jects move sufficiently separated from each other, but quickly brakes down under high object density (relative to the frame-to-frame displacement of the objects) or in the presence of near-by false detections. Bao et al. [16] have developed an extended version of the nearest neighbor algorithm for tracking nuclei / chromosomes in *C. elegans* embryos via Histone::GFP. Their goal was to trace the cell lineage up to hundreds of cells in one color channel and use a second color channel to monitor gene expression. Instead of only the nearest neighbor they consider a small number of potential matches with comparable distance as the nearest neighbor. To resolve conflicting situations, they exploit the fact that sister nuclei have similar intensity and size and can distinguish between young and older nuclei based on their shape. To accurately detect the nuclei they additionally take advantage of the fact that due to small frame-to-frame movement a nucleus in the current frame significantly overlaps with itself in the previous frame.

Sbalzarini and Koumoutsakos [91] present a more generic tracking algorithm for tracking fluorescent particles in 2D images with applications to studies of lipoprotein molecules, virus particles and quantum dots. Object detection begins by finding all local intensity maxima above a user-defined intensity percentile threshold and subsequently classifies between false and true positive calls based on zero-th and second order moments of the spot intensities. The zero-th order moment is the total intensity of the spot and the second order moment corresponds to the width of the spot. Two assumptions are made in this approach. First, the true positives form a dense cluster in the 2-dimensional momentum space (i.e., they all have similar momentum values) while the false positives are more spread around this main cluster. This essentially means that the true objects must be relatively uniform in their appearance. Second, the majority of the calls are true positives. Under these two assumptions there will be many other calls nearby a true positive in the momentum space that contribute to a score defined as a sum over the set of all calls. For a false positive on the other hand there are much fewer supporting calls, hence giving a lower score. A threshold on this score is then used to make the distinction between true and false positives. To link the detected objects into trajectories a cost function that includes object displacement and the

difference of the two intensity moments is defined. Interestingly, it includes no normalization for the different ranges of values of these measures of object coherence. The linking strategy then is to minimize the total linking cost under the constraint that each object must be linked either to another detected object or to a dummy object, which is a linear assignment problem [92]. To handle missed detections and object occlusions the linking is done not only between frames t and $t + 1$ but for multiple frames ahead, i.e., for all pairs of frames t and $t + r$ for $r = 2 \dots R$ and R being a user-defined value. However, in a tracking scenario with disappearing and newly emerging objects like centrosomes, this multi-frame linking strategy gives a higher chance to incorporate false positives into the trajectories even with modest false positive rates.

While the approaches discussed above establish the trajectories in a single pass, Jaqaman et al. [93] do so in two steps, also utilizing the linear assignment problem. They first link objects between consecutive frames. The cost function here is the squared Euclidean distance between the objects and conservative thresholds on this distance are used to lower the risk of false assignments. This results in many short track segments. In the second step, the track segments are linked to form complete trajectories. For this, problem-dependent cost functions and thresholds for gap closing and merge and split events are defined for pairs of track segments and globally minimized for all track segments and event types simultaneously. Thus, the second step innately takes temporal information into account. They demonstrate the applicability of the method for characterizing the motion behavior of single receptor molecules and for life-time analysis of endocytic structures. This tracking framework requires the object detection method, the assignment cost functions and the threshold parameters to be specified for the problem at hand.

In some of the tracking applications, it is sufficient to select a subset of (largely) error-free trajectories for the subsequent analysis. For example, to compute a velocity distribution of tracked microtubules, a good estimate does not require that all microtubules present in the image sequence be considered. Whenever it is valid to assume that a sufficiently large subset of the tracked objects is statistically identical to the entire population, biologically useful results can be obtained despite incomplete tracking. For the centrosome assay

incomplete or inaccurate tracking is not tolerable, because each centrosome in a *C. elegans* embryo has a distinct biological fate and the centrosome tracking algorithm is part of a fully automated image analysis pipeline in which subsequent steps depend on the established trajectories.

In summary, the use of problem-specific prior knowledge (and appropriate parameter settings) is indispensable for any of the discussed tracking methods and determines to a high degree the concrete implementation within the context of the respective biological study. One generally applicable and computationally efficient but abstract framework for object tracking is the linear assignment problem, which can be equivalently formulated as bipartite graph matching. In the next section we will use this framework for our application.

2.2.3 Overview of the Tracking Algorithm

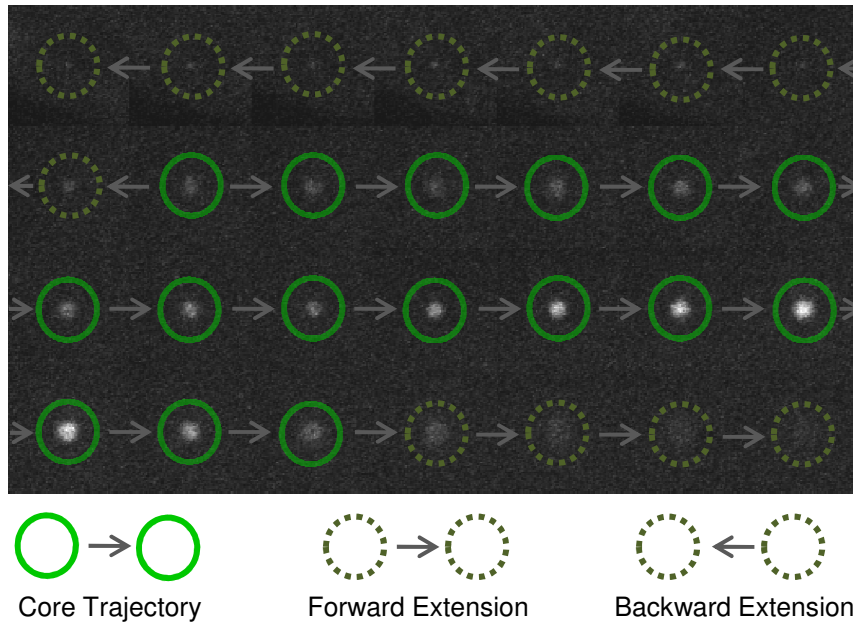


Figure 2.2: A typical centrosome trajectory established by first finding the core trajectory and then forward and backward extending the core trajectory.

Automated tracking of centrosomes is the computational core component of our assay to study the dynamics of centrosomes in developing *C. elegans* embryos. To robustly solve this problem, we formulate it as a graph matching problem and incorporate prior knowledge of the centrosome cycle (see Section

1.2.3). In that we use the novel approach of first reliably finding those subsegments of a trajectory in which the centrosome is bright and large, and then extending these ‘core’ trajectories backward in time to catch the initiation of the centrosome where it is very small and dim, and forward in time to catch the disappearance of the centrosome when it appears to ‘explode’ or dissolve (Figure 2.2).

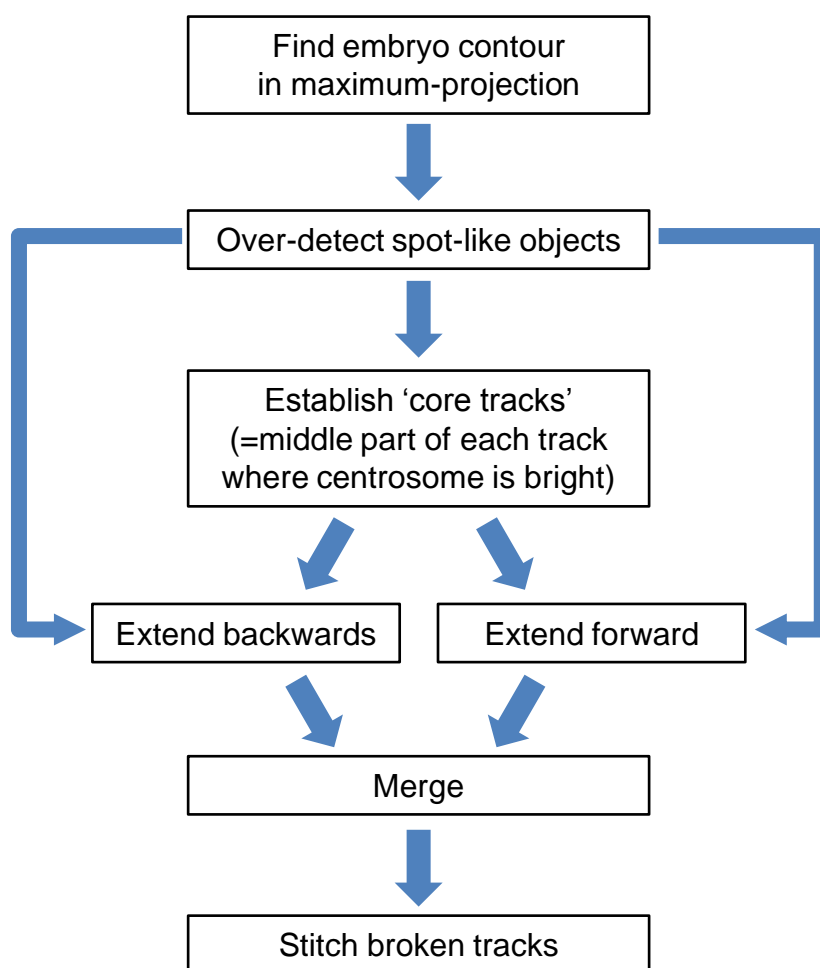


Figure 2.3: Workflow of the centrosome tracking algorithm.

The workflow of our algorithm is illustrated in Figure 2.3. It first finds the contour of the embryo and proceeds by detecting spot-like objects within the embryo region. In contrast to other tracking approaches, the object detection step is tuned to very high sensitivity so that missing detections (false negatives) virtually do not occur; however, this may result in a significant rate of false detections (false positives). We then link the objects between consecutive frames into trajectories with a two-stage approach. In the first stage we find

the core (sub-)trajectories that we are confident are correct, in that the object assignments based on appearance and displacement are mutually consistent and the individual spot calls used are clearly not noise. Every core trajectory, which is invariably the middle portion of a complete trajectory, is then extended in the second stage to find the best set of complete trajectories over the set of all called objects. Occasionally, two non-overlapping core trajectories are found for a given complete trajectory in which case, after the extension step, they abut in time. In a final phase, we recognize these situations and concatenate the trajectories involved. The key idea is to first understand what is certain, and then extrapolate the rest around that.

2.2.4 *Object (Over-)Detection*

We first over-detect possible centrosomes in each 3D stack by finding all objects that are to first approximation a Gaussian spot of some radius. Each stack I is first smoothed slightly to give I_S by applying a 3D-Gaussian filter of radius 0.6 pixels chosen as per [94]. We consider there to be an object at position p if (i) the value at p in I_S is locally maximal, and (ii) the normalized cross-correlation [95] of the raw image I with a Mexican hat filter with standard deviation σ centered at p is greater than τ . Several discrete values of σ are used and we find that setting τ to 0.70 (Figure 2.4) gives a very low false negative rate (i.e., it is rare to not detect a centrosome) while leaving us with a manageable number of false positive signals (e.g., noise and autofluorescent ‘spots’).

We then refine the center $(x(p), y(p))$ of each putative centrosome and compute its radius $\rho(p)$, its mean-illumination-to-background ratio $\eta(p)$, as well as its total illumination $\iota(p)$, by finding the best least squares fit of the spot at p to an elliptical 2D-Gaussian model with a (background) offset $b(p)$ using the trust-region-reflective algorithm [96] implemented in Matlab. We do so only in the 2D plane with the same z -value as p , primarily because 3D statistics are less reliable due to the relatively poor z -resolution of confocal microscopy as well as movement of the centrosome while it is imaged. All parameters are free and real-valued so that in the end we have a sub-pixel center $(x(p), y(p))$,

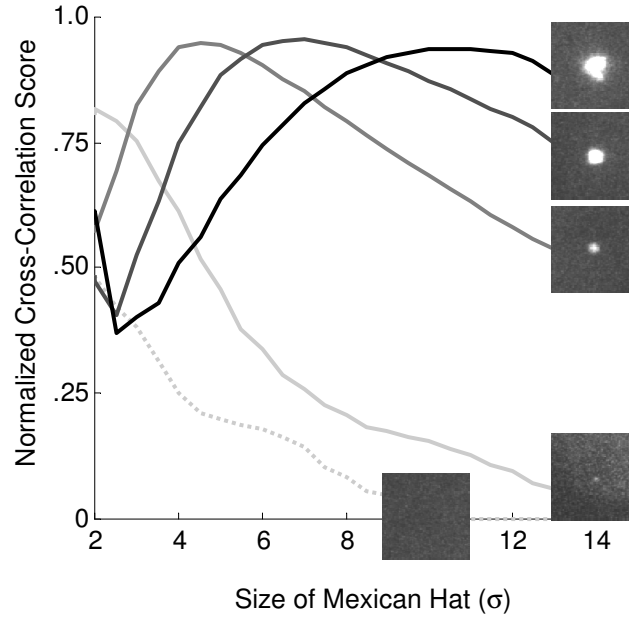


Figure 2.4: Response of Mexican hat filters of various sizes to centrosome and background image patches.

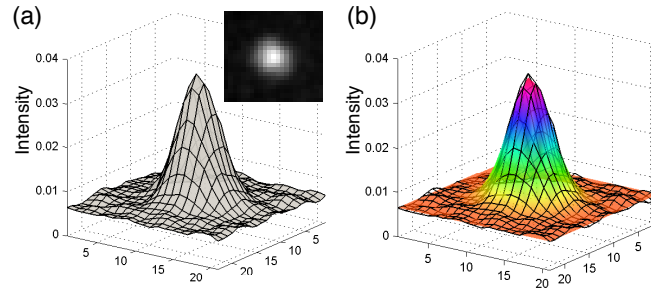


Figure 2.5: Measuring radius, total intensity and background level by Gaussian fitting. (a) An image of a centrosome and its intensity profile. (b) Gaussian fit (colored surface) overlaid on the centrosome intensity profile (meshed surface).

angle $\theta(p)$ of the principle axis, standard deviations $\sigma_p(p)$, $\sigma_m(p)$ along the principal and minor axes, peak height $h(p)$ and background offset $b(p)$. Figure 2.5 shows an object intensity profile and the fitted Gaussian, respectively. From this Gaussian, we then have radius $\rho(p) = 2\sqrt{\sigma_p(p) \cdot \sigma_m(p)}$, $\eta(p) =$ mean intensity within $\rho(p)$ divided by $b(p)$, and $\iota(p) =$ volume under the Gaussian without $b(p)$.

Occasionally, when centrosomes are very small and near each other, they are recognized as a single object. The signature of these cases are that the ratio $r = \sigma_p(p)/\sigma_m(p)$ is large and $\sigma_m(p)$ is small. In these cases, we try

fitting the signal to *two* Gaussian models and if we have a better residual and both σ -ratios are smaller than r then we accept the two spots scenario and associated Gaussian models.

Using a series of sized Mexican filters and a fixed threshold τ will detect the same centrosome multiple times at slightly different radii and different positions. To solve this problem, in a post-processing step we iteratively detect pairs of objects p and q that are closer together than the sum of their radii $\rho(p) + \rho(q)$ and discard the dimmer object (i.e., the one with the smaller η -value). This eliminates on average $54\% \pm 10\%$ (s.d.) ($n=40$ movies) of the initially detected objects and is clearly conservative as two centrosomes do not intersect in space.

Another problem that may cause tracking errors are objects outside the embryo such as the polar body or remains of the embryo's parent. We remove these by first detecting the contour of the embryo in the maximum intensity projection. Our contour detection algorithm starts with finding a rough estimate of the embryo region in the first frame of the movie using Otsu thresholding [97] and then applies marker-controlled watershed segmentation [98] to each frame to find the actual embryo contour. In this, we propagate the segmentation of the previous frame as marker for the next frame through time (see also Section 3.1). We then discard all object calls that are outside the convex hull of the detected embryo contour. Using the convex hull instead of the originally detected contour minimizes the risk of discarding centrosome objects close to the embryo boundary, which might be imperfectly segmented. This step eliminates on average $28\% \pm 19\%$ (s.d.) ($n=40$) of the objects from the last step.

We now have a set of possible centrosome centers and an estimate of their radius. Empirically, it is exceeding rare for a centrosome to not be in this set, and anywhere from 0 to 90% (average: $34\% \pm 14\%$ (s.d.), $n=40$) of the remaining spots are false positive signals (e.g., noise, etc.) depending on the time point and acquisition conditions.

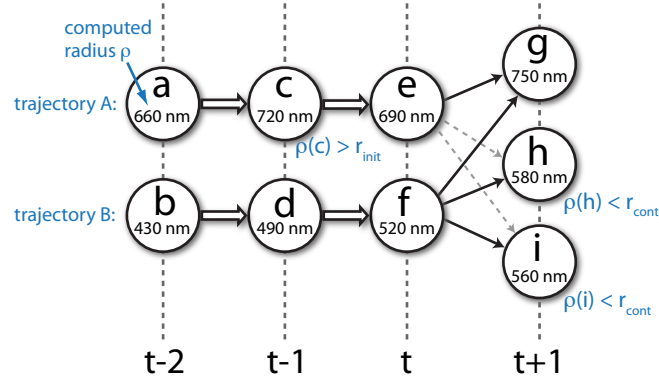


Figure 2.6: Illustration of a radius hysteresis situation. Two trajectories A and B have been established up to time t . Object c on trajectory A is larger than r_{init} (700 nm). The edges (e, h) and (e, i) that are associated with trajectory A and involve objects smaller than r_{cont} (600 nm) are thus radius hysteresis edges and will be removed from the matching hypothesis graph.

2.2.5 The Bipartite Graph Matching Paradigm

After the sensitive detection of putative centrosomes in all frames as described in the previous section, the next step is to link those objects that correspond to the same physical centrosome into trajectories and filter out false positives. We break this problem down into two steps. In the first step we establish only a segment of each centrosome's full trajectory where the centrosome is significantly brighter than the background and hence easy to distinguish from noise objects. We refer to this part of the trajectory as the 'core' trajectory. In the second step we extend the core trajectories backward and forward in time to obtain each centrosome's full trajectory. Both steps are implemented in the same bipartite graph matching framework for which we define an assignment cost function based on object displacement and appearance. Bipartite graph matching as used here has the property of maximizing the number of one-to-one object assignments between consecutive time points while minimizing the total assignment cost. This section describes the framework and how we incorporate prior knowledge. The next two sections then show how we use the framework to find reliable core trajectories and extend the core trajectories to full trajectories.

Let $V_t = \{p_1^t, p_2^t, \dots, p_{m_t}^t\}$ be the set of detected objects at time point $t = 1 \dots n$ and $V = V_1 \cup \dots \cup V_n$ be the set of all objects. Note that for each time point the number m_t of objects can be different. We define a weighted directed graph $G = (V, E, w)$ where the set of edges represents all possible matching hypotheses between objects detected at successive time points. There is an edge (p, q) between objects $p \in V_t$ and $q \in V_{t+1}$ if and only if:

1. $\|p - q\| \leq d_{\max}$: d_{\max} models the biophysical limit on centrosome velocity (300 nm/sec) and also takes into account cell movement within the embryo as well as movement of the entire embryo.
2. if $\rho(p) > r_{\text{init}}$ then $\rho(q) > r_{\text{cont}}$: Biologically the centrosome is always increasing or constant in size with time up to the beginning of its dissolution, but due to measurement and estimation error the computed radius varies. Empirically, we have observed that once the radius exceeds $r_{\text{init}} = 700$ nm (well above the resolution limit), it never becomes smaller than $r_{\text{cont}} = 600$ nm. We call this criterion radius hysteresis and in what follows we require that it not only be true of edges but also of sequences of edges, i.e., once a centrosome becomes larger than r_{init} on a trajectory it does not then become smaller than r_{cont} later in the trajectory (Figure 2.6).

These two simple conditions eliminate a large number of potential correspondences that are certainly incorrect.

The weight $w(p, q)$ of an edge of the graph is intended to reflect the chances that p and q represent the same centrosome. Biologically we know that centrosomes move in a relatively smooth manner and that facets of their appearance such as their radius and intensity also change gradually. Thus we expect p and q to be close together and for their appearance to be similar. We formalize this by defining the weight or ‘distance’ between p and q as:

$$w(p, q) = (1 - \alpha) d_S(p, q) + \alpha d_A(p, q) \quad (2.1)$$

where

$$d_S(p, q) = \|p - q\| / \sigma_S \quad (2.2)$$

and

$$d_A(p, q) = \sum_{a=1}^m \left| \frac{\varphi_a(p) - \varphi_a(q)}{\min(\varphi_a(p), \varphi_a(q))} - \mu_a \right| / m\sigma_a \quad (2.3)$$

where d_S is a normalized spatial (Euclidean) distance, d_A is a normalized relative-appearance distance and $\alpha \in [0, 1]$ is a free parameter (to be selected by the algorithm) that weights the contributions of these two aspects of coherence; μ_a are the means and σ_a and σ_S the standard deviations of the frame-to-frame differences of the respective feature. In our implementation we use two appearance features. $\varphi_1(p)$ is the radius of p and $\varphi_2(p)$ is the total intensity of p . The distances are normalized using the concept of a z-score [99] wherein the measure is the number of standard deviations above or below the mean value of a distribution of ‘raw’ scores. In our case, the set of edges D contributing to the distribution (and therewith to μ_a , σ_a and σ_S) is decided dynamically during the algorithm and so will be described later. Note that we do not subtract the mean of the spatial distance, μ_S , in d_S because this would favor links between objects that move by μ_S over links where the objects do not move significantly.

A solution S to the tracking problem is a set of edges in G that form a number of vertex-disjoint paths, each corresponding to a centrosome’s lifespan. We could add additional constraints on these paths for wild-type data, but since we want solutions for mutants as well, we can in general not require more structure than this. However, suppose that there were no false-positive objects in the graph, that is, every vertex truly represents a centrosome. Then our solution in this special case is the set of paths that involves the maximal number of edges, and if there are several, then we want the set of edges with minimum sum weight. This is equivalent to finding a minimum-weight bipartite matching between V_t and V_{t+1} for each time t , and we can do so using the Hungarian algorithm [100]. It is a standard technique to use bipartite matching even when there are false positive objects, and we will do so numerous times with different parameter settings as a way of discovering these false pos-

itives. Moreover, as we solve each bipartite matching for successive value of t , we eliminate any edges between V_t and V_{t+1} whose addition to the trajectories up to time t would result in a trajectory with radius hysteresis (Figure 2.6).

2.2.6 Finding Reliable Core Trajectories

The first step is to find the middle part of each centrosome's trajectory where the centrosome is large and bright. We refer to this part of a trajectory as the core trajectory. While it is desirable to establish as much of each trajectory as possible, it is more important that the core trajectories be error-free.

For several thresholds $\omega \in \Omega$, we consider the subgraph of G restricted to the objects that are brighter than ω , i.e., $V(\omega) = \{p \in V \mid \eta(p) \geq \omega\}$ ($\eta(p)$ is the mean-illumination-to-background-ratio, see Section 2.2.4 for the definition of $\eta(p)$). Note that the larger the threshold ω , the fewer false-positive objects there are in $V(\omega)$ until at some value, not yet known to us, practically no false-positives remain. We first apply the bipartite matching heuristic described above with $\alpha = 0$ and $\sigma_S = 1$ on $V(\omega)$ yielding a solution $S(\omega)$ that is the set of edges in the optimal matching at each time point. This solution is solely based on the spatial distances d_S between the objects. Given $S(\omega)$, we let it be the set of edges D for computing the means and standard deviations of the spatial and appearance feature differences that then determine normalized scores for d_S and d_A . For a series of different values of $\alpha \geq 0.5$, we then apply the bipartite matching heuristic where the weights of each edge are determined by $D = S(\omega)$ and α . We refer to this solution as $S(\omega, \alpha)$ and observe that it takes object appearance into account as well.

Our key idea is that if $S(\omega) = S(\omega, \alpha)$ or nearly so, and also the larger that α is, then the more likely it is that the edges in $S(\omega) \cap S(\omega, \alpha)$ are all correct. In the ideal case $S(\omega)$ would equal $S(\omega, 1)$ for some ω . So for a set of choices $(\omega, \alpha) \in \Omega \times A$ we compute a table of the number of conflicts between $S(\omega)$ and $S(\omega, \alpha)$ as $K(\omega, \alpha) = |S(\omega) - S(\omega, \alpha)|$ and find the smallest value $K^* = \min\{K(\omega, \alpha)\}$. Almost always $K^* = 0$ and there are multiple choices of ω and α that give the value K^* . Among these we select (ω^*, α^*) where α^* is the largest

α that gives us K^* for some ω , i.e., $\alpha^* = \max\{\alpha \mid \exists \omega : K(\omega, \alpha) = K^*\}$, and ω^* is the lowest possible brightness threshold ω , among those at blending factor α^* that give K^* , i.e., $\omega^* = \min\{\omega \mid V(\omega, \alpha^*) = K^*\}$. To conclude, we let the edges in $\chi = S(\omega^*) \cap S(\omega^*, \alpha^*)$ be our core trajectories and in most cases $\alpha^* \geq 0.8$. In effect we have found the level ω^* that minimizes the possibility of having false positives in $V(\omega^*)$ and hence the solution to the bipartite matching heuristic is most likely to be correct. The so established core trajectories include on average $86\% \pm 11\%$ (s.d.) ($n=40$) of the centrosome objects of the complete trajectories that we wish to establish.

