Quantitative Comparison of Mitotic Spindles by Confocal Image Analysis

Jeffery R. Price[a,1]
Deniz Aykac[a]
Shaun S. Gleason[a]
Karuna Chourey[b]
Yie Liu[b]

[a]Image Science and Machine Vision Group
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6010
{pricejr,aykacdb,gleasonss}@ornl.gov

[b]Mammalian Genetics Group
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6445
{liuy3,choureyk}@ornl.gov

Abstract

The mitotic spindle is a sub-cellular protein structure that facilitates chromosome segregation and is crucial to cell division. We describe an image analysis approach for comparing mitotic spindles that have been imaged three-dimensionally using confocal microscopy. The proposed approach is based upon a set of features that are computed from each image stack representing a spindle. We compare several spindle data sets that are characterized by varying biological (genotype) and/or environmental (drug treatment) conditions. Experimental results on positive and negative control data indicate that the proposed approach is effective at indicating spindle differences. Furthermore, in two of the experimental comparisons, the results definitively indicate spindle differences of which biologists were previously unable to observe subjectively.

Keywords: image analysis, subcellular imaging, confocal microscopy, fluorescence microscopy, mitotic spindle, gene knockout.

[1]Contact author. Tel: 865-574-5743, Fax: 865-576-8380, Email: pricejr@ornl.gov.

1 Introduction

The current state-of-the-art in imaging is capable of providing biologists with tremendous amounts of data that can easily overwhelm any quantitative analysis capabilities of a human. The research we present in this paper begins to address the need for automated, quantitative processing of image data in an area of important and currently evolving functional genomics research. Data related to functional genomics research is often captured in image and/or volumetric form by a variety of instruments and techniques. Some examples include traditional optical microscopy, atomic force microscopy [1], deconvolution microscopy [2], and laser-scanning confocal microscopy [3]. Presently, much of the analysis of the acquired image data is performed manually and subjectively – a well-trained scientist examines the image and attempts to make qualitative observations concerning its contents relative to what the scientist has seen before. To conduct statistically meaningful experiments, however, it is necessary that many images be produced and analyzed quantitatively, as has been noted for some time [4]. It is, of course, prohibitively time-consuming for a human to manually analyze a large number of images, and virtually impossible for them to do so quantitatively. Hence there is a need for image processing, data analysis, and/or pattern recognition in many image-related, functional genomics applications. The work we present in this paper begins to address that need as it relates to understanding the mitotic spindle.

The mitotic spindle is a crucial structure involved in cell division