2.2.7 Extending Core Trajectories to Full Trajectories

Given the core trajectory of each centrosome we now go back to the full set of called objects and extend the trajectories forward and backward in time to obtain the full trajectories. To this end, we apply the same core procedure of our tracking algorithm with edge weights that combine spatial distance and appearance distance as described above with the weighting factor $\alpha = 0.3$ and $D = \chi$. The choice of α was empirically chosen and indicates a slight preference towards minimizing distance changes over appearance changes during the extension phase. Moreover, we compute the mean μ_χ and standard deviation σ_χ of the weights of all the edges in the core trajectories and eliminate from $E - \chi$ all edges whose weight is more than 4 standard deviations above the mean, i.e., $\mu_\chi + 4\sigma_\chi$, as being implausible (E is the set of all edges as defined in Section 2.2.5).

In the following we will describe the track extension algorithm for the backward case. Forward extension is analogous. We extend core tracks backward iteratively starting with $t = n$ and working backward to 1. Suppose we have already iteratively extended backward to time t and let $B(t)$ be the set of edges in our current set of tracks where the induction starts with $B(n) = \chi$, the set of core tracks computed in the previous subsection. Note that all edges $(a, b) \in B(t)$ with $b \in V_{t' \leq t}$ are core trajectory edges.

To extend to time $t - 1$ we first find a bipartite matching between V_{t-1} (i.e., the set of all objects at time $t - 1$) and $U(t) = \{b \in V_t \mid \exists a : a \in V_{t-1} \wedge (a, b) \in$

$B(t)\} \cup \{b \in V_t \mid \exists c : c \in V_{t+1} \wedge (b, c) \in B(t)\}$ (i.e., the set all objects at time t that are adjacent to an edge in $B(t)$) using the Hungarian algorithm with the weights and edges described in the previous paragraph. We remove from consideration any edge whose addition to the established trajectories would result in a trajectory with radius hysteresis (Figure 2.6).

Suppose the algorithm returns the set of edges H_t as the best matching between V_{t-1} and $U(t)$. We must reconcile these edges against the core trajectory edges $\chi_t = (V_{t-1} \times V_t) \cap B(t)$ that are currently connecting time $t-1$ and t . An edge $e = (a, b) \in H_t$ is compatible (with $B(t)$) if either $e \in \chi_t$ (i.e., e is also a core edge) or there are no edges adjacent to a and no edges between any vertex at $t-1$ and b in $B(t)$ (i.e., a is not in any core trajectory and b is not already matched to another object at $t-1$). There is only one situation in which we will accept non-compatible edges in H_t as part of the new extension $B(t-1)$. A pair of edges (a, b) and (c, d) in H_t is an *exchange* (w.r.t. $B(t)$) if and only if $(a, d) \in B(t)$ and there are no edges adjacent to c or between any vertex at $t-1$ and b in $B(t)$. This scenario typically arises when a centrosome pair first forms and their spots are so dim that at time $t-1$ c is not above threshold ω^* and so a is inadvertently paired with d instead of b . During the extension phase c is present and the exchange straightens the problem out. Figure 2.7 illustrates an exchange. In conclusion, we add to our current set of edges all the compatible and exchange edges in H_t and remove those (core) edges that are replaced by an exchange edge to arrive at $B(t-1)$, formally:

$$B(t-1) = B(t) \cup \{e \in H_t \mid e \text{ is compatible or an exchange edge}\} \quad (2.4)$$

$$- \{e \in \chi_t \mid e \text{ is replaced by an exchange edge}\}.$$

We apply the same procedure forward in time to obtain a sequence $F(1) = \chi \cap B(1), F(2), \dots, F(n)$ of forward extensions. Note carefully that we start the forward induction from $\chi \cap B(1)$ (i.e., the set of all core trajectory edges that have not been replaced by an exchange edge during the backward extension phase) and not $B(1)$. So it may be that $B(1)$ and $F(n)$ are not mutually compatible exactly as defined above, that is, $B(1) \cup F(n)$ may have tracks with radius hysteresis or tracks that split or merge. First, we resolve every such fork in the merged result by eliminating the edge, which depends on whether it is in $B(1)$ or $F(n)$, with the higher weight. When such an elimination is required

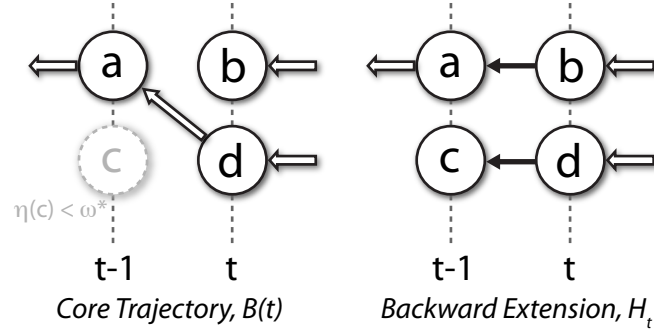


Figure 2.7: Illustration of an exchange during the backward extension phase (the edges are drawn in the opposite direction only to illustrate that we work backward in time). The core trajectory edge $(a, d) \in B(t)$ is replaced by the exchange edges $(a, b) \in H_t$ and $(c, d) \in H_t$.

we want to be sure that the extension procedure whose edge is eliminated would not have made a different extension in light of this. So we rerun the forward and backward extensions with the incompatible edges removed. We repeatedly do this until $B(1)$ and $F(n)$ are compatible, typically one or two iterations. We then further remove from $B(1)$ and $F(n)$ those extension edges with highest weight that would link (non-overlapping) core trajectories in the merged result in order to prevent introducing edge sequences with radius hysteresis. The result of the extension phase is the set of tracks $T = B(1) \cup F(n)$.

2.2.8 Trajectory Stitching

Occasionally, two or more non-overlapping core trajectories are found for a given centrosome, in which case, after the extension step, the trajectory is broken into several parts that abut in time. In a final phase we search to exhaustion for edges that link two trajectories fulfilling the three conditions below and add the edge, effectively 'stitching' together two trajectories.

1. The stitched trajectory must not be longer than a generous maximum bound of 45 minutes on the duration of the centrosome cycle.
2. The differences in position, radius, and integrated intensity across the stitching edge must not be greater than the respective maximum difference over all trajectories established so far.

3. The stitched trajectory must not have radius hysteresis ($r_{\text{init}} = 700$ nm, $r_{\text{cont}} = 600$ nm).

We then repeat the stitching process with less strict conditions for (a) trajectories that are shorter than a minimum of 10 minutes on the duration of the centrosome cycle, and (b) trajectories with initial radius greater than 600 nm. For these rare cases of certainly broken trajectories we allow twice the respective maximum difference (condition 2) and relax the radius hysteresis to $r_{\text{init}} = 1000$ nm and $r_{\text{cont}} = 550$ nm (condition 3).

2.2.9 Tracking Performance

Our corpus of centrosome movies comprises several hundred embryos between the 1- and the 16-cell stage with various centrosome markers and genetic conditions. To quantify the tracking performance we randomly selected 10 movies of multi-cell embryos for detailed visual inspection. Due to the small number of movies that cover the 8- and 16-cell stage in our corpus, we additionally inspected 3 random examples of such movies. Many papers on tracking systems including [16] identify three types of errors: failure to incorporate an object, incorporation of a false-positive object, and incorporation of an incorrect object in a given track. We use a finer error categorization because some kinds of errors are more severe than others. For example, a completely missed centrosome is a more severe tracking error than a trajectory that misses only a couple of frames at the beginning of the cycle when the centrosome is a tiny dim spot. We use the following error categories that cover the space of all observed errors.

- *Short track.* The track is correct but misses a few frames at either end.
- *Long track.* The track is correct but has a few extra frames at either end.
- *Broken track.* A centrosome is completely tracked but the trajectory is broken into two or more pieces. Stitching the involved tracks would correct the problem.

Table 1: Minor and major tracking error rates in per cent of the total number of tracks. The tracking results of randomly selected centrosome movies starting at different cell stages were analyzed.

Cell stage	n	Minor errors		Major errors				
		Short	Long	Broken	Fused	Chimera	Missing	Particle
1	8	0	0	0	0	0	0	0
2	32	0	3.1	0	0	0	0	0
4	48	10.4	2.1	0	0	0	0	0
8	48	10.4	2.1	0	0	0	0	0
16	64	17.2	3.1	1.6	0	1.6	3.1	0

- *Fused track*. Two or more centrosomes are in a track but splitting the track would correct the problem.
- *Chimera track*. The track contains more than one centrosome or non-centrosome object and it is not a long or fused track.
- *Missing track*. A centrosome is not tracked at all.
- *Particle track*. A non-centrosome object is tracked but involves no centrosomes.

Table 1 summarizes the evaluation results. Tracks that are a few frames too short are the most prevalent error, but it is also the least severe because it only involves faint, tiny centrosomes at the very beginning of the cycle or diffuse, disassembled centrosomes at the very end of the cycle. Fusion errors and particle tracks never occurred. More severe errors occurred only in the 16-cell stage. For example, two missed centrosomes were due to their being so deep in the stack, and hence so dim, that they were not picked up in the core trajectory extraction step.

2.3 MEASURING CENTROSOME SIZE

2.3.1 Definition of Centrosome Size and Measuring Algorithm

The total fluorescence is often used as a measure for the size of spot-like objects, which in principle is proportional to the amount of the fluorescently labeled protein. Given a Gaussian fit of the centrosome's 2D intensity profile (see Section 2.2.4), the total fluorescence above background can accurately be obtained as the full integral under the Gaussian. Unfortunately, as a measure for centrosome size it is only useful for a limited number of cases because it is severely affected by the centrosome's z-position. Figure 2.8 shows an extreme example. While the two marked centrosomes are identical in terms of protein amount, they appear in the image with a 10-fold intensity difference. The reason for this is that the light that comes from the centrosome closer to the bottom of the embryo has to travel a longer distance through the densely packed embryo before it reaches the microscope. As a result much more light is lost on the way and so the object appears significantly dimmer although the total protein amounts are the same. It is extremely difficult to computationally correct for this artifact because the optical properties of the embryo constantly change due to additional cell membranes when the cells divide as well as polarization events within the cells.

As an alternative measure for centrosome size that is much more robust against this imaging artifact we use the radius of the centrosome. Under the reasonable assumption that the centrosome is spherical (this is for example less the case for centrosomes in *Drosophila*) the radius can then readily be converted to volume in order to obtain a 3-dimensional estimate of centrosome size. In Section 2.2.4 we compute an approximation for the radius $\rho_G(p)$ of object p as twice the geometrical mean¹ of the standard deviations $\sigma_p(p)$ and $\sigma_m(p)$ along the principal and minor axes of the fitted Gaussian, i.e. $\rho_G(p) = 2\sqrt{\sigma_p(p) \cdot \sigma_m(p)}$. At the resolution of light microscopy, however, there is no clear boundary between a centrosome and the background. There-

¹ The geometrical mean of the two semi-axes of an ellipse is the radius of a circle with the same area as the ellipse.

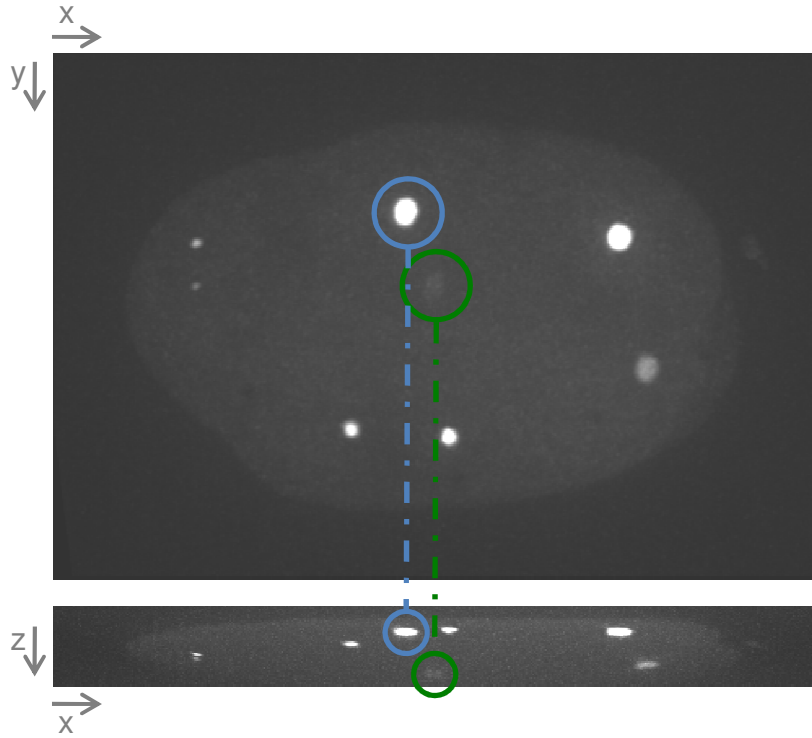


Figure 2.8: The intensity of the centrosomes strongly depends on the distance to the embryo surface. Centrosomes deeper down in the stack (green) appear dimmer than those close to the top of the embryo (blue). If the embryo was imaged flipped over the green centrosome would appear brighter than the blue one.

fore, the radius of the fitted Gaussian is actually only a measure that is relative to an unknown but fixed scaling constant. We empirically chose this constant to be 2 based on human perception of centrosome size. Note, that this constant cancels out when comparing the size of centrosomes relative to each other. Furthermore, while the model above is sufficient for the purposes of computationally inferring centrosome tracks, for the biological analysis of the data we undertake the following two refinements to obtain as accurate a measurement of centrosome size as is possible.

In the late stages of a centrosome's trajectory the intensity profile of the spot is often not Gaussian, typically having a dip in the center as illustrated in Figure 2.9(a). To more accurately measure the radius of the object in these cases, we assume to first approximation that the illumination of the spot is symmetric with respect to the sub-pixel center and elliptical iso-contours of the fitted Gaussian, and compute the radial intensity profile of the spot as a function of ellipse radius. That is for each ellipse centered about $(x(p), y(p))$

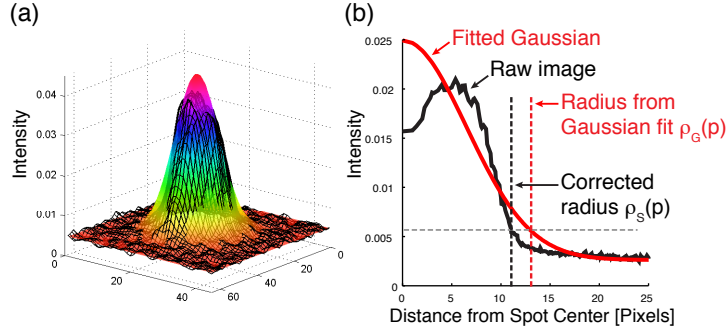


Figure 2.9: Refining centrosome radius. (a) A non-Gaussian centrosome intensity profile together with its Gaussian fit. A dip at the center of the profile results in a slight over-estimation of the object's radius. (b) Final centrosome radius is measured in a radial intensity profile based on the Gaussian fit.

with skew $\sigma_p(p)/\sigma_m(p)$ and angle $\theta(p)$, we compute the mean intensity along this ellipse. This is illustrated in Figure 2.9(b) along with the radial intensity profile of the fitted Gaussian. The spot radius $\rho_S(p)$ is then estimated as that value for which the radial intensity of the raw image equals the radial intensity of the fitted Gaussian at $\rho_G(p)$.

Finally, we correct for image distortion caused by the convolution of each light point with the point spread function of the microscope, which makes objects to appear bigger in the images than they actually are. An approximation is given by $\rho(p) = \sqrt{\rho_S(p)^2 - \rho_{PSF}^2}$ (convolution of a Gaussian object with a Gaussian PSF) where ρ_{PSF} is the width of the point spread function and $\rho(p)$ is our final estimate of the radius of p .

2.3.2 Validation of the Measuring Accuracy

To estimate the accuracy of our radius measuring algorithm we tested it against fluorescent beads of known size. Since the embryo has significantly different optical properties than a clear solution such as water or oil, we injected beads into embryos.² This is a rather delicate procedure and thus we did not select embryos in a specific cell-stage. The biggest beads that we could inject had a radius of 550 nm which is about half the size of centrosomes in the 4-cell embryo. A bead has a sharp physical boundary and so its intensity profile only

² This work was done in a collaboration with Horatio Fantana.

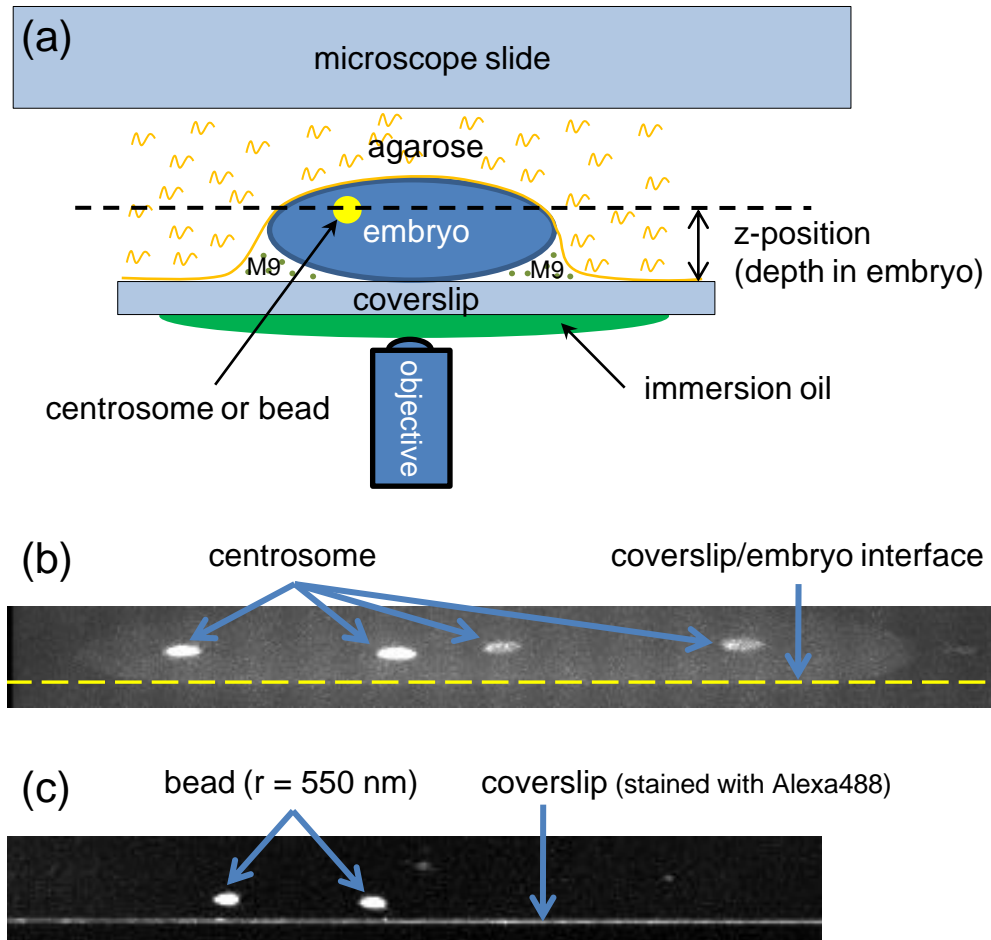


Figure 2.10: Imaging beads inside *C. elegans* embryos. (a) The setup for imaging beads is identical to imaging centrosomes. (b) Side view of an embryo with γ -Tubulin labeled centrosomes. (c) Side view of an embryo with injected beads of 550 nm radius.

appears Gaussian-like in the image because of the convolution with the point spread function. It is unknown whether the same is true for centrosomes, or whether additionally a non-uniform distribution of the labeled protein within the PCM determines its intensity profile. Nonetheless, a bead injected into an embryo should be accurate enough an approximation to analyze how the measured radius depends on the z-position of the object.

Figure 2.10(a) shows the imaging setup. It is identical for centrosomes and beads. To determine the z-position of an object relative to the coverslip / embryo interface, we need to recognize where this interface is. For embryos with fluorescently labeled centrosomes this can be done directly with an accuracy of one or two z-planes because the cytoplasm fluoresces as well, as shown in Figure 2.10(b). For beads that we injected into N2 worms (i.e., no fluorescent

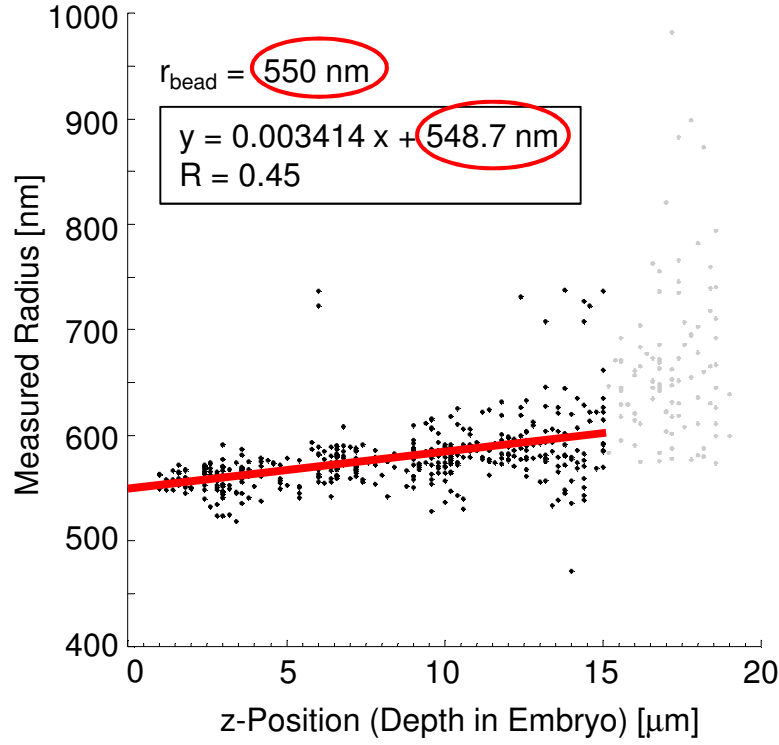


Figure 2.11: The measured radius of injected beads increases slightly with increasing z-position. The line is a least-square fit of all data points with z-position not greater than 15 μm which is the maximum relevant z-position of centrosomes up to the four-cell stage in our actual experiments.

labeling) and imaged with only about 10 to 20 per cent of the laser power used for imaging centrosomes, the cytoplasm is not visible. We therefore instead labeled the coverslip by adding Alexa Fluor 488 to the M9 buffer which non-specifically binds to the coverslip, see Figure 2.10(c).

We automatically identified all beads in a total of 46 embryos, measured the radius of each bead with the algorithm described above and computed the z-position relative to the coverslip / embryo interface. Beads that were very close to each other were discarded manually. Figure 2.11 shows that most of the beads were measured within 50 nm of the actual radius of 550 nm and that there is a slight tendency to over-estimate the radius with increasing z-position. We suspect that the reason for this is the eggshell around the embryo that could have the effect of an optical lens. The line fit extrapolated to z-position = 0 shows that the measured radius without the lens effect would be 548.7 nm, corresponding to a measurement error of less than one per cent.

We did not use this result to correct for the slight radius over-estimation for two reasons. First, it is unclear how this over-estimation depends on the size of the beads. Unfortunately, it was not possible to inject beads bigger than 550 nm radius. Second, for fluorescently labeled centrosomes, the correction would require the location of the embryo / coverslip interface in each stack. This is not always possible because the imaged z-range of the embryo often does not contain the embryo / coverslip interface (due to minimizing the number of z-planes) and for some centrosome markers the contrast between embryo and background is too low.

2.4 BIOLOGICAL IDENTITIES OF THE TRACKED CENTROSOMES

It is often necessary to know the biological identities of the tracked centrosomes, that is, the identity in the *C. elegans* lineage [101, 102] of the cell in which it is located and whether it is the anterior or the posterior associated pole of the spindle. We determine these identities by establishing a binary tree of trajectories representing the pair and parent/child relationships between the tracked centrosomes as follows.

1. For centrosome movies that start at the 1-cell stage the anterior/posterior orientation of the embryo is easily determined as the sperm always enters on the posterior side of the embryo [103]. Thus, the pole closest to the centrosomes detected in the first frame of the movie is the posterior pole. For movies that start at later cell stages, we additionally take into account the stereotypical timing of the cell divisions to figure out where the posterior pole is. For example, in the 4-cell stage, the cells ABa and ABp divide before EMS, and EMS divides before P₂.
2. The centrosomes are paired by finding an optimal bipartite matching between their trajectories. There is an edge between trajectory k and h if and only if (a) they overlap in time by at least 6 minutes, (b) the distance between the centrosomes is positively correlated with time (i.e., they tend to separate), and (c) their minimum centrosome separation distance is less than 12 μm . The weight $w(k, h)$ between k and h is the

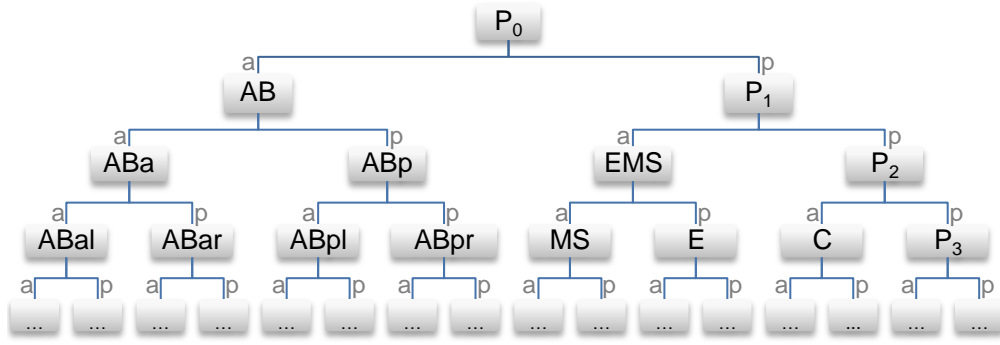


Figure 2.12: *C. elegans*' invariant cell lineage.

product of (a) the minimum centrosome separation distance in the first half of the trajectories and (b) the absolute difference in radii averaged over time. Intuitively, we expect a pair to be near each other and for their centrosomes to be at the same growth stage. For the pairs so established, we label the one closer to the posterior pole at the end of its trajectory as the posterior child and the other as the anterior child.

3. In the last step, we determine for each centrosome pair its (common) parent centrosome, again by computing a minimum-weight bipartite matching on trajectories where the edge weight is the distance between the last location of a potential parent and the first location of a potential child. Edges are only present between trajectories where the parent ends no later than 3 minutes after the child's begin, and the child begins no later than 15 minutes after the parent's end.

Given the invariant cell lineage of *C. elegans* (Figure 2.12) and its canonical naming scheme with respect to anterior/posterior localization, we then only need to traverse the resulting tree level by level to assign a cell name to each centrosome.

2.5 DETECTION OF THE NUCLEAR ENVELOPE BREAKDOWN

In order to compare the time series of centrosome statistics with each other we need to register them in time. It is common practice in biology to define such a registered time axis relative to a cell cycle event. We use nuclear envelope

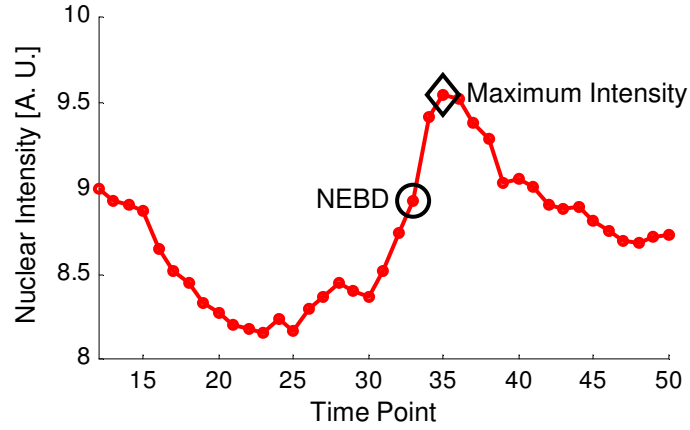


Figure 2.13: Typical time vs. nuclear intensity curve computed by the nuclear envelope breakdown detection algorithm.

break down (NEBD) [104] because it can be observed in our images without an additional marker. The nucleus is free of centrosomal proteins and so appears as a dark, circular region in the cytoplasm. As the nucleus breaks down, there is an inflow of GFP-labeled molecules from the cytoplasm into the nuclear region and its fluorescent intensity increases. The NEBD is the time point for which the rate of this inflow of luminosity is maximal.

As the dark nuclear region is difficult to delineate based purely on signal, we take advantage of prior knowledge. The nucleus is located roughly halfway between the two centrosomes and in the early stages its diameter is in the range of 5 to 8 μm . We thus have an estimate of an image region in which the darkest pixels correspond to the nucleus and can compute a time versus nuclear intensity curve.

Given a pair of centrosomes we consider a circle half way between them in the xy-dimension and consider the minimum intensity projection of the z-planes between them within that region. We estimate the intensity of the nucleus as the average of the darkest half of the pixels in this circle. Plotting this value through time yields a curve illustrated in Figure 2.13. We first find the time of the maximum intensity of the curve and then set as the NEBD the time of the maximum rate of change of the intensity in the preceding three minutes.

2.5.1 *Accuracy of NEBD detection*

To evaluate the performance of our automatic NEBD detection method we manually determined the NEBD for 30 randomly selected movies containing a total of 95 cells. During manual NEBD detection, we experienced that it was often ambiguous as to the exact frame in which NEBD occurred. Therefore, we considered automatic NEBD detection to be correct if it was within 1 frame of the manually determined NEBD. We measured exact agreement for 64 per cent and agreement within 1 frame for 34 per cent of the cases, yielding 98 per cent accuracy for NEBD detection.

3D-SEGMENTATION OF PH-DOMAIN::GFP IMAGES

To determine how centrosome size correlates with cell size, we marked cell membranes with GFP fused to a PH domain [105] and took z-stacks of embryos between the one and the four-cell stage. To quantify the volume of each cell type, the resulting 3D images needed to be segmented. For image acquisition we used two different techniques, spinning disk confocal and single plane illumination microscopy (SPIM) [106, 107], which differ in the time and effort needed for imaging as well as in the quality of the resulting images.¹ This chapter describes the two segmentation pipelines we have developed for PH-domain::GFP images. Both are based on marker-controlled watershed [98]. This extension of the watershed algorithm handles the problem of over-segmentation by defining one marker for each object to be segmented, i.e., a set of connected pixels within the respective object. The input image is then modified so that it has local minima only at each marker position. The segmentation obtained by the watershed transform of this image depends on the exact choice of the markers but has exactly as many regions as markers that were specified. In our case there is one marker for each cell and one marker for the background.

3.1 SPINNING DISK CONFOCAL MICROSCOPY IMAGES

For segmentation of PH-Domain::GFP [105] images acquired with a spinning disk confocal microscope, we have developed a completely automatic algorithm based on plane-by-plane segmentation using 2D marker-controlled wa-

¹ SPIM imaging was done in collaboration with Stephan Preibisch.

tershed. Figure 3.1 shows several planes of a typical z-stack together with the segmentation result. The following problems complicate the segmentation.

1. The membrane signal almost completely attenuates towards the bottom of the embryo.
2. The resolution is highly anisotropic.
3. While in most parts of the images, the membranes have sharp, peak-like profiles, some membrane parts in the lower half of the stack, especially those at the outer embryo contour, are very blurred and their profiles are step-like.
4. Small bright objects may be present inside the cells.

One important thing that simplifies the segmentation task is that the cells do not lie on top of each other up to the 4-cell stage. The segmentation is performed in 2D plane-by-plane and based on two main ideas.

First, the signal changes relatively little between two subsequent z-planes. We exploit this by propagating the segmentation of one z-plane as marker to the next. The segmentation begins with the image I_z half way through the embryo (plane #8 in Figure 3.1). For I_z we compute the 2D watershed segmentation W_z (i.e., a binary image) from an externally provided marker M_z . For all other planes we automatically derive a marker from the respective adjacent plane that is already segmented. That is, for the plane I_{z+1} , we compute a marker M_{z+1} by slightly shrinking the binary regions in W_z using morphological erosion. For I_{z+1} we then compute the watershed segmentation W_{z+1} from M_{z+1} . This process continues until the top of the stack is reached or no cell marker remains as a result of the erosion. The same procedure is applied towards the bottom of the stack. We refer to this procedure as marker-controlled watershed with marker propagation.

The second idea is to perform the segmentation in a hierarchical manner. We first only separate the embryo from the background, i.e., we find the outer contour of the embryo. Here, the marker has only two regions, one for the embryo that extends over all cells and one for the background. In the second step, we impose the detected embryo contour on the input image and then segment

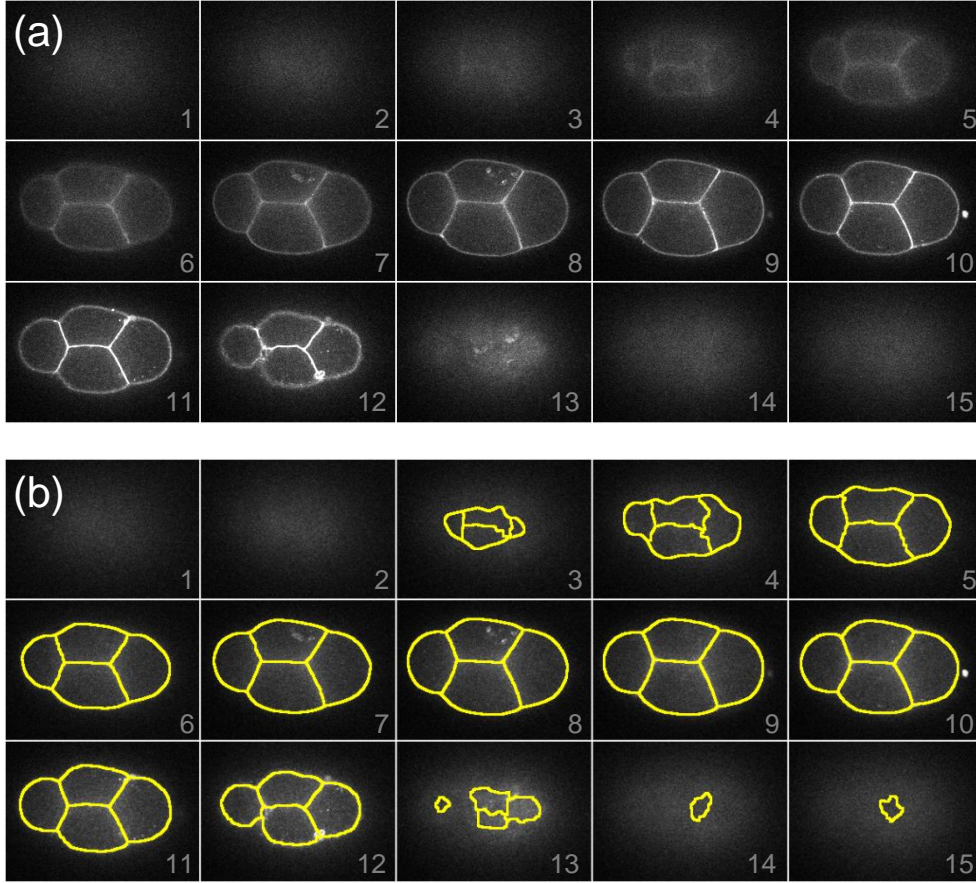


Figure 3.1: Spinning disk confocal stack of an embryo with PH-domain::GFP labeled cell membranes. Montage showing every fourth plane of the stack. (a) Original images. (b) Cell contours detected with our hierarchical plane-wise watershed segmentation algorithm.

the individual cells in this modified image. The rationale behind the two-step segmentation is that the intensity profiles of the membranes at the outer embryo contour are different from those in the interior of the embryo. Both steps use marker-controlled watershed with marker propagation as described in the previous paragraph.

The complete workflow of the algorithm is illustrated in Figure 3.2. As pre-processing steps, we apply background subtraction and noise reduction. Background subtraction is based on fitting a two-dimensional third-order polynomial through background sample points. To reduce noise while largely preserving edge information we perform anisotropic diffusion [108]. Additionally, for the detection of the outer embryo contour a Sobel filter amplifies edge information. The algorithm then proceeds with the actual hierarchical

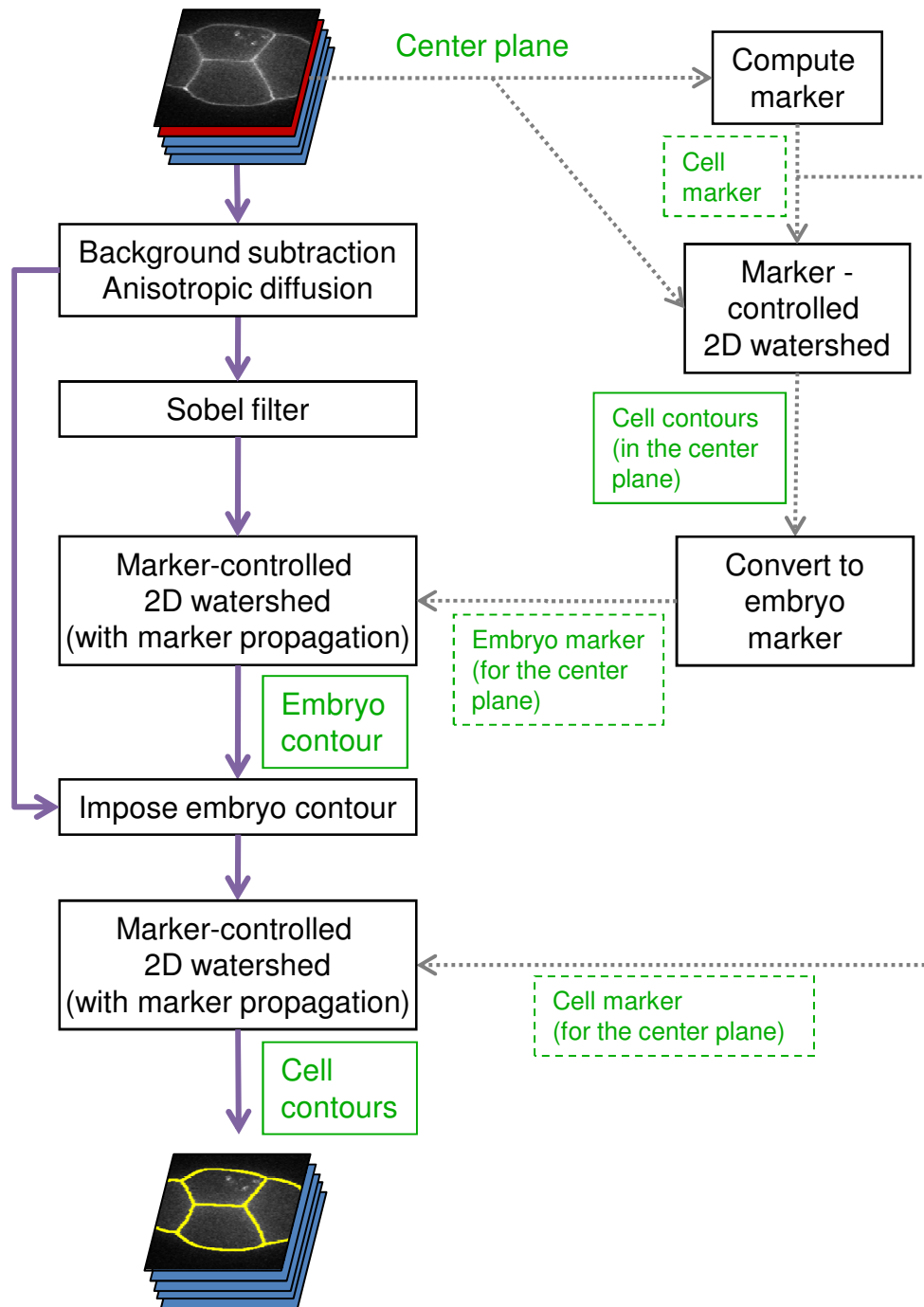


Figure 3.2: Workflow of the segmentation algorithm for PH-domain::GFP images acquired with spinning disk confocal microscope.

watershed segmentation as described above. The right part of the workflow computes a marker for each cell in the center plane to automatically initialize the watershed cascade. This will be described in the next subsection.

3.1.1 *Automatic initialization*

Automatic initialization of the segmentation pipeline requires finding the center plane, i.e., the plane half way through the embryo, and computing a marker for each cell in this image. To determine the center plane we perform a rough segmentation of the embryo region using Otsu thresholding [97] and find the center plane as the plane in which the embryo area is maximal. In this image the membrane signal is usually significantly above background but small weak parts may occur. To amplify the membrane signal in these regions we compute the maximum-intensity projection of the sub-stack that contains the center plane and a few planes above and below it. On this projection we then apply the seven filtering and thresholding steps illustrated in Figure 3.3 to compute a marker for each cell.

3.2 SINGLE PLANE ILLUMINATION MICROSCOPY IMAGES

Thanks to the isotropic resolution of the PH-domain::GFP images acquired with single plane illumination microscopy (SPIM), we can directly perform the watershed segmentation in 3D; marker propagation and hierarchical segmentation are not necessary. The higher image quality comes at the cost of a more time-consuming imaging procedure in which it is much more difficult to find early embryos. Figure 3.4 shows several planes of a typical SPIM stack together with the segmentation result. One difficulty with these images is the high degree of blur caused by small morphological changes of the cells while multiple views are acquired. The registration and fusion process that combines these views into one isotropic image cannot fully correct for these changes. Figure 3.5 illustrates the workflow of the segmentation algorithm. It consists of a series of pre-processing steps to amplify the membrane signal and then applies marker-controlled 3D watershed. The marker is supplied manu-

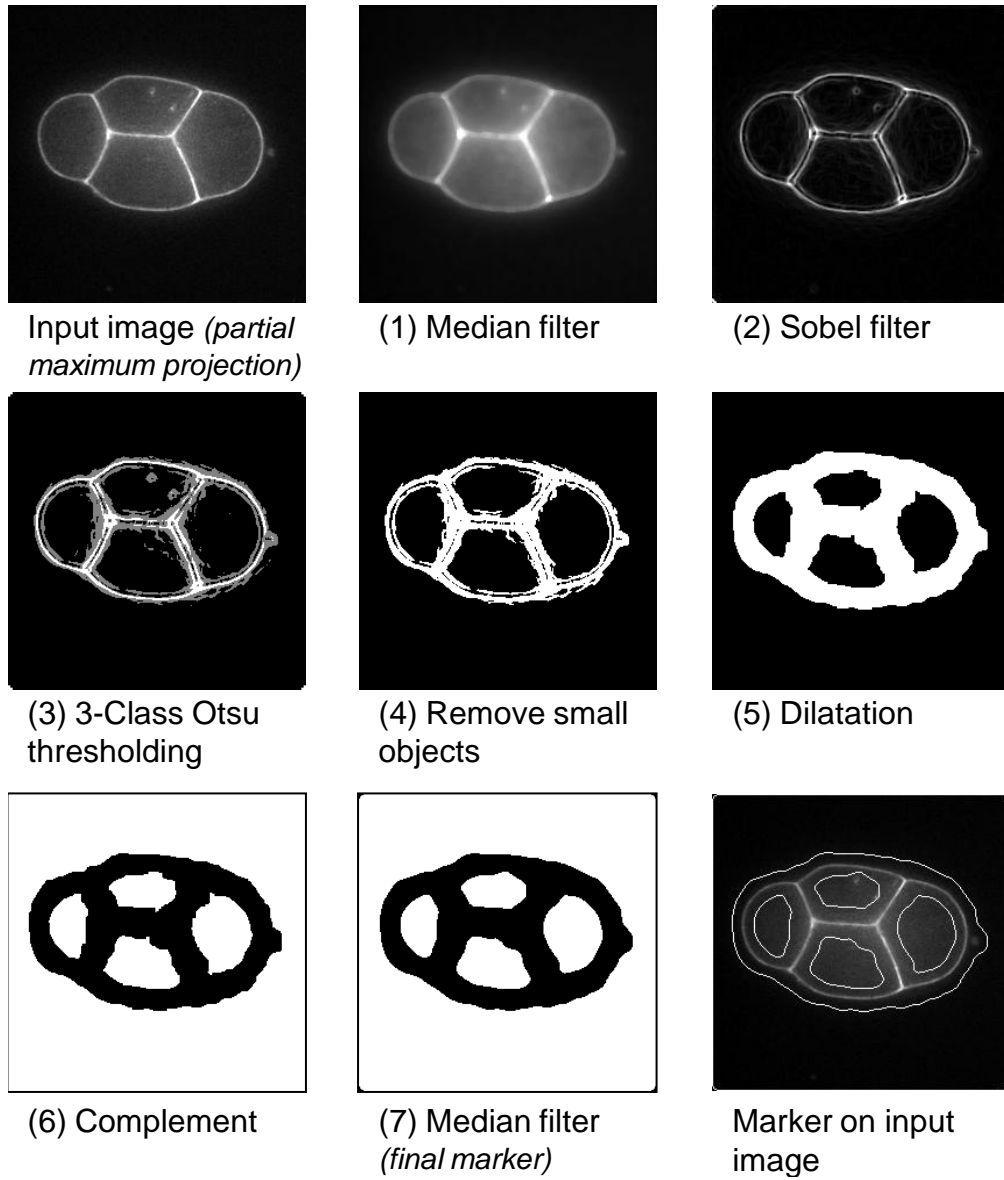


Figure 3.3: Computation of an initial watershed marker for each cell in the center plane.

ally, which requires one click for each cell. The SPIM image segmentation pipeline is also applicable for embryos after the 4-cell stage when the cells lie on top of each other. Figure 3.6 shows volume renderings of the segmentation result for a 4-cell and an 8-cell embryo, respectively.

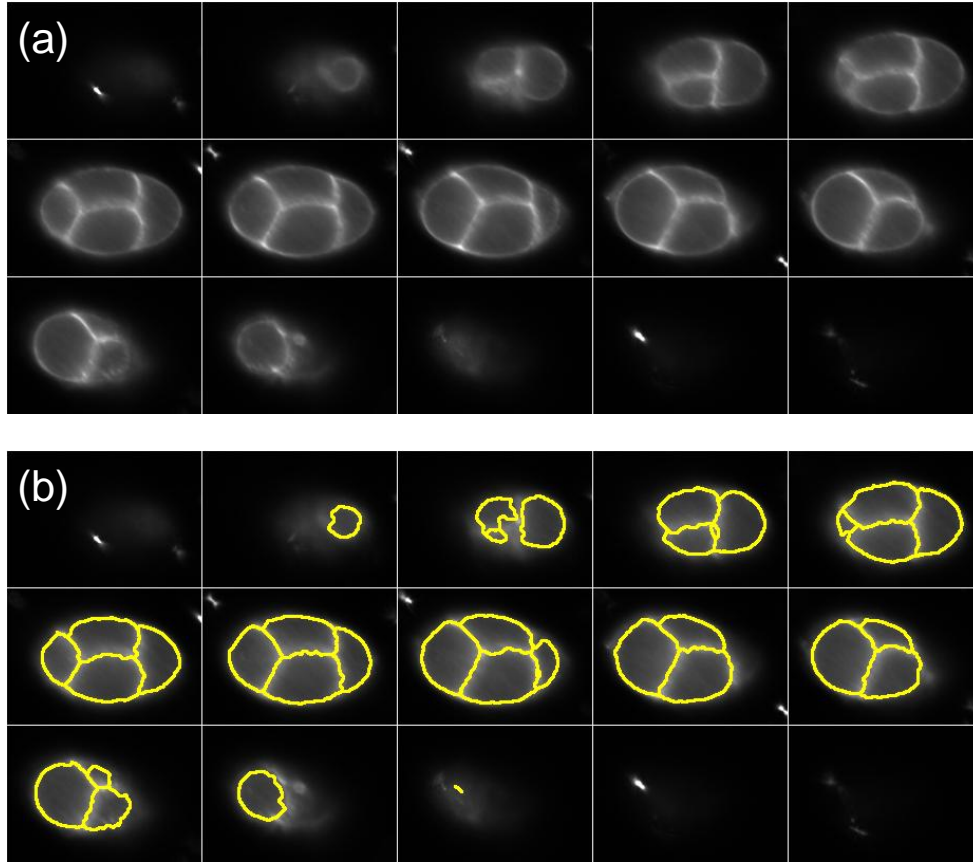


Figure 3.4: Registered and fused single plane illumination stack of a 4-cell embryo with PH-domain::GFP labeled cell membranes. Montage showing every eighth plane of the stack. (a) Original images. (b) Cell contours detected with our algorithm based on marker-controlled 3D watershed segmentation.

3.3 CELL VOLUMES IN *C. ELEGANS* EMBRYOS

While both segmentation pipelines, for spinning disk and SPIM images, produce minor segmentation errors the overall segmentation accuracy is well-suited for the biological analysis. Imperfect segmentation mainly occurred at the lower and upper end of the stacks and for the case of SPIM images when

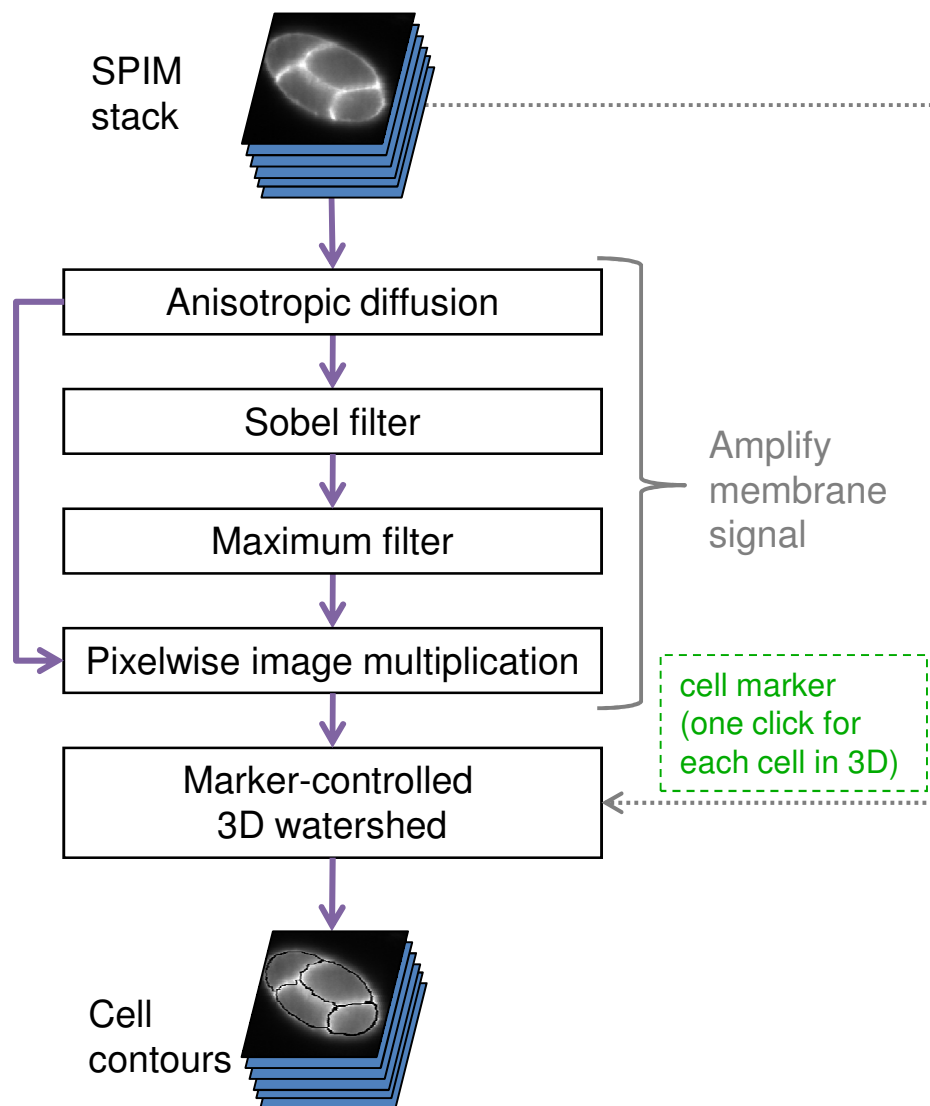


Figure 3.5: Workflow of the segmentation algorithm for PH-domain::GFP images acquired with single plane illumination microscope.

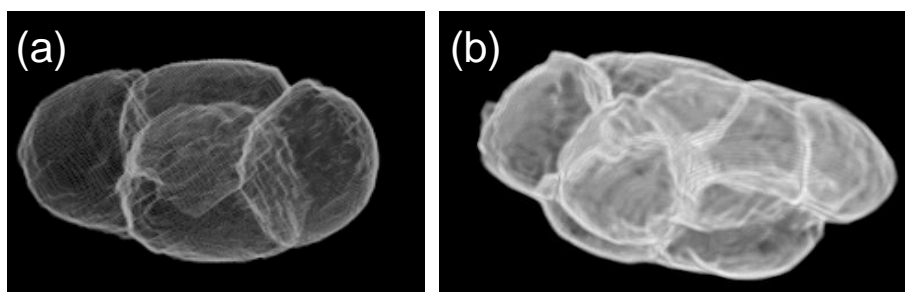


Figure 3.6: 3D view of the segmentation results for PH-domain::GFP images acquired with SPIM. (a) 4-cell embryo. (b) 8-cell embryo.

a bead (required for registration of the different views) was directly adjacent to the embryo.

To quantify the volume of each segmented cell, we model the volume enclosed between two adjacent z-planes as a prismatoid. Thus, the total volume of a cell is

$$V = \frac{1}{2} \sum_i \Delta z \cdot (A(z_i) + A(z_{i+1})) \cdot \Delta^2_{xy} \quad (3.1)$$

where $A(z_i)$ is the area in pixels of the cell in plane z_i , Δz is the distance between to adjacent z-planes and Δxy the pixel size in x and y. Figure 3.7 shows a graph of cell volumes for each cell type in the one- to the four-cell stage.

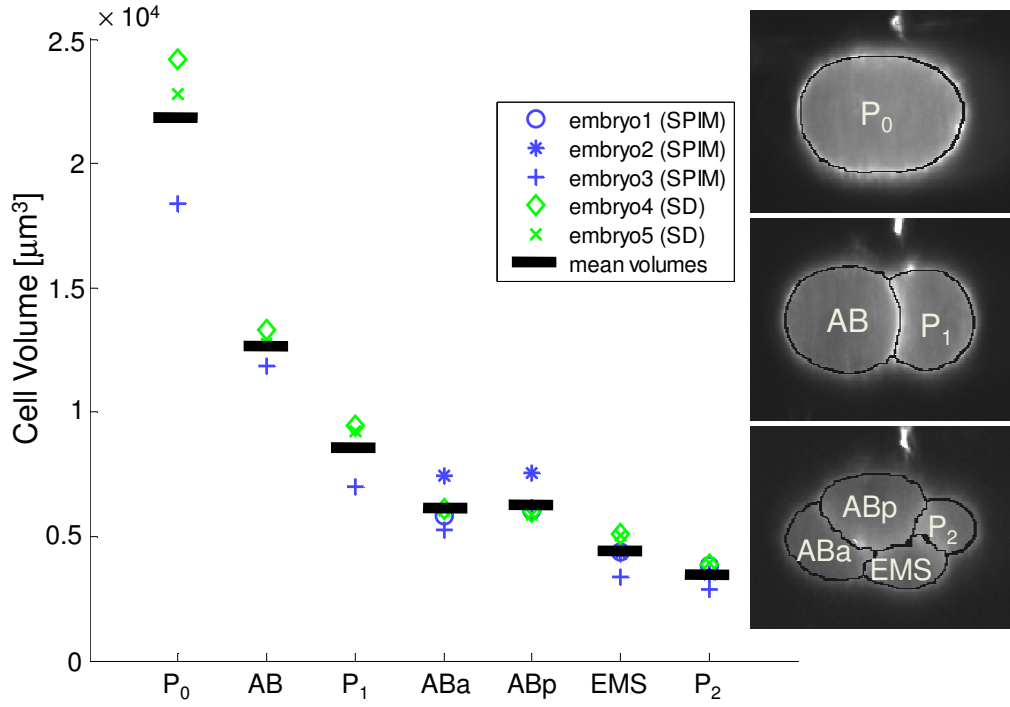


Figure 3.7: Graph showing the computed volumes for all cell types between the 1- and 4-cell stage. Three embryos using SPIM and two embryos using spinning disk confocal microscope (SD) were filmed at 1-minute intervals. Embryo 1 and 2 were filmed only in the 4-cell stage; embryo 3, 4 and 5 from the 1- to 4-cell stage. For each individual embryo multiple samplings (i.e., 3D stacks) of the same cell stage were segmented with the respective segmentation algorithm described in Section 3.1 and 3.2, respectively, and the computed cell volumes averaged. The images shown in the right are selected planes of SPIM stacks (embryo3) together with the identified cell boundaries.

WHAT SETS THE SIZE OF A CENTROSOME?

The ways in which cells set the size of their intracellular structures is an important but largely unsolved problem [70]. To approach this problem we set out to investigate the mechanisms by which the centrosome scales in size during early embryonic divisions in *C. elegans*. The assay that we described in Chapter 2 allows us to obtain quantitative, time-resolved data of centrosome growth. In this chapter we will first examine how centrosomes grow in wild-type *C. elegans* embryos over multiple rounds of cell division. We will then use genetic perturbation techniques to alter parameters such as cell size, centrosome number per cell or the amounts of centrosome components to study how these parameters affect centrosome size. Supported by our data, we propose a limiting component hypothesis in which the available amount of one or more structural centrosome components dictates centrosome size.

4.1 CENTROSOME DISINTEGRATION BEGINS AT THE METAPHASE-ANAPHASE TRANSITION

We first looked at centrosome growth in one-cell embryos labeled with γ -TUB::GFP. The P₀ wild-type growth curve in Figure 4.1(a) shows that after an initiation phase that starts at about 650 seconds before nuclear envelope breakdown (NEBD) the centrosome continuously increases in size through the cell cycle, before it eventually disintegrates by loss of PCM material. Such a kinetic profile complicates the quantification of centrosome size, because it does not suggest a clearly defined time point at which the centrosome reaches its final size. The disintegration of the centrosome in *C. elegans* is in part due

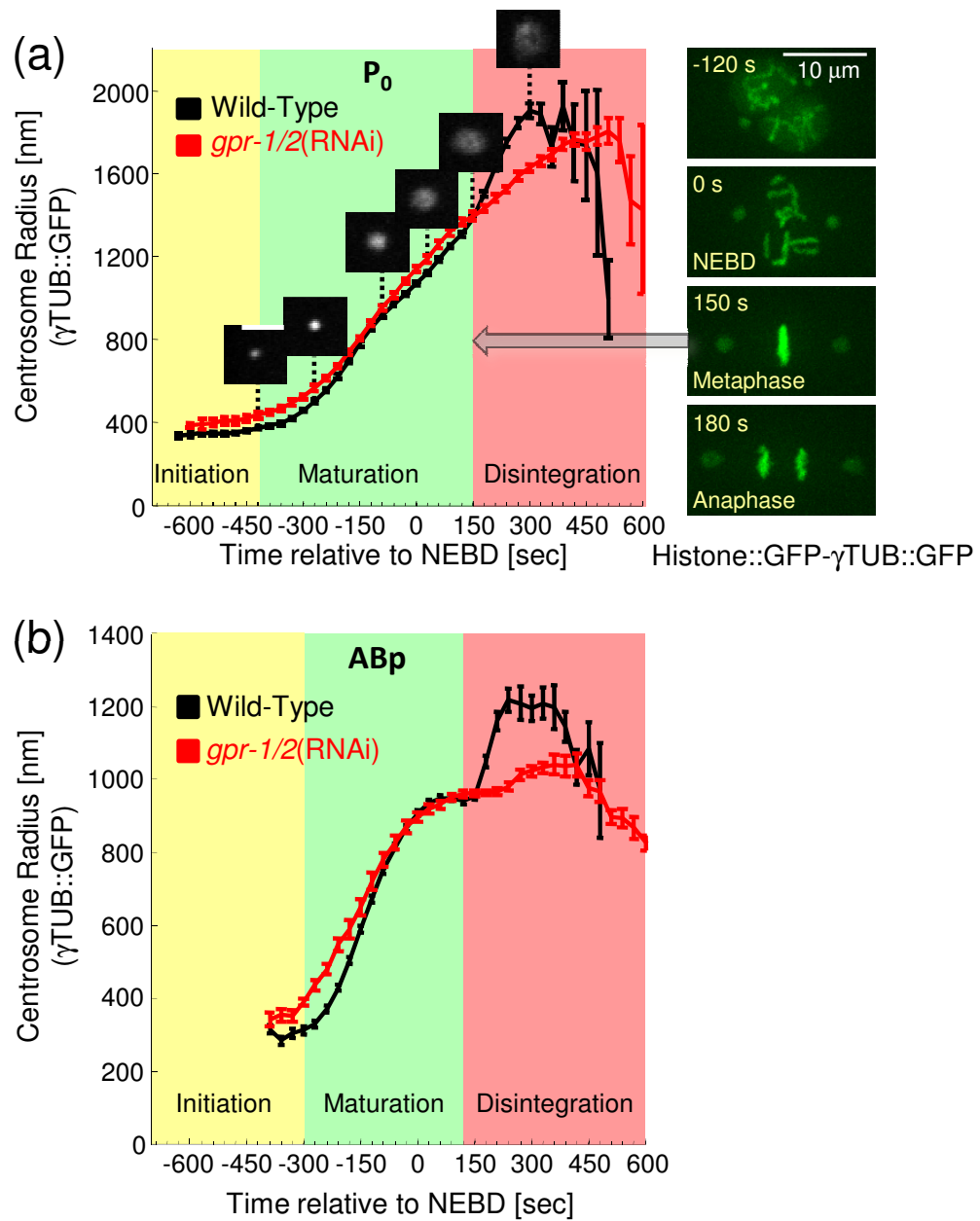


Figure 4.1: Centrosome growth in wild type and in the absence of cortical force generation. (a) Centrosome radius over time in P_0 from an average of 20 wild type and 7 *gpr-1/2*(RNAi) embryos. Note that the radius in the initiation phase is slightly above the resolution limit. Error bars are standard error of the mean. Scale bar in centrosome clipping is 3 μ m. Image sequence are maximum projections of an embryo expressing γ -Tubulin::GFP and Histone::GFP, illustrating that metaphase correlates with the transition from maturation to disintegration. (b) Centrosome radius over time in ABp from an average of 20 wild type and 7 *gpr-1/2*(RNAi) embryos.

cortical pulling forces mediated by microtubules that tear the PCM apart. To identify the onset of disintegration we looked at centrosome growth in the absence of cortical pulling forces using *gpr-1/2*(RNAi) [109]. This showed that centrosome growth in P_0 (Figure 4.1(a)) and ABp (Figure 4.1(b)) are almost identical with and without force generation until about 150 seconds after nuclear envelope breakdown for P_0 and 120 seconds for ABp. Thus, any increase in size after this point might be a consequence of pulling forces and not an active recruitment of new PCM components. We then manually scored the time point of metaphase relative to nuclear envelope breakdown in wild-type Histone::GFP- γ -Tubulin::GFP embryos. Because the metaphase time point coincided with the divergence point of the wild-type and *gpr-1/2*(RNAi) growth curves for the respective cell type, we defined final centrosome size as the size at metaphase. Microtubule-mediated pulling forces are activated at the metaphase-anaphase transition to segregate the chromosomes. One possibility is that the same mechanism is also involved in disintegration of the centrosomes.

4.2 THE GROWTH KINETICS OF CENTROSOMES IN WILD-TYPE EMBRYOS CHANGE THROUGH DEVELOPMENT

We next tracked and measured the centrosomes through development up to the 16-cell stage in γ -TUB::GFP embryos. Figure 4.2 shows the growth curves averaged over the centrosomes within each cell stage. As expected, final centrosome size continuously decreases from one cell stage to the next. Moreover, our time-resolved data reveals two important features of how centrosomes grow. First, as cells become smaller the centrosomes grow slower. Second, starting at the four-cell stage, centrosome size plateaus before onset of centrosome disintegration.

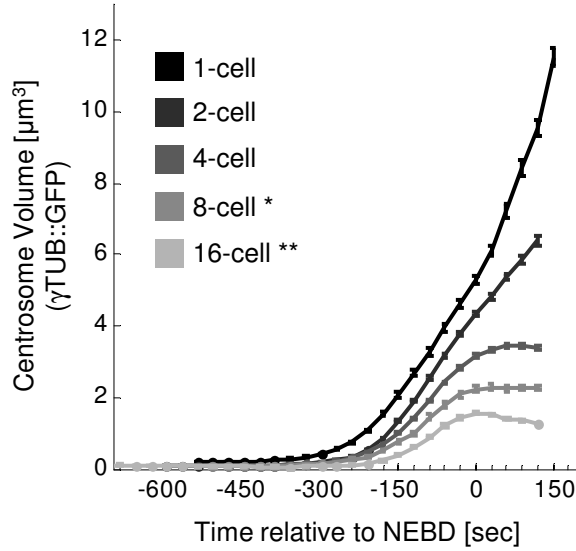


Figure 4.2: Centrosome growth through development. Growth rate and final centrosome volume decrease from the 1- to the 16-cell stage. The data is comprised of an average of 25 embryos imaged from the 1- to 4-cell stage, and 6 embryos imaged from the 8- to 16-cell stage, all expressing γ -Tubulin::GFP. Error bars are standard error of the mean. (*) Due to signal attenuation along the z-axis, not all 16 centrosomes of an embryo in the 8-cell stage contribute to the graph. (**) Similarly, not all 32 centrosomes of an embryo in the 16-cell stage contribute to the graph; only AB-Lineage centrosomes were analyzed.

4.3 CENTROSOMES BECOME SMALLER IN EACH ROUND OF CELL DIVISION

To see whether the decrease in centrosome size continues after the 16-cell stage, we monitored centrosomes in a dividing embryo until all 558 embryonic cells are present. Proceeding from the one to the 558-cell stage takes about 7 hours. Unfortunately, imaging the embryo over such a long period was only possible by increasing the imaging interval from 40 seconds to 5 minutes as otherwise the laser exposure was lethal for the embryo (see also Section 2.1). At this low temporal resolution it was not possible to reliably track all centrosomes and determine the nuclear envelope breakdown in each cell. However, we could still use our centrosome detection algorithm (see Section 2.2.4) to identify and measure the centrosomes in each stack individually. To exclude noise objects and inaccurate radius measurements of extremely

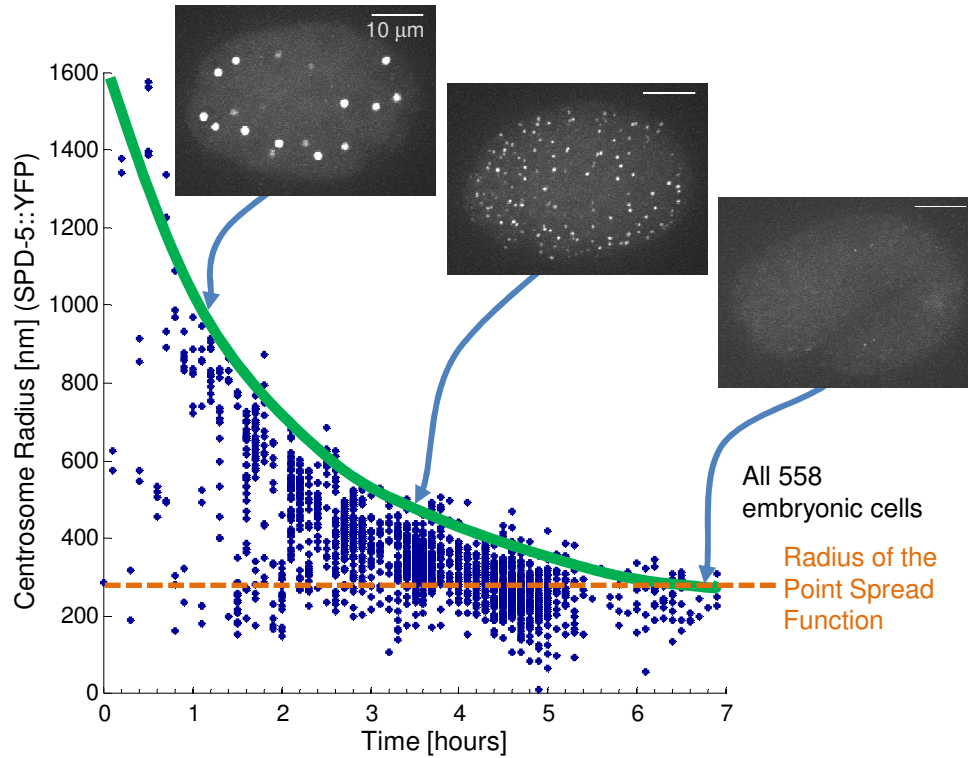


Figure 4.3: Centrosomes become smaller throughout embryonic development. An embryo with SPD-5::YFP labeled centrosomes was imaged from the 1- to the 558-cell stage at 5-minute intervals. The dashed orange line indicates the mean radius measured for sub-resolution beads injected into embryos. The solid green line is a manually drawn trend line of the maximum centrosome radius over time taking into account that some of the identified centrosomes are at the disintegration phase. Snapshots of embryos are maximum projections.

dim centrosomes we discarded all detected objects with peak intensity less than 1.5 times above background intensity. The remaining objects are all centrosomes that are either growing, plateauing or being pulled apart (disintegration). Although only an approximation and keeping in mind the resolution limit of the microscope, this result suggest that centrosome size decreases continuously throughout entire embryonic development (Figure 4.3).

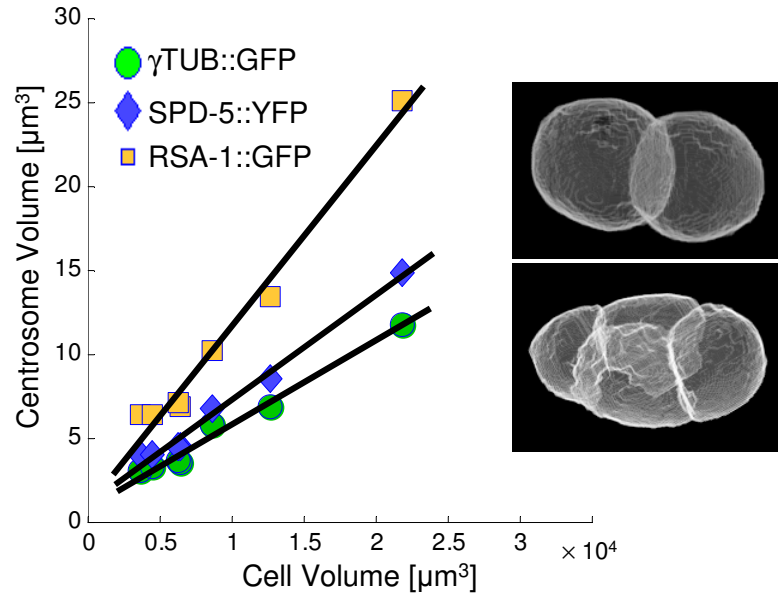


Figure 4.4: Centrosome volume versus respective cell volume for three different centrosome markers. Linear fits are least square fits. Final centrosome size was measured at metaphase time depending on cell type (e.g., 150 s after NEBD for P_0). Volume measurements for individual cell types were obtained as described in Chapter 3 (n=5). Inset images showing reconstructions of an embryo in the 2- and 4-cell stage.

4.4 CENTROSOME VOLUME SCALES PROPORTIONATELY WITH CELL VOLUME

To see in more detail how the decrease in cell size and centrosome size through development correlate with each other we marked cell membranes with GFP fused to a PH domain [105] and quantified the cell volumes using the segmentation algorithms described in Chapter 3. For three different centrosome markers, γ -TUB::GFP [110, 62, 111], RSA-1::GFP [112] and SPD-5::YFP [42] we computed final centrosome size (i.e., size at metaphase) in each cell. Plotting cell volume against centrosome volume (Figure 4.4) showed that the relationship between cell volume and centrosome volume is approximately proportionate.

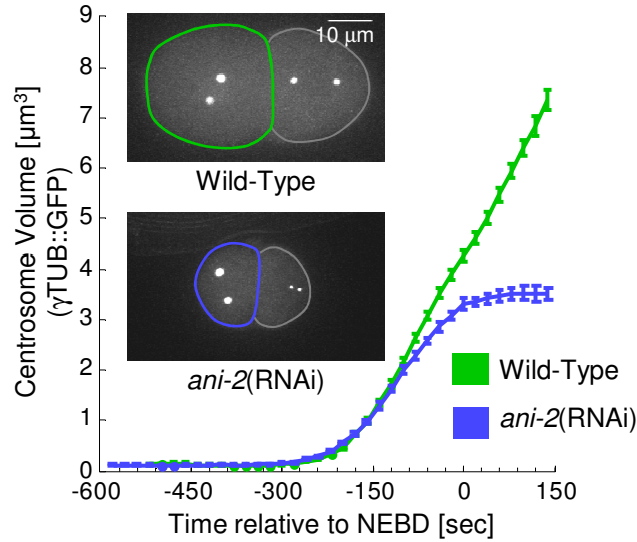


Figure 4.5: Centrosome size plateaus in AB cells of *ani-2*(RNAi) embryos ($n = 8$) that are $42\% \pm 5\%$ (s.d.) of wild-type ($n = 20$) size. Assuming ellipsoidal shape, embryo volume was estimated by measuring the embryo area in maximum intensity z-projections and the embryo thickness in maximum intensity y-projections. Snapshot images are maximum projections.

4.5 CENTROSOME SIZE IS INDEPENDENT OF LINEAGE

One possibility is that centrosome growth is controlled in a developmental manner, i.e., it depends on the cell type. The alternative is that centrosome growth is independent of cell type, but is instead an (indirect) consequence of cell size. To distinguish these possibilities, we varied the size of embryos using *ani-2*(RNAi) [113]. *ani-2* encodes a *C. elegans* anilin, which is required for precise determination of oocyte size. Smaller embryos make smaller cells, allowing us to compare the size of centrosomes in a varied range of cell sizes for any particular cell type. In AB cells of embryos that are $42\% \pm 5\%$ (s.d.) of wild type size as shown in Figure 4.5, centrosomes initially grow slower than centrosomes in wild-type cells. After this initial growth, the size then reaches a plateau, which is never seen in centrosomes of AB cells of wild type size. All other cells (P_0 , P_1 , ABa, ABp, EMS and P_2) showed a similar variation in growth rates upon depletion of ANI-2 (data not shown). Interestingly, the changes in centrosomal growth rate during development, and changes in growth rate in smaller *ani-2*(RNAi) cells are similar. We measured

the maximum growth rates (see Materials and Methods, Chapter 7) of centrosomes through development, and after *ani-2*(RNAi). This yielded a maximum growth rate of centrosomes in two-cell wild-type embryos of $0.022 \mu\text{m}^3/\text{sec}$, while centrosomes in four-cell embryos grow at $0.017 \mu\text{m}^3/\text{sec}$. Centrosomes in two-cell *ani-2*(RNAi) embryos that are $42\% \pm 5\%$ (s.d.) of wild-type size grow at $0.018 \mu\text{m}^3/\text{sec}$. These experiments show that centrosome size is independent of cell type. The plateau in growth that occurs in later divisions is thus a consequence of cell size and not of a developmental shift.

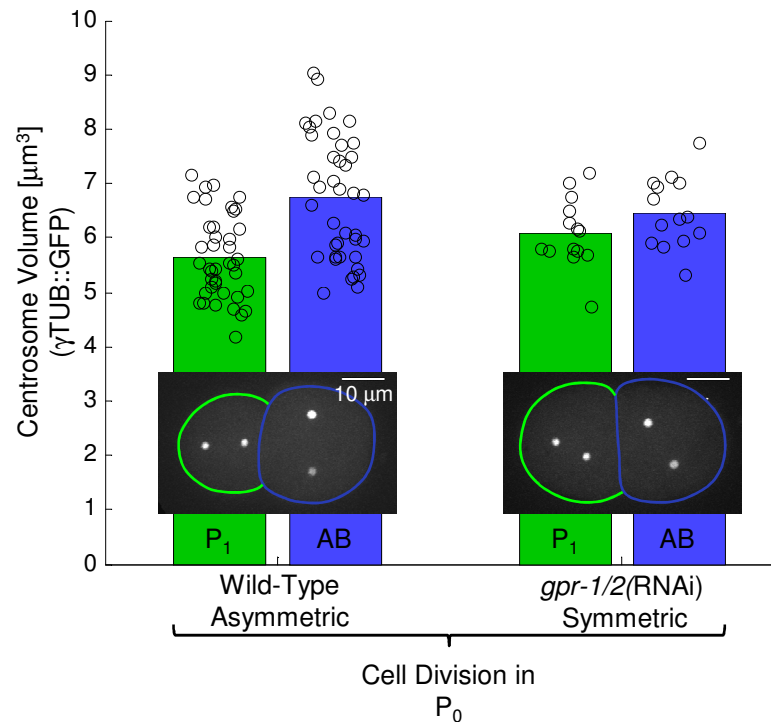


Figure 4.6: Centrosomes in AB and P₁ are approximately equal in size when P₀ divides symmetrically. Bar chart showing centrosome volume at metaphase in AB and P₁ for 20 wild-type and 7 *gpr-1/2*(RNAi) embryos. Two-sided t-test: Null hypothesis: P₁ and AB have equal means and equal but unknown variance. Alternative: the means are not equal. Wild-type: p-value = $3.116 \times 10^{-6} \Rightarrow$ highly significant difference between P₁ and AB. *gpr-1/2*(RNAi): p-value = 0.129 \Rightarrow no significant difference between P₁ and AB.

How sensitively centrosome size responds to changes in cell size can be seen by looking at AB and P₁ cells that are equal in size. In wild-type embryos, P₀ divides *asymmetrically* into a bigger AB cell and a smaller P₁ cell. Accordingly, the centrosomes in AB are bigger than in P₁ (Figure 4.6). *Symmetric* division of

P_0 occurs upon *gpr-1/2*(RNAi) [109]. *gpr-1* and *gpr-2* are two extremely similar genes that encode proteins in *C. elegans* required for the posterior displacement of the spindle, the position of which defines the division axis. In two-cell *gpr-1/2*(RNAi) embryos, not only are the AB and P_1 cells approximately equal in size but their centrosomes are as well (Figure 4.6).

4.6 TOTAL CENTROSOME VOLUME IS CONSTANT ACROSS CELL STAGES

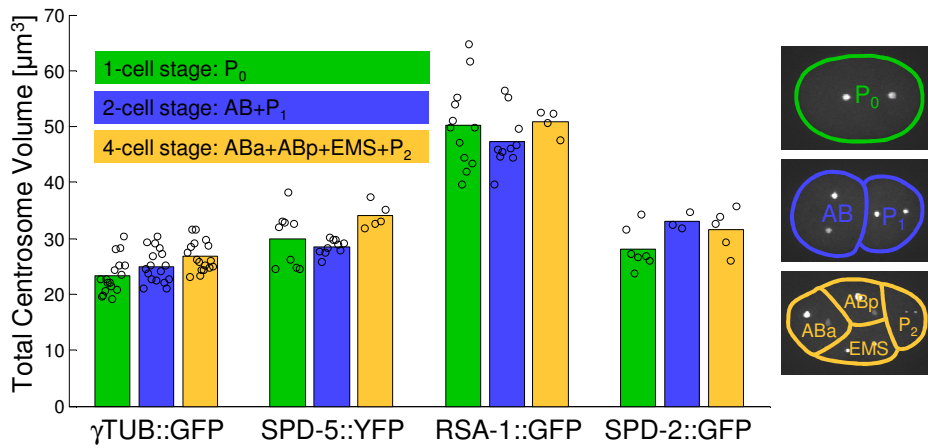


Figure 4.7: Total centrosome volume is conserved through development. Comparison of the total sum of centrosome volumes between the 1-, 2- and 4-cell stages. Centrosome volume was determined in each individual cell at metaphase. $\gamma\text{-TUB}$: $n = 17/17/17$, SPD-5: $n = 9/10/5$, RSA-1: $n = 12/10/4$, SPD-2: $n = 7/3/5$ for 1-/2-/4-cell embryos.

The data presented in the previous sections, in particular the fact that centrosomes tend to plateau when the cells are small, suggests that centrosome size is limited by the availability of one or more of its components. Simplistically speaking, like a brick house that cannot be built any bigger once all available bricks are used up, a centrosome cannot further increase in size when no more of its structural components are available. Assuming a uniform distribution of these components in the PCM, this model implies that the *total* volume of all centrosomes in any one cell stage should be constant as the cytoplasm is partitioned into many cells during development. To test this idea, we quantified total centrosome volume up to the four-cell stage (Figure 4.7). These measurements show that for the four centrosomal components we mea-

sured, total centrosome volume stays constant in the embryo. In other words, the volumes of the individual centrosomes approximately half in each round of cell division. The graph additionally shows that the measured centrosome size depends on the marker with RSA-1 taking up significantly more volume than the other three markers. Beyond the four-cell stage, the attenuation of the fluorescent signal along the z-direction does not allow reliable measurement of all centrosomes. We could therefore neither prove nor disprove whether the conservation of centrosome volume is also true for higher cell stages. However, the following theoretical consideration leaves the possibility that this could indeed be the case up to the 558-cell stage when all embryonic cells are present. A necessary condition for this is that the centrosomes must then still be big enough to accommodate the centriole pair in the PCM. Assuming a radius of 1500 nm for a centrosome in the one-cell stage and ignoring cell death events during development, each cell divides on average about nine times. This means that after halving centrosome volume nine times, its radius would be about 190 nm. This is indeed still big enough to accommodate the centriole pair.

4.7 TOTAL CENTROSOME VOLUME IS INDEPENDENT OF CENTROSOME NUMBER

A second consequence of the limiting component idea is that the total centrosome volume should also stay constant if one increases or decreases the number of centrosomes per cell. The individual centrosomes should hence become accordingly smaller or bigger, respectively. We were able to manipulate the number of centrosomes using two mutant alleles of ZYG-1. ZYG-1 is a kinase that determines the number of centrioles per cell, and thereby the number of centrosomes [43]. Using the temperature-sensitive mutant allele *zyg-1(b1)* [43], we could decrease the number of centrosomes to one in one-cell embryos (Figure 4.8(d)). This resulted in centrosomes of twice the volume as in wild-type embryos with two centrosomes in P₀ (Figure 4.8(a)).

Conversely, with the mutant allele *zyg-1(it29)* (unpublished line by Kevin F. O'Connell) we could randomly increase the number of centrosomes in the two-

cell stage (Figure 4.8(d)). As expected, the more centrosomes there are in a cell, the smaller their final size and the slower the centrosomes grow (Figure 4.8(b)). Again, all centrosomes together add up to the same volume at metaphase as the centrosomes in wild-type embryos, independent of centrosome number (Figure 4.8(c)).

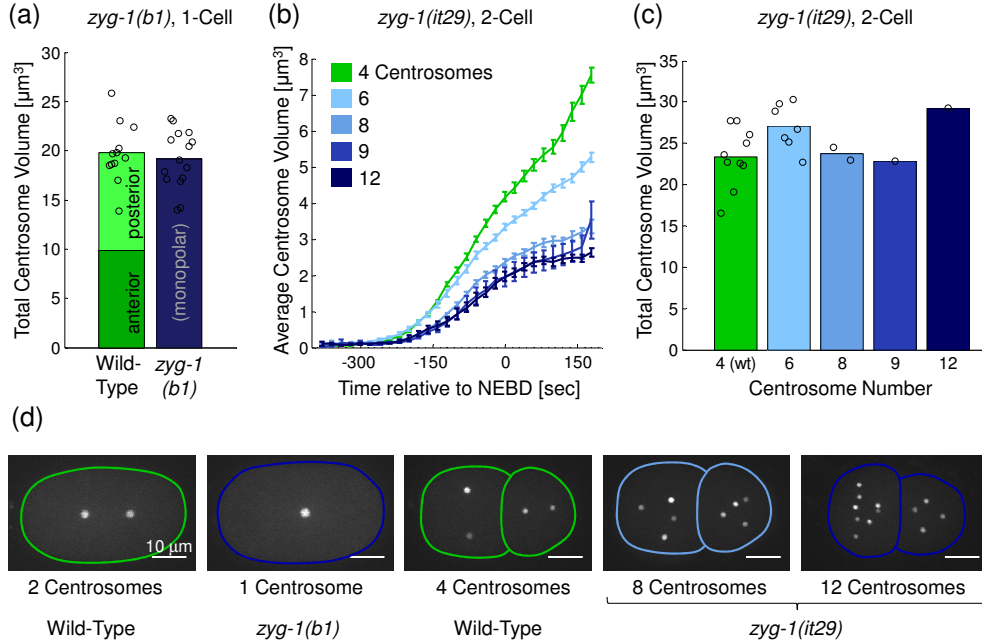


Figure 4.8: Total centrosome volume is independent of the number of centrosomes per cell. (a) Total centrosome volume at metaphase in 1-cell wild-type ($n = 12$) versus *zyg-1(b1)* embryos ($n = 15$). (b) Centrosome growth in 2-cell wild type versus *zyg-1(it29)* embryos. Each curve is an average of all centrosomes in the 2-cell stage. Wild-type: 10 embryos, *zyg-1(it29)*: 7/2/1/1 embryos with 6/8/9/12 centrosomes in the 2-cell stage. Error bars are standard error of the mean. (c) Total centrosome volume at metaphase in the same embryos as in (b). (d) Representative images of the respective embryos are maximum projections.

Finally, we also examined centrosome size in cells in which we varied the relative size of the two centrosomes by depleting the centriolar protein SAS-4 [44]. SAS-4 is a centriole duplication factor [44] and thus becomes crucial after the one-cell stage when two new centrioles are required for normal development in *C. elegans*. Depending on the duration of *sas-4*(RNAi) this results either in a partial phenotype with two centrosomes of different size or in a full phenotype with only one centrosome per cell (Figure 4.9). In the case of unequal centrosomes the smaller centrosome harbors a defective daughter

centriole while the other centrosome inherits the intact mother centriole [44]. Again the total centrosome volume is approximately the same as in wild-type embryos with two equally big centrosomes (Figure 4.9(a)) per cell.

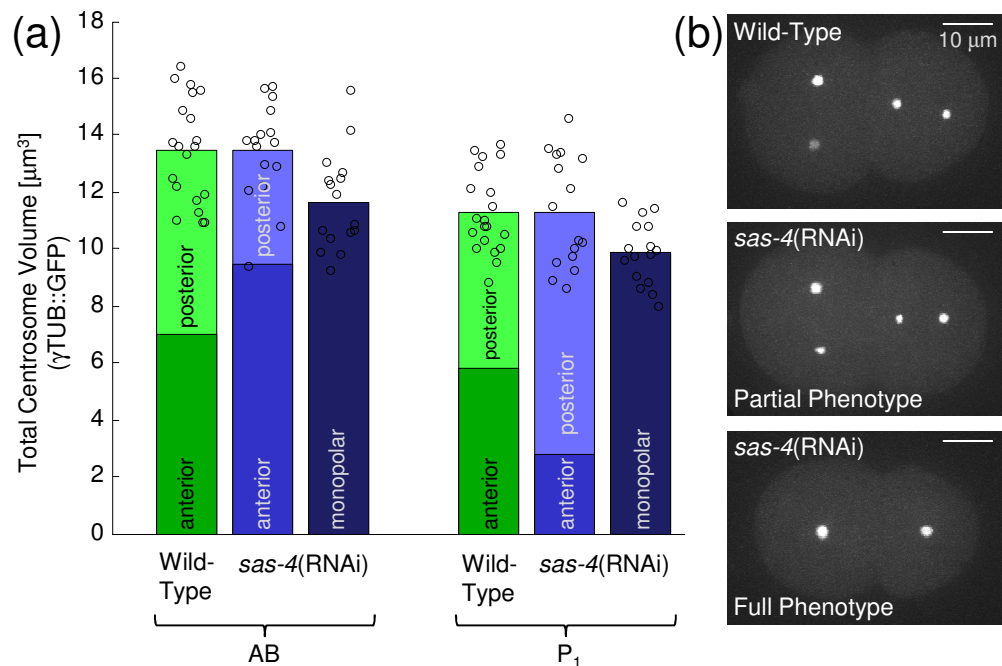


Figure 4.9: Total centrosome volume is independent of centrosome number per cell. (a) Total centrosome volume in AB and P₁ at metaphase in *sas-4*(RNAi) embryos with partial (bipolar spindle but unequally sized centrosomes) and full phenotype (monopolar spindle). Wild type: 20 embryos, *sas-4*(RNAi) partial phenotype: 16 embryos, *sas-4*(RNAi) full phenotype: 16 embryos. (b) Representative images of the respective embryos are maximum projections.

Taken together these three experiments clearly show that the total centrosome volume remains constant independent of the number of centrosomes per cell.

4.8 SEVERAL COMPONENTS ARE REQUIRED FOR A FULL SIZE CENTROSOME

The results of the previous sections strongly support our idea of a limiting component. We thus tried to identify which proteins could set centrosome size by altering the amounts of several proteins in γ-TUB::GFP embryos. We were particularly interested in the proteins SPD-2 [60, 40], SPD-5 [42] and AIR-

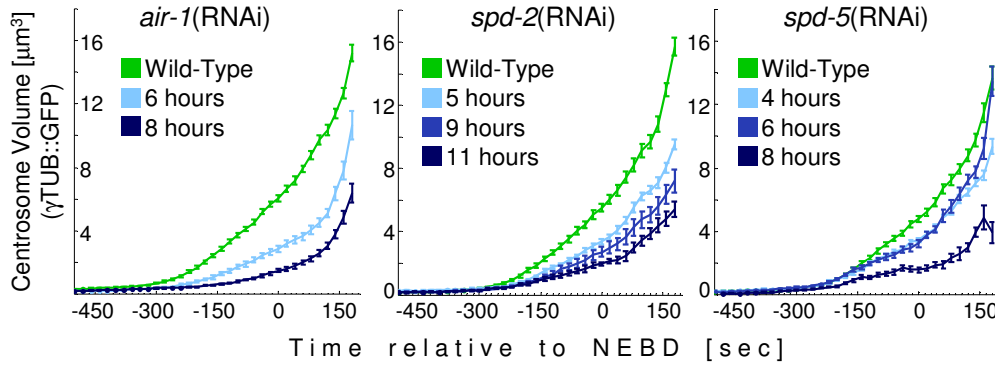


Figure 4.10: Three potential candidates may act limiting on centrosome size. Graph showing centrosome growth in 1-cell embryos expressing γ -Tubulin::GFP and subjected to RNAi against the respective endogenous gene. *air-1*: $n = 18/16/10$ embryos, *spd-2*: $n = 12/5/5/3$, *spd-5*: $6/4/4/6$ (order of the given numbers corresponds to the order in the legend of the respective graph). Error bars are standard error of the mean.

1 [41] as these have previously been shown to be required for the localization of all other proteins to the centrosome. Indeed, partial knock down of these three components by RNAi decreases the rate of centrosome growth in similar ways (Figure 4.10).

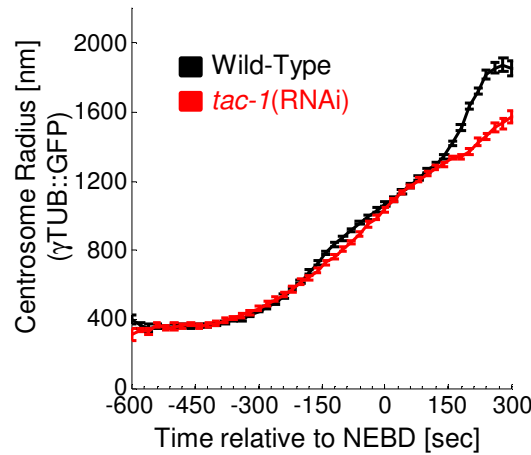


Figure 4.11: Centrosome growth in P_0 in the absence of long microtubules (*tac-1*(RNAi), 6 embryos) is identical to wild-type (24 embryos) up to centrosome disintegration at 150 sec after NEBD. Error bars are standard error of the mean.

We also tested TAC-1, a centrosomal protein that is required for long astral and long spindle microtubules [114], and found that centrosome size in P_0 is not affected by loss of TAC-1 (Figure 4.11) up to the onset of centrosome

disintegration at about 150 seconds after nuclear envelope breakdown. Thus, centrosome size is independent of microtubule length.

To see whether transcription of a potential limiting component plays a role for centrosome size we inhibited transcription using *ama-1*(RNAi). *ama-1* [115, 116, 117] encodes the large subunit of RNA polymerase II which catalyses the transcription of DNA to mRNA. In one- to four-cell *ama-1*(RNAi) embryos, the centrosome growth curves are identical to wild-type (Figure 4.12). This suggests that centrosome growth, at least up to the four-cell stage, entirely relies on maternally provided mRNA amounts. Unfortunately, we could not investigate whether translation of a potential limiting component plays a role for centrosome size without affecting cell cycle progression.

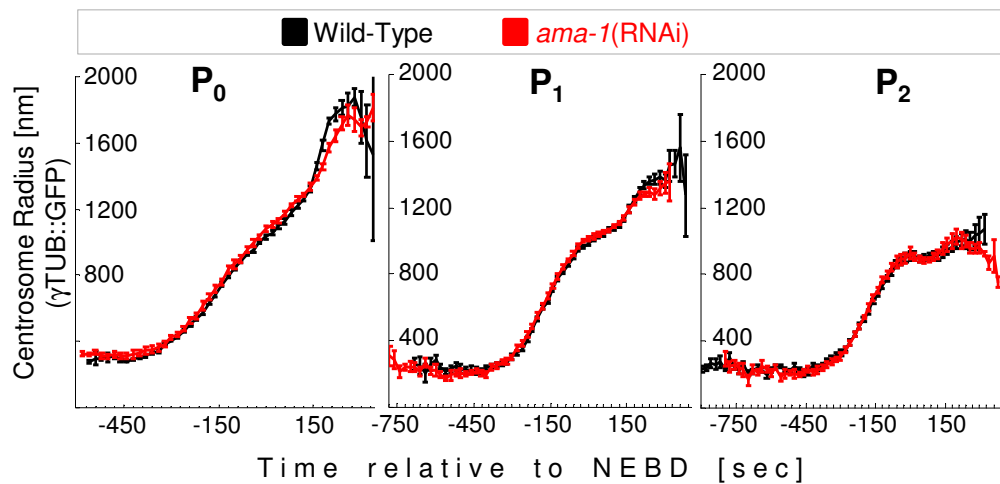


Figure 4.12: Inhibiting transcription (*ama-1*(RNAi), 4/5/4 embryos for P₀/P₁/P₂) does not affect centrosome growth in the one- to four-cell stage (wild-type: 5/9/8 embryos for P₀/P₁/P₂). Error bars are standard error of the mean.

4.9 OVER-EXPRESSION OF SPD-2 LEADS TO BIGGER CENTROSOMES

If the amount of a protein directly determines centrosome size, then one would expect bigger centrosomes upon raising the amount of the protein. One technique to increase protein levels is to introduce a GFP transgene into the genome. The total amount of the protein is then the sum of the expression of the endogenous and the additionally introduced exogenous gene. There are two possible effects on centrosome size. (1) Either centrosome size remains the

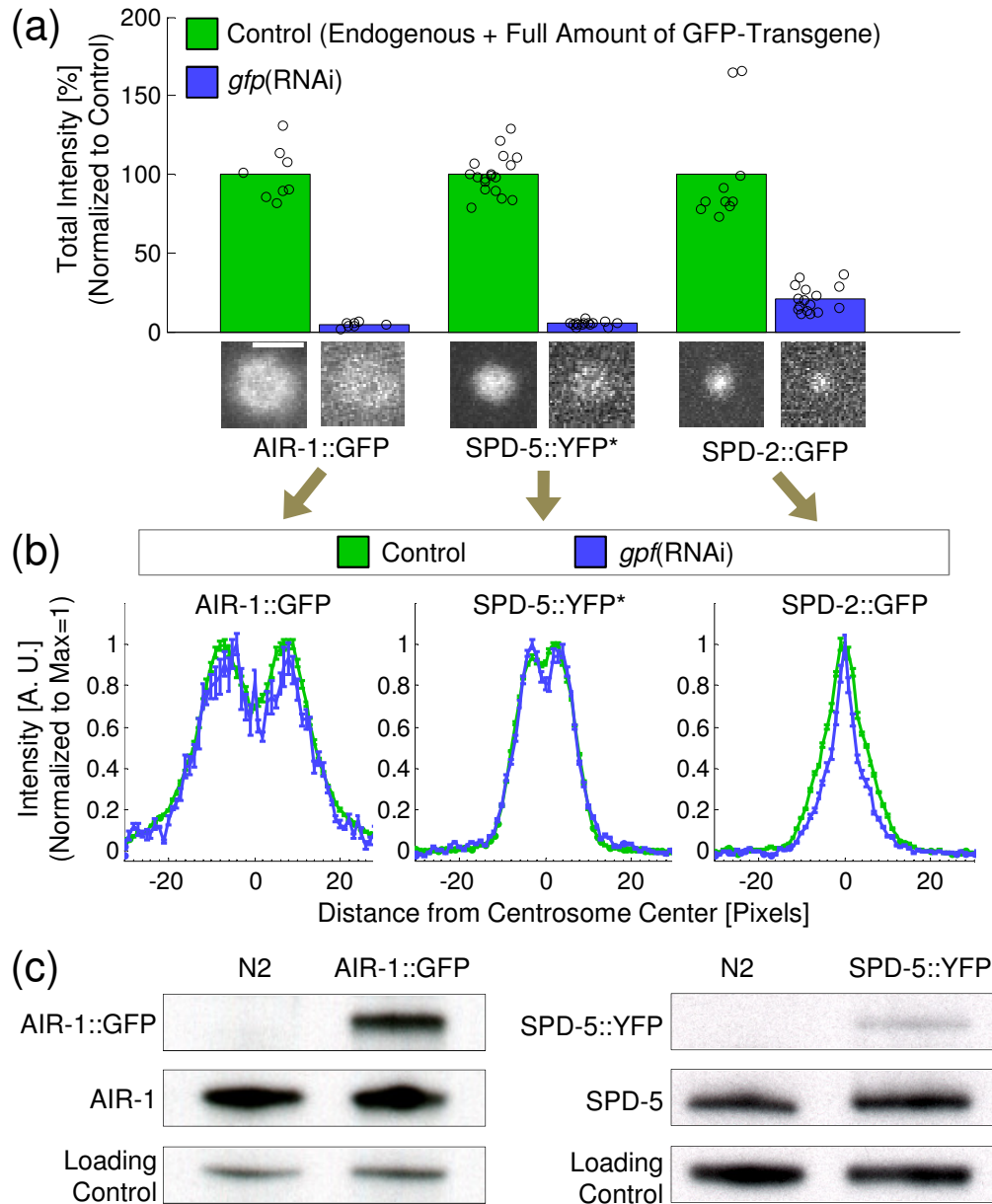


Figure 4.13: Centrosome size is sensitive to SPD-2::GFP. (a) Bar chart showing total centrosome intensity before and after partial RNAi against GFP. All values are normalized to the mean of the respective control embryos. AIR-1: $n = 4/4$, SPD-5: $n = 9/7$, SPD-2: $n = 5/8$ control/*gfp(RNAi)* embryos. Images underneath the bar chart represent metaphase centrosomes all at the same magnification and each displayed at its individual minimum-to-maximum intensity range. Scale bar is 3 μm . (b) Line scans (see Chapter 7) through centrosomes with and without RNAi against GFP. Error bars are standard error of the mean. (c) Western blots indicating the expression levels of the endogenous and respective exogenous gene. While AIR-1 levels are significantly increased in the presence of the respective transgene, SPD-5 levels are not. (*) RNAi against GFP is also effective against YFP.

same despite the availability of additional protein, or (2) the centrosomes become bigger. We cannot measure centrosome size in embryos that have only the endogenous protein (i.e., N2) as their centrosomes are not fluorescently labeled. Nonetheless, we can distinguish between the two possible effects by partially depleting only the exogenous protein using RNAi against GFP so that the centrosomes are still bright enough to be measurable. If centrosome size remains the same, then RNAi against the GFP tagged protein will have no effect on centrosome size. In the case of bigger centrosomes, RNAi against the GFP tagged protein will result in smaller centrosomes.

Figure 4.13 shows the results of this experiment for the three proteins AIR-1, SPD-2 and SPD-5. Since RNAi against GFP resulted in dramatically dimmer centrosomes (Figure 4.13 (a)), there was significantly more spread in the radii computed by our measuring algorithm (Section 2.3) than for the control embryos with bright centrosomes. We therefore decided to score centrosome size qualitatively by computing line scans through the sub-pixel accurate centrosome centers determined by the Gaussian fit (Figure 4.13 (b)). This clearly showed that for AIR-1 and SPD-5 the additional transgene had no effect on centrosome size, while for SPD-2 centrosome size was affected in the presence of the transgene. To test the significance of this result, we quantified the amounts of endogenous and exogenous protein by western blotting (Figure 4.13 (c)). Introducing the respective GFP transgene almost doubled the total amount of AIR-1 in the embryo, but the levels of SPD-5 increased by only about 10 to 20 per cent. We conclude that the amount of AIR-1 does not directly determine centrosome size. For SPD-5 we cannot make a definitive statement.

Interestingly, for SPD-2 we found a compensation mechanism that tightly regulates the total amount of SPD-2 protein, even in the presence of an additional SPD-2::GFP transgene. Unlike AIR-1, the embryo down-regulates the amount of endogenous SPD-2 to keep the total amount of SPD-2 constant (Figure 4.14). We could, however, employ a second over-expression technique to overcome this regulation. By optimizing transgene codon usage [118] we could raise SPD-2 protein expression beyond wild type levels (Figure 4.15(a)). Briefly, codon optimization exploits the fact that the same amino acid can be encoded

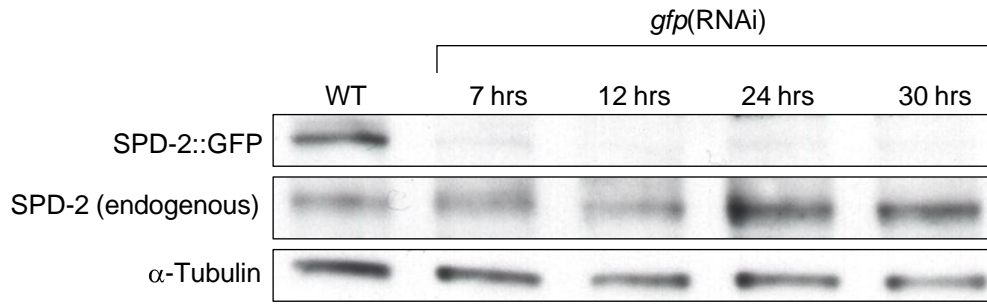


Figure 4.14: SPD-2 amounts are tightly regulated in the embryo. Worms expressing a SPD-2::GFP fosmid (TH231) were subjected to increasing times of *gfp(RNAi)*. Each lane was loaded with lysate corresponding to 40 gravid adult hermaphrodites. A SPD-2 antibody was used for blotting that recognizes both endogenous and transgene. α -Tubulin was used as a loading control.

by more than codon (i.e., a series of three nucleotides on the DNA) and that different codons are expressed at different levels. By designing a transgene with a more frequent usage of optimal codons than in wild-type the protein can be over-expressed. We then tracked and measured centrosomes in one- and four-cell embryos of the over-expression line (SPD-2::GFP[CAIo.37]) and found that increasing SPD-2 protein amount indeed increased centrosome size (Figure 4.15 (b) - (f)). For the four-cell stage the growth curve in Figure 4.15(f) additionally shows that the increase in centrosome size is due to an increased growth rate, i.e., the centrosome grows faster but then plateaus at the same time as the wild-type centrosome. For the one-cell stage an alternative interpretation of the growth curve in Figure 4.15(d) is that the centrosome becomes bigger because it starts to grow earlier in the presence of additional SPD-2.

Taken together, the experiments in this and the previous section demonstrate a direct role of SPD-2 in regulating centrosome size.

4.10 CYTOPLASMIC INTENSITY ANALYSIS

If the amount of SPD-2 depleted from the cytoplasm dictates the size of the centrosome, we would expect that centrosome growth correlates with the cytoplasmic concentration of SPD-2. As long as the centrosome grows the concentration should drop and as soon as centrosome size plateaus so should the

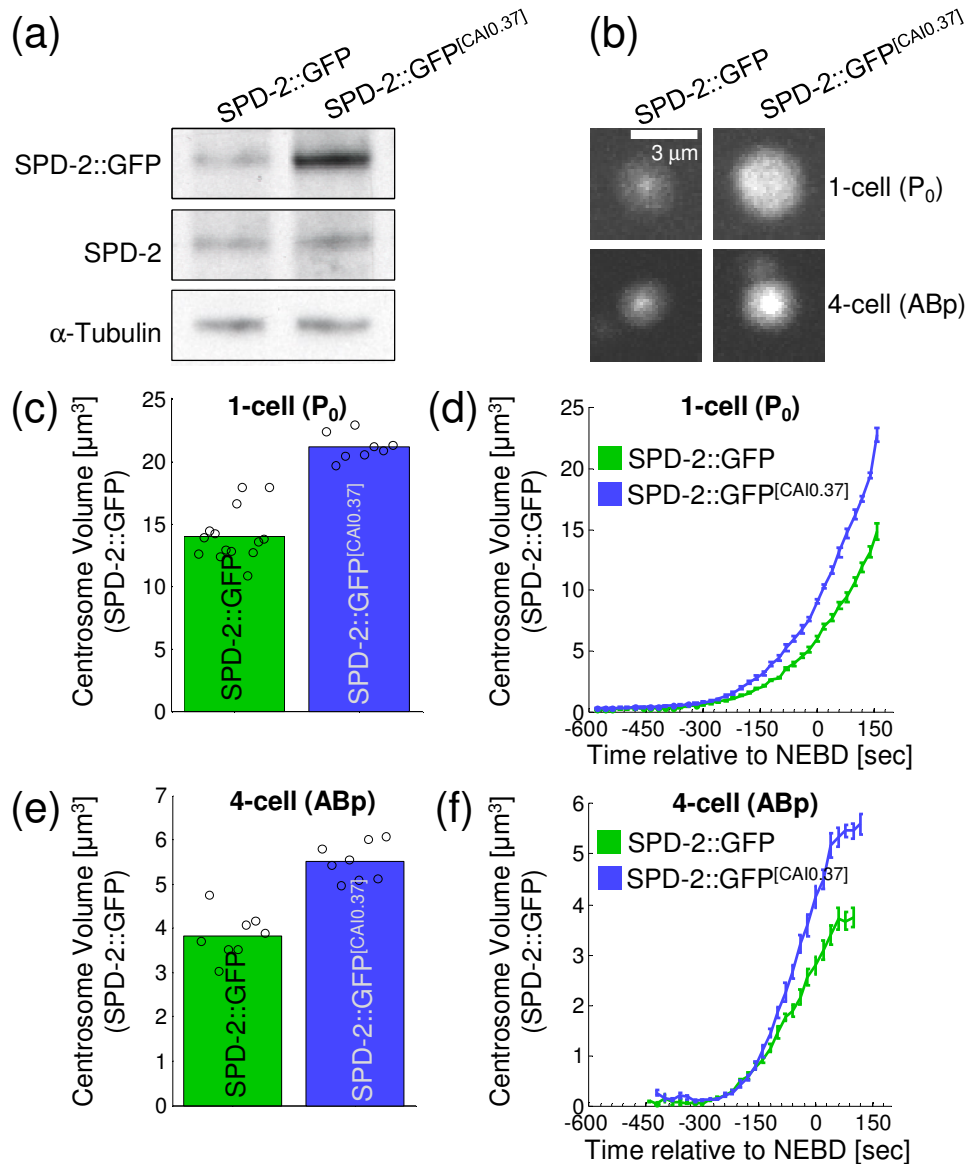


Figure 4.15: The size of centrosomes is sensitive to total SPD-2 amounts. (a) Western blot illustrating that codon-optimization of SPD-2 (SPD-2::GFP^[CA10.37]) results in higher amounts of SPD-2 transgene compared to non-optimized SPD-2 (SPD-2::GFP). α -Tubulin was used as a loading control. (b) Representative images (maximum projections) of centrosomes at metaphase in 1- and 4-cell embryos expressing different levels of SPD-2::GFP. (c) Centrosome volume in P_0 at metaphase in SPD-2::GFP embryos expressing different levels of SPD-2. Green bar: embryos expressing endogenous SPD-2 and a SPD-2::GFP transgene with native codon sequence, $n = 7$ embryos. Blue bar: embryos expressing endogenous SPD-2 and a SPD-2::GFP transgene with codon optimized sequence, $n = 4$ embryos. (d) Centrosome growth in P_0 for the same embryos as in (c). Error bars are standard error of the mean. (e)/(f) equivalent to (c)/(d) for centrosomes in ABp with $n = 4 / 4$ embryos for native / codon optimized SPD-2::GFP.

concentration. The cytoplasmic concentration is in principle proportional to the cytoplasmic intensity in the images. Unfortunately, due to imaging limitations, in particular photo-bleaching and signal attenuation in the z-direction, our data does not allow us to show a correlation between centrosome grow and cytoplasmic concentration.

To check whether there is any detectable depletion of SPD-2 (i.e., a decline in cytoplasmic SPD-2::GFP intensity), we performed the following experiment. Instead of taking time-lapse movies, we imaged one-cell embryos only once at the beginning of the centrosome cycle and once at the end. We refer to these time points as the early and the late time point, respectively. This should prevent almost any photo-bleaching. To measure and correct for auto-fluorescence we also imaged N2 worms. Firstly, this showed that autofluorescence makes up about 85 per cent of the entire cytoplasmic signal (Figure 4.16 (a)). As a consequence, even if the entire amount of SPD-2 was depleted from the cytoplasm, the intensity decline would at most be 15 per cent. Presumably, however, due to the binding constant of SPD-2 only a fraction of cytoplasmic SPD-2 can be recruited to the centrosome. Moreover, there is high variability in the measured intensity decline from the early to the late time point of the individual embryos (Figure 4.16 (b)). Therefore, it seems that the noise level in our images does not allow reliable quantification of how much the cytoplasmic concentration drops as the centrosome grows.

4.11 FLUORESCENCE RECOVERY AFTER PHOTO-BLEACHING EXPERIMENTS

In an effort to understand where within the PCM new molecules are incorporated we conducted FRAP experiments. FRAP is a technique that allows one to investigate the diffusion and motion of fluorescently labeled molecules. The idea is to deliberately destroy the fluorophores in a region of interest with a short laser impulse (photo-bleaching) and then to analyze the recovery of fluorescence in that region. Recovery occurs due to molecules from outside the photo-bleaching region that are still fluorescent.

Upon photo-bleaching of the entire centrosome, there are three possible ways how the fluorescence could recover: from inside out, from outside in

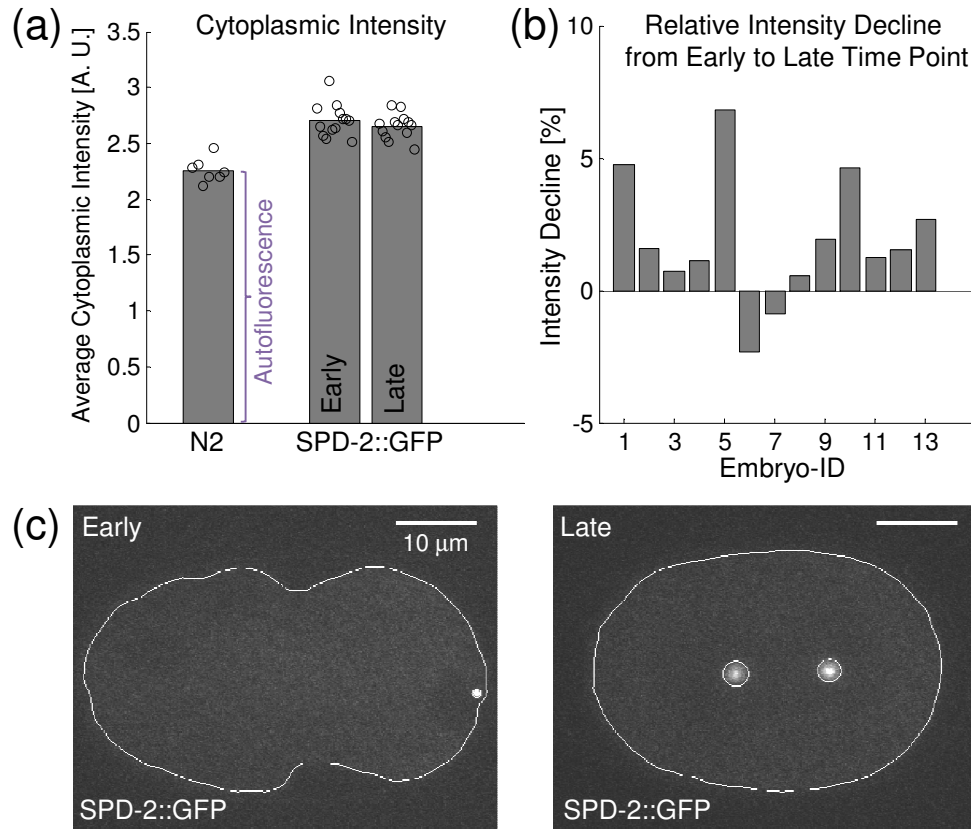


Figure 4.16: Cytoplasmic intensity in N2 (caused by autofluorescence) and SPD-2::GFP one-cell embryos. The embryo area has been identified by automatic segmentation based on background subtraction and thresholding of every single plane of the z-stack. The plane with maximum area was selected to compute the cytoplasmic intensity as the average intensity within the embryo but outside the centrosomes. (a) Autofluorescence makes up about 85 per cent of the cytoplasmic SPD-2::GFP signal. The absolute SPD-2::GFP intensities differ only slightly between the beginning (early) and the end (late) of the centrosome cycle. N2: $n = 7$, SPD-2::GFP: $n = 13$. (b) Cytoplasmic intensity decline from early to late time point for the 13 imaged SPD-2::GFP embryos after subtracting mean autofluorescence. (c) Maximum embryo-area planes of an embryo at the early and the late time point overlayed with the identified embryo and centrosome contours.

or evenly all over the centrosome region. Note that only the fluorophores are destroyed but the actual centrosome structure is (hopefully) not affected by photo-bleaching. Recovery from inside out would suggest that molecules are incorporated only at the center of the PCM (possibly via an interaction with the centrioles), whereas recovery from outside in would suggest that molecules are incorporated only at the 'surface' of the PCM (and then possibly internally reorganized to get to the interior of the PCM). Recovery evenly all over the centrosome is more problematic. The obvious interpretation is that molecules can be incorporated anywhere in the PCM. However, incorporation only at the center or only at the 'surface' is also possible in this scenario because we cannot exclude internal re-organization of the PCM on time scales that are not captured in our images.

We performed FRAP experiments for SPD-2::GFP one-cell embryos (Figure 4.17). To maximize the temporal resolution the embryos were not imaged in their full depth but only a sufficiently large subregion containing the centrosomes (19 z-planes at 500 nm steps), which allowed us to image at 4 second intervals. The image sequence of a single photo-bleached centrosome is shown in Figure 4.17(a). To quantify the recovery behavior we computed for circular regions of various radii centered at the centrosome center how long it takes for each region to recover to one third of its fluorescence before photo-bleaching (Figure 4.17(b)-(d)). This showed that recovery speed is independent of the distance to the centrosome center (Figure 4.17(d)), suggesting that recovery occurs evenly all over the centrosome. Thus, we cannot make a definite statement as to where molecules are incorporated into the PCM. In further FRAP experiments the temporal resolution will need to be further increased and ways to correct for the superposition of fluorescence recovery with the actual centrosome maturation process should be explored.

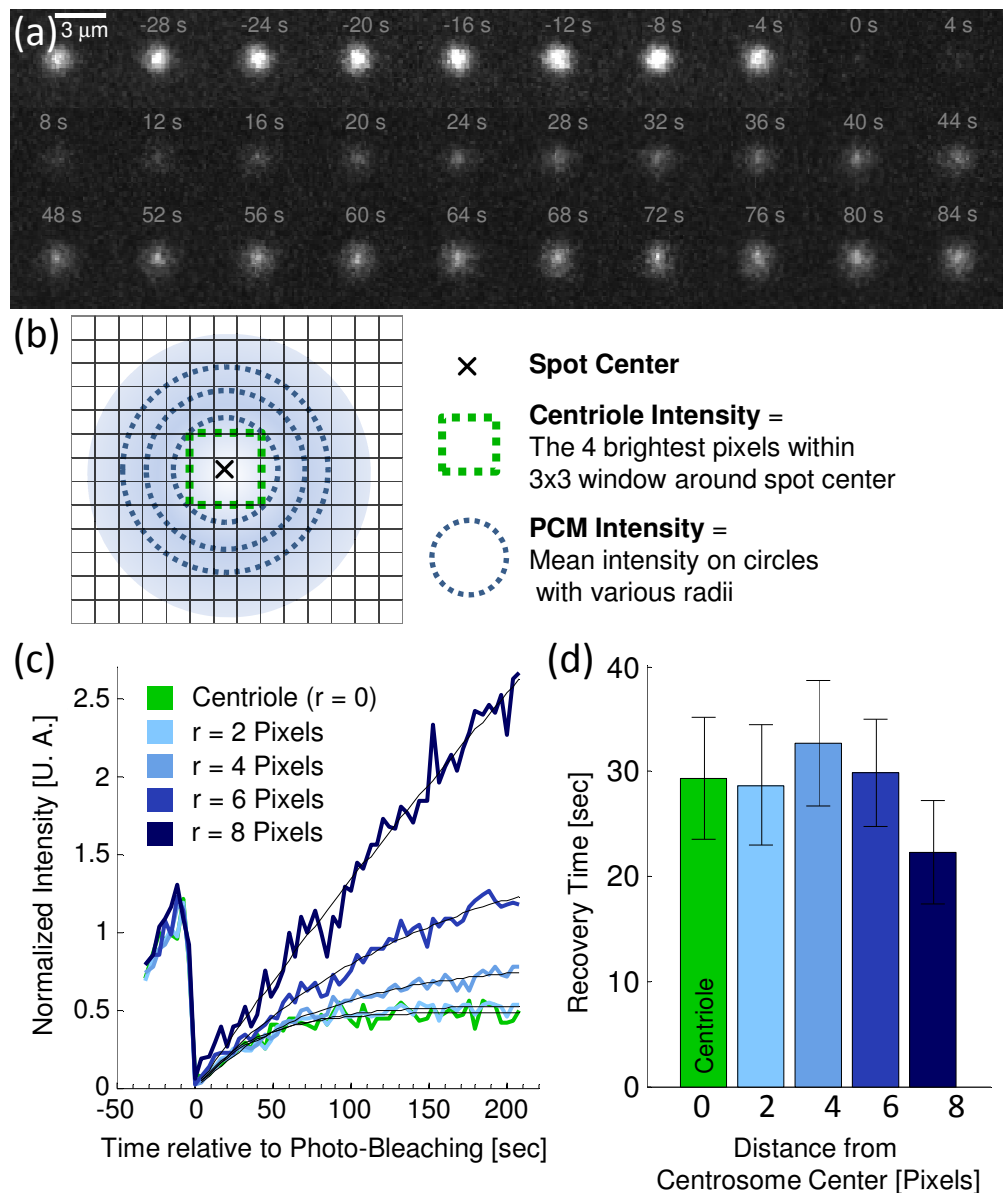


Figure 4.17: Fluorescence recovers evenly over the centrosome after photo-bleaching. (a) Images (z-planes through the center) of a photo-bleached SPD-2::GFP centrosome imaged at 4 second intervals. (b) Measuring local fluorescence within the centrosome for quantification of fluorescence recovery after photo-bleaching. (c) Fluorescence recovery curves computed as illustrated in (b) for the centrosome shown in (a). The pixel size of the images is 133 nm. Each intensity curve is normalized by its mean intensity at the last three time points before photo-bleaching and fitted with an exponential function $I(t) = A \cdot (1 - e^{-\tau \cdot t})$. (d) Recovery time for each centrosome region measured as the time it takes to recover to one third of its fluorescence before photo-bleaching ($n=3$). Error bars are standard error of the mean.

SUMMARY AND DISCUSSION

In animal cells, the centrosome is a dynamic organelle with several important functions. Towards a more detailed understanding of the molecular organization of the centrosome, we investigated the mechanism by which the size of the centrosome is regulated in *C. elegans* embryos. The key idea to improve upon previous studies of this problem was to quantitatively describe how the centrosome grows during the cell cycle over multiple rounds of cell divisions in the early embryo. To facilitate this study, we designed an image-based assay for centrosomes in *C. elegans* embryos which required the development and implementation of several computer algorithms for:

- detecting and tracking fluorescently labeled centrosomes in 3D time-lapse images
- measuring the size of the centrosomes
- detecting the nuclear envelope break down
- determining the identity of the cells within *C. elegans*' invariant cell lineage
- characterizing protein amounts via intensity analysis, which required segmenting centrosome images into background, embryo and centrosome region
- segmenting cells in images of fluorescently labeled cell membranes (PH-domain::GFP).

In combination with genetic perturbation techniques this assay allowed us to systematically study the influence on centrosome size of individual centro-

some components and several parameters such as cell size, centrosome number or microtubule behavior. The analysis of the image data was completely automated apart from determining cell identities and the nuclear envelope breakdown in embryos with non-wild-type morphology (i.e., not exactly two centrosomes per cell). However, uncertainties in the data still arises due to biological variance, imaging artifacts that we could not correct for as well as measurement errors. This made it often necessary to manually inspect the images and verify all analysis results in detail to identify outliers and biological side effects such as cell cycle delay, which can be a time-consuming procedure. Nonetheless, the development of the centrosome assay made it possible to collect much more and less biased quantitative data of centrosome growth than would have been possible with manual analysis. The following sections first summarize and discuss the algorithmic aspects of the centrosome tracker and then the important biological observations and conclusions derived from the tracking data.

5.1 TRACKING ALGORITHM

The centrosome tracking algorithm (Section 2.2) is the computational core component of the assay developed in this thesis. It was designed with the goal of robustly tracking fluorescently labeled centrosomes in 3D image sequences of multi-cellular wild-type and mutant *C. elegans* embryos. This goal has been achieved; the tracking results for about one thousand centrosome movies required only minimal manual corrections and/or parameter adjustments.

The tracking algorithm first detects all spots that possibly correspond to centrosomes. It then links the detected objects between consecutive time points in a novel two-step approach. In the first step only objects bright enough to clearly not be noise are linked into trajectories. This yields one or more segments of each centrosome's full trajectory. In the second step these partial trajectories are extended forward and backward in time. In this respect our algorithm is most similar to the approach by Jaqaman et al. [93]. They also first assemble track segments that they are confident are correct; however, by

using user-defined thresholds on the edge weights rather than restricting the set of detected objects to a certain subset as in our case. In the second step, we iteratively extend the track segments and then only stitch broken tracks that abut in time to obtain full trajectories, whereas they bridge gaps by linking track segments end-to-start or end/start-to-middle into full trajectories to handle missed objects as well as objects that split or merge. In both methods user-defined thresholds on the edge weights are used in the second linking step.

One important aspect to obtain reliable track segments is the use of two competing measures of object coherence, the spatial distance and changes in appearance between consecutive time points. An object assignment in the first linking step is accepted only if the assignment occurs for both distance measures, i.e., the assignment is stable under the two different distance measures. This greatly minimizes the risk of false assignments. The use of two distance measures also enabled us to adaptively determine suitable values (from a range of given values) for two of the most crucial parameters for finding truly reliable track segments. Another important aspect is the choice of minimum-weight bipartite matching as the strategy to compute these assignments. Minimum-weight bipartite matching guarantees spatially minimal total assignment cost under a maximal number of unique object assignments between two consecutive time points. To handle a variable number of objects we rely on heuristics to prune the matching hypothesis graph so that a maximal number of assignments usually does not result in wrong assignments. However, this is not guaranteed. By considering only bright objects in the first linking step, virtually all false positive calls are removed from consideration but also true positive calls may temporally be filtered. In combination with computing a maximal number of object assignments, this can lead to serious assignment errors. Fortunately, many of these errors can be corrected during the track extension phase by the concept of exchange edges (see Section 2.2.7).

Overall, the tracking performance relies on three principles. First, in our layered approach we initially focus on sub-problems that can be solved with higher confidence than the tracking problem as a whole. These reliable partial solutions are then incrementally refined until the overall tracking problem

is solved. However, while the general idea of a layered approach could also be applied to other tracking problems, our concrete approach is tailored to the problem of centrosome tracking. Second, we make extensive use of prior knowledge but rely only on properties that are also valid for mutants. One example of this is the concept of the radius hysteresis (see Section 2.2.5), which implements the centrosome's tendency to increase or stay constant in size over time. Through this concept object assignments are made not only based on two consecutive time points, but instead, one aspect of the entire history (or future in the case of backward tracking) of the centrosome is taken into account. Third, our object detection algorithm is tuned so that false negative calls are extremely rare and occur at most at the very beginning and end of the centrosome cycle. As a consequence, however, the number of false positive calls is relatively high. This detection philosophy is in contrast to many other approaches such as [91, 119] that try to adjust the detection parameters so as to minimize the overall detection error by trading-off the number of false positive and false negative calls. We deal with false positive object calls through the two-step linking approach. In the first step, false positives (and some true positives) are temporarily filtered out based on their static properties, in this case the peak-to-background ratio. In the second step, now having statistical knowledge about the frame-to-frame behavior of the objects (i.e., dynamic properties), the matching graph can be pruned, thereby isolating virtually all false positives.

5.2 DETECTION OF SUBTLE PHENOTYPES INVALIDATES HYPOTHESES IN PREVIOUS STUDIES

The development and use of computational methods was an important part in this project. Beyond organizing the huge amount of image data and automating the creation of thousands of graphs, our centrosome assay has brought crucial advantages over previous studies with manual data analysis. First, due to the complete automation of the image analysis pipeline, including cell naming and detection of the nuclear envelope breakdown, we could process a much larger number of embryos. Second, although limited by imag-

ing artifacts and variations in imaging conditions, the Gaussian fitting based quantification method (see Section 2.3) provided consistent and objective measurements of centrosome size. Taken together, this made it possible to detect subtle phenotypes that were often not obvious to the human eye. For example, in two-cell *sas-4*(RNAi) embryos one centrosome assembles around a defective centriole and is smaller than in wild-type, while the other centrosome assembles around an intact centriole. Based on this observation made by Kirkham et al. in [44], they discuss that “the amount of SAS-4 at centrioles proportionally dictates the amount of centrosomal PCM”. What they missed is that the other centrosome with the intact centriole becomes bigger than in wild-type. In fact, the volumes of the two *sas-4*(RNAi) centrosomes add up to the same total volume as the two wild-type centrosomes (see Figure 4.9). While the centriole might still have a role in setting centrosome size, this means that SAS-4 does not proportionally dictate the amount of centrosomal PCM, thus invalidating the hypothesis proposed in [44].

5.3 LIMITING COMPONENT HYPOTHESIS FOR CENTROSOME SIZE

Tracking centrosome growth in wild-type embryos has yielded several interesting results. Based on measurements up to the 16-cell stage, we found that

- beginning at the four-cell stage, centrosomes size plateaus before metaphase
- centrosomes grow slower in smaller cells.

Based on measurements up to the 4-cell stage, we found that

- centrosome volume scales proportionately with cell volume.
- the total volume of all centrosomes at metaphase is constant at any one cell stage.

These observations suggest an attractively simple mechanism by which centrosome size is controlled by the availability of one or more structural centrosome components in the cytoplasm. In this limiting component model, the embryo is initially provided with a fixed amount of each protein including

the limiting component(s) by the mother. The centrosome grows by recruiting the limiting component(s) from the cytoplasm. As a consequence, the concentration of the limiting component(s) in the cytoplasm drops. When this concentration reaches a certain level, defined by the binding constant of the limiting component(s), there will be an equilibrium between molecules being incorporated into the PCM and molecules being released from the PCM. Thus centrosome size plateaus. At the end of the cell cycle, when the centrosome disintegrates, all PCM components are released back to the cytoplasm and cytokinesis will divide the cytoplasmic bag proportionately to cell volume. As a result, the total amount of the limiting component(s) is distributed proportionately to the volume of the daughter cells.

The rate at which the PCM grows is likely be determined by the cytoplasmic concentration of one or more centrosome components. The concentration in turn is determined by the total amount of the component and the volume of the cell. This could explain why the centrosomes grow slower in smaller cells. As a consequence of proportionately distributing the limiting component(s) to the daughter cells, the total amount in each daughter cell will be lower than in the mother cell but the initial concentration will be the same. However, removing one molecule from a small volume will result in a larger concentration decline than removing one molecule from a bigger volume. Thus, over time the concentration drops more quickly in small cells and hence the PCM grows slower.

Our experiments have shown that SPD-2 is one such component that determines centrosome size. Not only does partial depletion of SPD-2 result in smaller centrosomes, but - as one would expect from a limiting component - increasing the amount of SPD-2 indeed results in bigger centrosomes. Our finding that the embryo keeps tight control over SPD-2 levels by a compensation mechanism (see Figure 4.14), and the role of SPD-2 in centriole duplication and PCM assembly [45], are further evidence that SPD-2 has a direct role in determining centrosome size. This is also consistent with previous work that suggest a role of the centriole in determining PCM size [45, 44].

In P_0 , AB and P_1 centrosome size does not plateau. The centrosomes continuously grow until disintegration at anaphase (see Figure 4.1). This might

be because the available amount of the limiting component(s) in these cells is higher than can be recruited to the centrosomes and thus the cytoplasmic concentration does not drop to the critical level where centrosome size plateaus. Supporting this idea, in smaller two-cell *ani-2*(RNAi) embryos (see Figure 4.5), as well as in two-cell embryos with more centrosomes (see Figure 4.8(b)), centrosome size does plateau. Contradictory to this idea, however, is our observation that the total centrosome volume is independent of centrosome number (see Figures 4.7, 4.8(a)/(c) and 4.9(a)). An abundance of the limiting component(s) in the one- and two-cell stage would imply that the total centrosome volume increases when the number of centrosomes increases (up to a level where the component actually becomes limiting). In line with this, one would also expect that centrosome size plateaus upon sufficiently long RNAi against the limiting component, but we have never seen a plateau in P_0 under such conditions (see Figure 4.10).

The difference between making cells smaller and RNAi against a selected centrosome component is that in the former case the amount of *all* components is reduced while in the latter case only the amount of the *selected* component is reduced. One possibility is that SPD-2 is not the only component that directly controls centrosome size. In a scenario with multiple limiting components, the amount of SPD-2 could only limit the growth rate, but the centrosome would eventually run out of another component. Supporting this idea, increased amounts of SPD-2 result in increased growth rate in 4-cell embryos while centrosome size still plateaus (see Figure 4.15(f)). Moreover, SPD-2 is not the only component required for a full-size centrosome (see Figure 4.10). However, there cannot be more than one component of which increased amounts are required for bigger centrosomes, because increasing SPD-2 amounts alone was sufficient for bigger centrosomes (see Figure 4.15) despite all other proteins remaining at their wild-type levels.

Alternatively, cell size could be a crucial factor in the process that transports PCM components to the centrosome. In a scenario with SPD-2 being the only limiting component, SPD-2 could indeed be present in excessive amounts in P_0 , AB and P_1 but the supply to the centrosome depends on cell size. We know little about the diffusion behavior of the individual centrosome components

and possibly involved active transport mechanisms via microtubules that are important for PCM assembly and maturation. A more detailed study of the cytoplasmic protein concentrations, and in particular their local distributions, will be necessary to evolve the simple idea of a limiting component to a more complex model that explains these conflicting observations.

Centrosomes are not the only organelle whose size scales with cell size. In *C. elegans* embryos mitotic spindle length also responds to changes in cell size [69, 120], and examples in other systems include the nucleus and mitochondria for which the organelle-to-cell volume ratio is held constant [121, 122]. A limiting component could be a general mechanism to ensure that the size of intracellular structures adapts to cell size in a simple and robust way. It abstracts the problem of regulating organelle size to the problem of controlling the amounts of the organelle's structural component(s).

DIRECTIONS FOR FUTURE WORK

6.1 TRACKING ALGORITHM

The tracking algorithm developed in this thesis was specifically designed for centrosomes in *C. elegans* embryos. One generally applicable idea thereby was to find a set of object links that are stable under two different measures of object coherence. For a more general tracking algorithm, this idea could be extended to other dimensions of the algorithm such as the linking strategy. Reliable track segments could then be established using only links that are stable under both, different measures of object coherence and different linking strategies. Moreover, specifically to further improve the establishment of centrosome core trajectories linking errors might be avoided by finding links that are stable under different subsets of the detected objects.

6.2 AUTOMATED IMAGE ACQUISITION AND GENOME-WIDE SCREEN

Our centrosome assay is automated in terms of image analysis, computing statistics over the extracted features and visualization of the results in the form of graphs and animations. Image acquisition, however, is done manually and represents the rate-limiting step in our experiments. Increasing the throughput by (semi-)automated image acquisition would improve the assay in two ways. First, the data quality could be improved by imaging control embryos in each individual imaging session. This would exclude imaging parameters such as the laser power or the room temperature as source of any spread in the data. Second, the number of experiments could be increased

and even a genome-wide screen could be setup to obtain a complete picture of which genes affect centrosome size.

One of the most time-consuming imaging steps is finding an embryo in the right developmental stage, usually an early one-cell embryo. This step is independent of the fluorescent centrosome label as it is performed using transmitted light microscopy. Automatically finding suitable embryos is thus one step towards automated image acquisition that would also be of interest for many other *C. elegans* research projects.

6.3 IMAGING WITH SINGLE PLANE ILLUMINATION MICROSCOPY

One limitation in quantifying centrosome dynamics over development was that accurate quantification of all centrosomes was possible only up to the four-cell stage due to the attenuation of the signal in the z-direction. Moreover, imaging with sufficient time-resolution was possible only up to the 16-cell stage due to limited laser exposure the embryo can cope with. Both limitations might be overcome by using single plane illumination microscopy (SPIM) [106] in conjunction with the multi-view registration and fusion method described in [107]. In Chapter 3, we used this method to image PH-domain::GFP embryos. In SPIM, the embryos are placed in a glass capillary rather than on a microscope slide. One practical problem is therefore that finding a suitable embryo is even more time-consuming because a larger volume has to be searched.

6.4 EXPERIMENTS IN THE CONTEXT OF THE LIMITING COMPONENT HYPOTHESIS

One of the most important questions in the context of the limiting component hypothesis is whether only one or several components are directly involved in regulating centrosome size. Based on our over-expression experiments (see Section 4.9) we could rule out AIR-1 as a potential limiting component, but not SPD-5. This was because introducing a SPD-5 transgene increased the total amount of SPD-5 protein only slightly. Increasing the levels of SPD-5 by codon optimization is a promising alternative.

For SPD-2 and SPD-5 we have seen that RNAi against either component leads to smaller centrosomes but that centrosome size does not plateau in one-cell embryos (see Figure 4.10). It would thus be worthwhile to perform double RNAi against SPD-2 and SPD-5 as this might create similar conditions as in embryos of smaller size where centrosome size plateaus (see Figure 4.5).

One question we could not address is: How does centrosome size depend on the time it has to grow? One could examine centrosome growth at different temperatures, but this affects the dynamics of all processes in the embryo at the same time. Another possibility would be to initiate centrosome growth earlier, or to delay the disintegration process. The disintegration of the centrosome is one aspect of the centrosome cycle that is currently only poorly understood. Our centrosome assay could be used to identify regulators of centrosome disintegration and then investigate to what extent centrosome growth depends on time.

Finally, whether the idea of a limiting component holds true for other organelles could be investigated by systematically increasing and decreasing the amounts of the components of which they are assembled.

MATERIALS AND METHODS

WORM STRAINS AND RNA INTERFERENCE

Maintenance of *C. elegans* worm strains was carried out according to standard protocols [123]. The following worm strains were used: Wild-type Bristol N2, TH27 (γ -Tubulin::GFP), OD58 (PH-domain::GFP) [105], TH97 (SPD-5::YFP), TH225 (Air-1::GFP ; fosmid), DH1 (*zyg-1(b1)*), *zyg-1(it29)*, TH257 (SPD-2::GFP ; CAIo.27), TH303 (SPD-2::GFP CAI 0.37), TH231 (SPD-2::GFP ; CAIo.27 fosmid). The three SPD-2::GFP lines used in this study (TH231, TH257 and TH303) differ in their codon adaptation index (CAI). The N-terminus of the *spd-2* transgene in TH303 was codon optimized (Geneart) with respect to the codon bias in *C.elegans*, so that the overall CAI score increased from the native value of 0.27 (TH231 and TH257) to 0.37. In addition, in TH257 and TH303, the transgenic *spd-2* sequences feature a 551 bp C-terminal stretch that was also subjected to codon optimization. However, for this fragment, codons were chosen in a way that would not increase the CAI score but simply achieve maximal divergence to the native codons. In this way, respective lines were made resistant to the *spd-2*(RNAi) feeding clone that was used in this study (see below).

All genes with the exception of TH225 (Air-1::GFP ; fosmid) and TH231 (SPD-2::GFP ; CAIo.27 fosmid), were cloned into a tagging plasmid (pAZ132), and expression was driven by the *pie-1* promotor. TH225 and TH231 were generated by recombineering in a fosmid containing *air-1* or *spd-2*, respectively, under its endogenous promotor, tagged with GFP at the C-terminus [124]. All transgenic lines were created by microparticle bombardment as described [125]. Worms were subjected to both RNAi by injection, and RNAi by

feeding as previously described [126, 127]. The RNAi method used for each gene, as well as the region targeted is indicated in the following list:

RNAi by injection

Primer sequences for the regions targeted are as follow, with T₃/T₇ extensions underlined:

ani-2:

TAATACGACTCACTATAGGGAGACCA CCAACGACTCCAAACGTCAGATA
AATTAACCCTCACTAAAGGGTCTCGTCCG TTTCTTGTTTCT

Template: N2 genomic DNA

RNAi by feeding

The regions targeted by the feeding clone are flanked by the following primers:

sas-4:

Forward primer: AACTGGTGGCACTGGATTC

Reverse primer: ATGAACGATTTAGGCGTTGG

gfp:

Forward primer: ATGAGTAAAGGAGAAGAACT

Reverse primer: TTTGTATAGTTCATCCATGC

Template: TH303

spd-2:

Forward primer: GAATGCATCAATGGCAGATG

Reverse primer: AATTTTGTGCCGGTACTTCG

LIVE CELL IMAGING

For live imaging of centrosomes and cell boundaries, worms were dissected on glass cover slips in M9 buffer and then mounted on 2 % agar pads. Imaging was conducted at 25°C and carried out on a spinning disk confocal system

consisting of an Olympus IX8 inverted stand microscope body equipped with a Prior NanoScanZ, a Yokogawa CSU-X1 spinning disk head, a 488 nm solid-state laser Coherent Sapphire (50mW) and an Andor iXon EM+ DU-897 BV back illuminated EMCCD camera for detection. The lens used was an Olympus UPlanSApo 60x 1.20 W. The microscope was controlled by the iQ 1.10.2 software.

For live-cell imaging of cell boundaries with SPIM, worms expressing a fusion between GFP and a PH-domain (OD58) were dissected on glass cover slips in M9 buffer and embryos mounted in a capillary filled with 1% low melting point agarose in PBS with 0.1 % Tween. Z-Stacks (x/y pixel size: 0.365 μm , z-step size: 1 μm) from three different views (0°, 120° and 240°) were acquired at 1-minute intervals using a prototype selective plane illumination microscope developed by Carl Zeiss MicroImaging and equipped with a Zeiss 40 \times /0.8 Achroplan objective. The individual views were registered and fused into single views as described in [107] to give a final pixel size of 0.365 μm in all three dimensions.

CENTROSOME GROWTH RATES

An approximation to the time-derivative at each sampling time point of the centrosome growth curves was computed by locally fitting a straight line to the curve in a window of ± 60 seconds centered at the respective time point.

LINE SCANS

The line scan of each individual centrosome is an average over various scanning lines angled at 0 to 90 degrees in 10 degree steps through the sub-pixel accurate center of the centrosome spot. The average of such line scans was then computed over several centrosomes. To account for different overall intensity levels of the centrosomes each average line scan was divided by its maximum intensity.

WESTERN BLOTTING

For comparing expression levels, 40 adult gravid hermaphrodites were picked into 10 μ l of water and snap frozen in liquid nitrogen, then thawed, mixed with 10 μ l sample buffer (125 mM Tris, pH 6.8, 6% SDS wt/vol, 10% vol, 20% glycerol vol/vol) and loaded onto a NuPage 3-8% Tris-Acetate gel (for SPD-2) or a NuPage Novex 4-12% Bis-Tris gel (for AIR-1) from Invitrogen. Antibodies recognizing both endogenous and the respective exogenous protein were used. A nonspecific band produced by the antibody was used as a loading control for AIR-1::GFP and SPD-5::YFP. Immunoblots were probed using primary antibodies at 1 μ g/ml and detected using a HRP-conjugated secondary antibody (1:10,000; Bio-Rad Laboratories).

BIBLIOGRAPHY

- [1] H. Peng, "Bioimage informatics: a new area of engineering biology," *Bioinformatics*, vol. 24, no. 17, p. 1827, 2008.
- [2] H. Dankert, L. Wang, E. Hoopfer, D. Anderson, and P. Perona, "Automated monitoring and analysis of social behavior in *Drosophila*," *Nature methods*, vol. 6, no. 4, pp. 297–303, 2009.
- [3] W. Geng, P. Cosman, M. Palm, and W. Schafer, "Caenorhabditis elegans egg-laying detection and behavior study using image analysis," *EURASIP journal on applied signal processing*, vol. 2005, p. 2240, 2005.
- [4] K. Huang, P. Cosman, and W. Schafer, "Automated detection and analysis of foraging behavior in *Caenorhabditis elegans*," *Journal of neuroscience methods*, vol. 171, no. 1, pp. 153–164, 2008.
- [5] "The diadem challenge." <<http://diademchallenge.org>>, last accessed: August 01 2010.
- [6] H. Peng, Z. Ruan, D. Atasoy, and S. Sternson, "Automatic reconstruction of 3D neuron structures using a graph-augmented deformable model," *Bioinformatics*, vol. 26, no. 12, p. i38, 2010.
- [7] B. Roysam, W. Shain, and G. Ascoli, "The central role of neuroinformatics in the national academy of engineering's grandest challenge: Reverse engineer the brain," *Neuroinformatics*, vol. 7, no. 1, pp. 1–5, 2009.
- [8] J. Rittscher, "Characterization of Biological Processes through Automated Image Analysis," *Annual review of biomedical engineering*, 2010.
- [9] S. Megason, "In toto imaging of embryogenesis with confocal time-lapse microscopy," *Methods Mol Biol*, vol. 546, pp. 317–332, 2009.

- [10] B. Null, C. Liu, M. Hedehus, S. Conolly, and R. Davis, "High-resolution, in vivo magnetic resonance imaging of *Drosophila* at 18.8 Tesla," *PLoS One*, vol. 3, no. 7, 2008.
- [11] S. Preibisch, R. Ejsmont, T. Rohlfing, and P. Tomancak, "Towards digital representation of drosophila embryogenesis," in *Proceedings of 5th IEEE International Symposium on Biomedical Imaging*, pp. 324–327, 2008.
- [12] S. Preibisch, S. Saalfeld, T. Rohlfing, and P. Tomancak, "Bead-based mosaicing of single plane illumination microscopy images using geometric local descriptor matching," in *Proceedings of The International Society for Optical Engineering*, vol. 7259, 2009.
- [13] P. Keller, A. Schmidt, J. Wittbrodt, and E. Stelzer, "Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy," *Science*, vol. 322, no. 5904, p. 1065, 2008.
- [14] D. Pastor, M. Luengo-Oroz, B. Lombardot, I. Gonzalvez, L. Duloquin, T. Savy, P. Bourguine, N. Peyrieras, and A. Santos, "Cell tracking in fluorescence images of embryogenesis processes with morphological reconstruction by 4D-tubular structuring elements," in *Conference proceedings:... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 1, p. 970, 2009.
- [15] O. Dzyubachyk, R. Jelier, B. Lehner, W. Niessen, and E. Meijering, "Model-Based Approach for Tracking Embryogenesis in *Caenorhabditis Elegans* Fluorescence Microscopy Data," in *Conference proceedings:... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 1, p. 5356, 2009.
- [16] Z. Bao, J. Murray, T. Boyle, S. Ooi, M. Sandel, and R. Waterston, "Automated cell lineage tracing in *Caenorhabditis elegans*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 8, p. 2707, 2006.

- [17] F. Long, H. Peng, X. Liu, S. Kim, and E. Myers, "A 3D digital atlas of *C. elegans* and its application to single-cell analyses," *nature methods*, 2009.
- [18] E. Preston, B. Mericle, S. Batzoglou, E. Myers, and S. Kim, "Analysis of Cell Fate from Single-Cell Gene Expression Profiles in *C. elegans*," *Cell*, vol. 139, pp. 623–633, 2009.
- [19] C. Bakal, J. Aach, G. Church, and N. Perrimon, "Quantitative morphological signatures define local signaling networks regulating cell morphology," *science*, vol. 316, no. 5832, p. 1753, 2007.
- [20] J. Boysen, S. Fanning, J. Newberg, R. Murphy, and A. Mitchell, "Detection of Protein-Protein Interactions Through Vesicle Targeting," *Genetics*, vol. 182, no. 1, p. 33, 2009.
- [21] J. Newberg, J. Li, A. Rao, F. Pontén, M. Uhlén, E. Lundberg, and R. Murphy, "Automated analysis of human protein atlas immunofluorescence images," in *Proceedings of the Sixth IEEE international conference on Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1023–1026, IEEE Press, 2009.
- [22] K. Huang and R. Murphy, "Boosting accuracy of automated classification of fluorescence microscope images for location proteomics," *BMC bioinformatics*, vol. 5, no. 1, p. 78, 2004.
- [23] C. Wilson, Y. Si, C. Thompsons, A. Smellie, M. Ashwell, J. Liu, P. Ye, D. Yohannes, and S. Ng, "Identification of a small molecule that induces mitotic arrest using a simplified high-content screening assay and data analysis method," *Journal of Biomolecular Screening*, vol. 11, no. 1, p. 21, 2006.
- [24] C. Granas, B. Lundholt, F. Loechel, H. Pedersen, S. Bjorn, V. Linde, C. Krogh-Jensen, E. Nielsen, M. Praestegaard, and S. Nielsen, "Identification of RAS-mitogen-activated protein kinase signaling pathway modulators in an ERF1 redistribution (R) screen," *Journal of biomolecular screening*, vol. 11, no. 4, p. 423, 2006.

- [25] B. Neumann, M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok, and J. Ellenberg, "High-throughput RNAi screening by time-lapse imaging of live human cells," *Nature methods*, vol. 3, no. 5, pp. 385–390, 2006.
- [26] B. Neumann, T. Walter, H. Jean-Karim, *et al.*, "Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes," *Nature*, vol. 464, no. 7289, pp. 721–727, 2010.
- [27] A. Fire, S. Xu, M. Montgomery, S. Kostas, S. Driver, and C. Mello, "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*," *Nature*, vol. 391, no. 6669, pp. 806–811, 1998.
- [28] J. Shawe-Taylor and N. Cristianini, *Kernel methods for pattern analysis*. Cambridge Univ Pr, 2004.
- [29] C. Brangwynne, C. Eckmann, D. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Julicher, and A. Hyman, "Germline P granules are liquid droplets that localize by controlled dissolution/condensation," *Science*, vol. 324, no. 5935, p. 1729, 2009.
- [30] A. Altinok, M. El-saban, A. J. Peck, L. Wilson, S. C. Feinstein, B. S. Manjunath, and K. Rose, "Activity analysis in microtubule videos by mixture of hidden markov models," *IEEE CVPR 2006*, vol. 2, pp. 1662–1669, 2006.
- [31] M. Saban, A. Altinok, A. Peck, C. Kenney, S. Feinstein, L. Wilson, K. Rose, and B. Manjunath, "Automated tracking and modeling of microtubule dynamics," in *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006*, pp. 1032–1035, 2006.
- [32] S. Hadjidemetriou, D. Toomre, and J. Duncan, "Motion tracking of the outer tips of microtubules," *Medical image analysis*, vol. 12, no. 6, pp. 689–702, 2008.
- [33] I. Smal, K. Draegestein, N. Galjart, W. Niessen, and E. Meijering, "Particle filtering for multiple object tracking in dynamic fluorescence microscopy images: application to microtubule growth analysis," *IEEE transactions on medical imaging*, vol. 27, no. 6, p. 789, 2008.

- [34] É. van Beneden, *Recherches sur la maturation de l'oeuf, la fécondation, et la division cellulaire*. Librairie CLEMM, 1883.
- [35] T. Boveri, "Zellenstudien II: Die Befruchtung und Teilung des Eies von *Ascaris megalocephala*," *Jena Zeit. Naturw*, vol. 22, pp. 685–882, 1888.
- [36] J. Lüders and T. Stearns, "Microtubule-organizing centres: a re-evaluation," *Nature reviews. Molecular cell biology*, vol. 8, pp. 161–7, Feb. 2007.
- [37] S. Doxsey, D. McCollum, and W. Theurkauf, "Centrosomes in cellular regulation," *Annual review of cell and developmental biology*, vol. 21, p. 411, 2005.
- [38] J. Manneville and S. Etienne-Manneville, "Positioning centrosomes and spindle poles: looking at the periphery to find the centre," *Biology of the Cell*, vol. 98, pp. 557–565, 2006.
- [39] E. Nigg, ed., *Centrosomes in development and disease*, ch. 12. Vch Verlagsgesellschaft MbH, 2004.
- [40] C. Kemp, K. Kopish, P. Zipperlen, J. Ahringer, and K. O'Connell, "Centrosome maturation and duplication in *C. elegans* require the coiled-coil protein SPD-2," *Developmental cell*, vol. 6, no. 4, pp. 511–523, 2004.
- [41] E. Hannak, M. Kirkham, A. A. Hyman, and K. Oegema, "Aurora-A kinase is required for centrosome maturation in *Caenorhabditis elegans*," *The Journal of cell biology*, vol. 155, pp. 1109–16, Dec. 2001.
- [42] D. R. Hamill, A. F. Severson, J. C. Carter, and B. Bowerman, "Centrosome maturation and mitotic spindle assembly in *C. elegans* require SPD-5, a protein with multiple coiled-coil domains," *Developmental cell*, vol. 3, pp. 673–84, Nov. 2002.
- [43] K. O'Connell, C. Caron, K. Kopish, D. Hurd, K. Kemphues, Y. Li, and J. White, "The *C. elegans* *zyg-1* gene encodes a regulator of centrosome duplication with distinct maternal and paternal roles in the embryo," *Cell*, vol. 105, no. 4, pp. 547–558, 2001.

- [44] M. Kirkham, T. Müller-Reichert, K. Oegema, S. Grill, and A. A. Hyman, "SAS-4 is a *C. elegans* centriolar protein that controls centrosome size.," *Cell*, vol. 112, pp. 575–87, Feb. 2003.
- [45] M. Song, L. Aravind, T. Müller-Reichert, and K. O'Connell, "The conserved protein SZY-20 opposes the Plk4-related kinase ZYG-1 to limit centrosome size," *Developmental cell*, vol. 15, no. 6, pp. 901–912, 2008.
- [46] R. Heald, R. Tournebise, A. Habermann, E. Karsenti, and A. A. Hyman, "Spindle assembly in *Xenopus* egg extracts: respective roles of centrosomes and microtubule self-organization," *The Journal of cell biology*, vol. 138, no. 3, p. 615, 1997.
- [47] N. Mahoney, G. Goshima, A. Douglass, and R. Vale, "Making microtubules and mitotic spindles in cells without functional centrosomes," *Current Biology*, vol. 16, no. 6, pp. 564–569, 2006.
- [48] R. Basto, J. Lau, T. Vinogradova, A. Gardiol, C. Woods, A. Khodjakov, and J. Raff, "Flies without centrioles," *Cell*, vol. 125, no. 7, pp. 1375–1386, 2006.
- [49] N. Rusan and G. Rogers, "Centrosome function: Sometimes less is more," *Traffic*, vol. 10, no. 5, pp. 472–481, 2009.
- [50] F. Motegi, N. Velarde, F. Piano, and A. Sugimoto, "Two phases of astral microtubule activity during cytokinesis in *C. elegans* embryos," *Developmental cell*, vol. 10, no. 4, pp. 509–520, 2006.
- [51] C. R. Cowan and A. A. Hyman, "Centrosomes direct cell polarity independently of microtubule assembly in *C. elegans* embryos," *Nature*, vol. 431, no. 7004, pp. 92–96, 2004.
- [52] V. Hachet, C. Canard, and P. Gönczy, "Centrosomes promote timely mitotic entry in *C. elegans* embryos," *Developmental cell*, vol. 12, no. 4, pp. 531–541, 2007.
- [53] A. Debec, W. Sullivan, and M. Bettencourt-Dias, "Centrioles: active players or passengers during mitosis?," *Cellular and Molecular Life Sciences*, pp. 1–22, 2010.

- [54] E. Nigg and J. Raff, "Centrioles, centrosomes, and cilia in health and disease," *Cell*, vol. 139, no. 4, pp. 663–678, 2009.
- [55] D. Zyss and F. Gergely, "Centrosome function in cancer: guilty or innocent?," *Trends in cell biology*, vol. 19, no. 7, pp. 334–346, 2009.
- [56] Y. Bobinnec, a. Khodjakov, L. M. Mir, C. L. Rieder, B. Eddé, and M. Bornens, "Centriole disassembly in vivo and its effect on centrosome structure and function in vertebrate cells.," *The Journal of cell biology*, vol. 143, pp. 1575–89, Dec. 1998.
- [57] A. Dammermann, T. Müller-Reichert, L. Pelletier, B. Habermann, A. Desai, and K. Oegema, "Centriole assembly requires both centriolar and pericentriolar material proteins.," *Developmental cell*, vol. 7, pp. 815–29, Dec. 2004.
- [58] D. Mazia, P. J. Harris, and T. Bibring, "The Multiplicity of the Mitotic Centers and the Time-Course of Their Duplication and Separation.," *The Journal of biophysical and biochemical cytology*, vol. 7, pp. 1–20, Feb. 1960.
- [59] G. Sluder and C. L. Rieder, "Centriole number and the reproductive capacity of spindle poles.," *The Journal of cell biology*, vol. 100, pp. 887–96, Mar. 1985.
- [60] L. Pelletier, N. O, E. Hannak, C. Cowan, B. Habermann, M. Ruer, T. Mu, and A. A. Hyman, "The *Caenorhabditis elegans* Centrosomal Protein SPD-2 Is Required for both Pericentriolar Material Recruitment and Centriole Duplication," *Current*, vol. 14, pp. 863–873, 2004.
- [61] R. Palazzo, J. Vogel, B. Schnackenberg, D. Hull, and X. Wu, "Centrosome maturation," *Current Topics in Developmental Biology*, vol. 49, pp. 449–470, 1999.
- [62] S. Strome, J. Powers, M. Dunn, K. Reese, C. Malone, J. White, G. Seydoux, and W. Saxton, "Spindle dynamics and the role of γ -tubulin in early *Caenorhabditis elegans* embryos," *Molecular biology of the cell*, vol. 12, no. 6, p. 1751, 2001.

- [63] M. Boxem, Z. Maliga, N. Klitgord, N. Li, I. Lemmens, M. Mana, L. de Lichtervelde, J. Mul, D. van de Peut, M. Devos, *et al.*, "A protein domain-based interactome network for *C. elegans* early embryogenesis," *Cell*, vol. 134, no. 3, pp. 534–545, 2008.
- [64] J. Azimzadeh and M. Bornens, "Structure and duplication of the centrosome," *Journal of Cell Science*, vol. 120, no. 13, p. 2139, 2007.
- [65] M. Bettencourt-Dias and D. Glover, "Centrosome biogenesis and function: centrosomics brings new understanding," *Nature Reviews Molecular Cell Biology*, vol. 8, no. 6, pp. 451–463, 2007.
- [66] F. Zhu, S. Lawo, A. Bird, D. Pinchev, A. Ralph, C. Richter, T. Müller-Reichert, R. Kittler, A. A. Hyman, and L. Pelletier, "The mammalian SPD-2 ortholog Cep192 regulates centrosome biogenesis.," *Current biology : CB*, vol. 18, pp. 136–41, Jan. 2008.
- [67] P. Meraldi and E. Nigg, "The centrosome cycle," *FEBS letters*, vol. 521, no. 1-3, pp. 9–13, 2002.
- [68] G. Sluder, "Two-way traffic: centrosomes and the cell cycle," *Nature Reviews Molecular Cell Biology*, vol. 6, no. 9, pp. 743–748, 2005.
- [69] G. Greenan, C. P. Brangwynne, S. Jaensch, J. Gharakhani, F. Jülicher, and A. A. Hyman, "Centrosome size sets mitotic spindle length in *Caenorhabditis elegans* embryos.," *Current biology : CB*, vol. 20, pp. 353–8, Feb. 2010.
- [70] W. F. Marshall, "Cellular length control systems.," *Annual review of cell and developmental biology*, vol. 20, pp. 677–93, Jan. 2004.
- [71] M. Gorjánácz, A. Jaedicke, and I. W. Mattaj, "What can *Caenorhabditis elegans* tell us about the nuclear envelope?," *FEBS letters*, vol. 581, pp. 2794–801, June 2007.
- [72] A. Biosciences, "Fluorescence imaging: principles and methods," 2010. Handbook.
- [73] J. Lichtman and J. Conchello, "Fluorescence microscopy," *Nature Methods*, vol. 2, no. 12, pp. 910–919, 2005.

- [74] R. Heintzmann and G. Ficz, "Breaking the resolution limit in light microscopy," *Methods in cell biology*, vol. 81, pp. 561–580, 2007.
- [75] E. Betzig, G. Patterson, R. Sougrat, O. Lindwasser, S. Olenych, J. Bonifacino, M. Davidson, J. Lippincott-Schwartz, and H. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, p. 1642, 2006.
- [76] M. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature methods*, vol. 3, no. 10, pp. 793–796, 2006.
- [77] W. Johnston, A. Krizus, and J. Dennis, "The eggshell is required for meiotic fidelity, polar-body extrusion and polarization of the *C. elegans* embryo," *BMC biology*, vol. 4, no. 1, p. 35, 2006.
- [78] Y. Kalaidzidis, "Intracellular objects tracking," *European journal of cell biology*, vol. 86, no. 9, pp. 569–578, 2007.
- [79] B. Müller Junior and R. Anido, "Distributed real-time soccer tracking," in *Proceedings of the ACM 2nd international workshop on Video surveillance & sensor networks*, p. 103, ACM, 2004.
- [80] D. Sage, F. Neumann, F. Hediger, S. Gasser, and M. Unser, "Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics," *IEEE transactions on image processing*, vol. 14, no. 9, 2005.
- [81] I. Smal, W. Niessen, and E. Meijering, "Advanced particle filtering for multiple object tracking in dynamic fluorescence microscopy images," in *4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007*, pp. 1048–1051, 2007.
- [82] Z. Khan, T. Balch, and F. Dellaert, "MCMC-based particle filtering for tracking a variable number of interacting targets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1805–1918, 2005.

- [83] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on signal processing*, vol. 50, no. 2, 2002.
- [84] A. Genovesio, T. Liedl, V. Emiliani, W. Parak, M. Coppey-Moisan, and J. Olivo-Marin, "Multiple particle tracking in 3-D+t microscopy: method and application to the tracking of endocytosed quantum dots," *IEEE Transactions on Image Processing*, vol. 15, no. 5, 2006.
- [85] I. Smal, W. Niessen, and E. Meijering, "Particle filtering for multiple object tracking in molecular cell biology," in *IEEE Nonlinear Statistical Signal Processing Workshop*, pp. 44–1, 2006.
- [86] W. Godinez, M. Lampe, S. Wörz, B. Müller, R. Eils, and K. Rohr, "Probabilistic tracking of virus particles in fluorescence microscopy image sequences," *Bildverarbeitung für die Medizin 2008*, pp. 448–452, 2008.
- [87] J. Yoon, A. Bruckbauer, W. Fitzgerald, and D. Klennerman, "Bayesian inference for improved single molecule fluorescence tracking," *Biophysical journal*, vol. 94, no. 12, pp. 4932–4947, 2008.
- [88] H. Gribben, P. Miller, J. Zhang, and M. Browne, "Poisson Kalman Particle Filtering for Tracking Centrosomes in Low-Light 3-D Confocal Image Sequences," in *Machine Vision and Image Processing Conference, 2009. IMVIP'09. 13th International*, pp. 83–88, 2009.
- [89] E. Cuevas, D. Zaldivar, and R. Rojas, *Kalman filter for vision tracking*. Freie Univ., Fachbereich Mathematik und Informatik, 2005.
- [90] R. Ghosh and W. Webb, "Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules," *Biophysical journal*, vol. 66, no. 5, pp. 1301–1318, 1994.
- [91] I. Sbalzarini and P. Koumoutsakos, "Feature point tracking and trajectory analysis for video imaging in cell biology," *Journal of Structural Biology*, vol. 151, no. 2, pp. 182–195, 2005.
- [92] R. Burkard and E. Cela, *Linear assignment problems and extensions*. Karl-Franzens-Univ. Graz & Techn. Univ. Graz, 1998.

- [93] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. Schmid, and G. Danuser, "Robust single-particle tracking in live-cell time-lapse sequences," *Nature methods*, vol. 5, no. 8, pp. 695–702, 2008.
- [94] D. Thomann, *Algorithms for detection and tracking of objects with super-resolution in 3D Fluorescence Microscopy*. PhD thesis, Swiss Federal Institute of Technology Zurich, 2003.
- [95] R. Haralick and L. Shapiro, *Computer and Robot Vision (Volume II)*. Boston, MA, USA: Addison-Wesley, 1992.
- [96] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on Optimization*, vol. 6, no. 2, pp. 418–445, 1996.
- [97] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, pp. 285–296, 1975.
- [98] K. Parvati, B. Rao, and M. Das, "Image segmentation using gray-scale morphology and marker-controlled watershed transformation," *Discrete Dynamics in Nature and Society*, vol. 2008, 2009.
- [99] "Stats4students." <<http://www.stats4students.com/Essentials/Standard-Score/Overview.php>>, last accessed: August 29 2010.
- [100] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the Society of Industrial and Applied Mathematics*, vol. 5, pp. 32–38, 1957.
- [101] J. Sulston, E. Schierenberg, J. White, and J. Thomson, "The embryonic cell lineage of the nematode *Caenorhabditis elegans*," *Developmental biology*, vol. 100, no. 1, pp. 64–119, 1983.
- [102] R. Schnabel and J. Priess, "Specification of the early embryonic cell fates in *C. elegans*," in *The Nematode C. elegans, Volume 2* (M. B. Riddle D., Blumenthal T. and P. J., eds.), Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1997.
- [103] B. Goldstein and S. Hird, "Specification of the anteroposterior axis in *Caenorhabditis elegans*," *Development*, vol. 122, no. 5, p. 1467, 1996.

- [104] M. Gorjánácz, A. Jaedicke, and I. Mattaj, "What can *Caenorhabditis elegans* tell us about the nuclear envelope?," *FEBS letters*, vol. 581, no. 15, pp. 2794–2801, 2007.
- [105] A. Audhya, F. Hyndman, I. X. McLeod, A. S. Maddox, J. R. Yates, A. Desai, and K. Oegema, "A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in *Caenorhabditis elegans*," *The Journal of cell biology*, vol. 171, pp. 267–79, Oct. 2005.
- [106] J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt, and E. Stelzer, "Optical sectioning deep inside live embryos by selective plane illumination microscopy," *Science*, vol. 305, no. 5686, p. 1007, 2004.
- [107] S. Preibisch, S. Saalfeld, J. Schindelin, and P. Tomancak, "Software for bead-based registration of selective plane illumination microscopy data," *Nature Methods*, vol. 7, no. 6, pp. 418–419, 2010.
- [108] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [109] L. E. Krueger, J.-C. Wu, M.-F. B. Tsou, and L. S. Rose, "LET-99 inhibits lateral posterior pulling forces during asymmetric spindle elongation in *C. elegans* embryos," *The Journal of cell biology*, vol. 189, pp. 481–95, May 2010.
- [110] Y. Bobinnec, M. Fukuda, and E. Nishida, "Identification and characterization of *Caenorhabditis elegans* gamma-tubulin in dividing cells and differentiated tissues," *Journal of cell science*, vol. 113 Pt 21, pp. 3747–59, Nov. 2000.
- [111] E. Hannak, K. Oegema, M. Kirkham, P. Gönczy, B. Habermann, and A. A. Hyman, "The kinetically dominant assembly pathway for centrosomal asters in *Caenorhabditis elegans* is gamma-tubulin dependent," *The Journal of cell biology*, vol. 157, pp. 591–602, May 2002.

- [112] A.-L. Schlaitz, M. Srayko, A. Dammermann, S. Quintin, N. Wielsch, I. MacLeod, Q. de Robillard, A. Zinke, J. R. Yates, T. Müller-Reichert, A. Shevchenko, K. Oegema, and A. A. Hyman, "The *C. elegans* RSA complex localizes protein phosphatase 2A to centrosomes and regulates mitotic spindle assembly," *Cell*, vol. 128, pp. 115–27, Jan. 2007.
- [113] A. S. Maddox, B. Habermann, A. Desai, and K. Oegema, "Distinct roles for two *C. elegans* anillins in the gonad and early embryo," *Development (Cambridge, England)*, vol. 132, pp. 2837–48, June 2005.
- [114] M. Srayko, S. Quintin, A. Schwager, and A. Hyman, "Caenorhabditis elegans TAC-1 and ZYG-9 form a complex that is essential for long astral and spindle microtubules," *Current Biology*, vol. 13, no. 17, pp. 1506–1511, 2003.
- [115] D. Bird and D. Riddle, "Molecular cloning and sequencing of *ama-1*, the gene encoding the largest subunit of *Caenorhabditis elegans* RNA polymerase II," *Molecular and Cellular Biology*, vol. 9, no. 10, p. 4119, 1989.
- [116] I. Maeda, Y. Kohara, M. Yamamoto, and A. Sugimoto, "Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi," *Current Biology*, vol. 11, no. 3, pp. 171–176, 2001.
- [117] M. Zhong, W. Niu, Z. Lu, M. Sarov, J. Murray, J. Janette, D. Raha, K. Sheaffer, H. Lam, E. Preston, *et al.*, "Genome-Wide Identification of Binding Sites Defines Distinct Functions for *Caenorhabditis elegans* PHA-4/FOXA in Development and Environmental Response," *PLoS Genetics*, 2010.
- [118] S. Redemann *et al.*, "Codon adaptation-based control of protein expression levels in *c. elegans*." submitted for publication, 2010.
- [119] V. Racine, A. Hertzog, J. Jouanneau, J. Salamero, C. Kervrann, and J. Sibarita, "Multiple-target tracking of 3D fluorescent objects based on simulated annealing," in *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 2006, pp. 1020–1023, 2006.

- [120] S. Dumont and T. J. Mitchison, "Force and length in the mitotic spindle.," *Current biology : CB*, vol. 19, pp. R749–61, Sept. 2009.
- [121] F. R. Neumann and P. Nurse, "Nuclear size control in fission yeast.," *The Journal of cell biology*, vol. 179, pp. 593–600, Nov. 2007.
- [122] K. Tanaka, T. Kanbe, and T. Kuroiwa, "Three-dimensional behaviour of mitochondria during cell division and germ tube formation in the dimorphic yeast *Candida albicans*," *Journal of cell science*, vol. 73, no. 1, p. 207, 1985.
- [123] S. Brenner, "The genetics of *Caenorhabditis elegans*," *Genetics*, vol. 77, no. 1, p. 71, 1974.
- [124] M. Sarov, S. Schneider, A. Pozniakovski, A. Roguev, S. Ernst, Y. Zhang, A. Hyman, and A. Stewart, "A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*," *Nature Methods*, vol. 3, no. 10, pp. 839–844, 2006.
- [125] V. Praitis, E. Casey, D. Collar, and J. Austin, "Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*," *Genetics*, vol. 157, no. 3, p. 1217, 2001.
- [126] R. Kamath, M. Martinez-Campos, P. Zipperlen, A. Fraser, and J. Ahringer, "Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*," *Genome Biol*, vol. 2, no. 1, pp. 0002–1, 2001.
- [127] B. Sönnichsen, L. Koski, A. Walsh, P. Marschall, B. Neumann, M. Brehm, A. Alleaume, J. Artelt, P. Bettencourt, E. Cassin, *et al.*, "Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*," *Nature*, vol. 434, no. 7032, pp. 462–469, 2005.

DECLARATION

Hiermit erkläre ich, dass...

ich an keinem zurückliegenden erfolglosen Promotionsverfahren teilgenommen habe. Die vorliegende Arbeit ohne unzulässige Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt wurde; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind in der Arbeit als solche kenntlich gemacht. Die vorliegende Arbeit wurde weder im Inland noch im Ausland in gleicher oder in ähnlicher Form einer anderen Prüfungsbehörde zum Zwecke einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt und auch noch nicht veröffentlicht. Das Manuskript wurde eigenständig erstellt; bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich Unterstützungsleistungen von Professor Dr. Anthony Hyman und Dr. Eugene Myers erhalten; die im Manuskript praesentierten biologischen Daten beruhen auf Experimenten und deren Auswertung, die gemeinsam mit Herrn Markus Decker durchgeführt wurden. Es waren keine weiteren Personen bei der geistigen Herstellung der vorliegenden Arbeit beteiligt, insbesondere habe ich auch nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Es erfolgten keine weder unmittelbar noch mittelbar geldwerten Leistungen im Zusammenhang mit dem Inhalt der vorliegenden Dissertation.

Dresden, September 2010

Steffen Jaensch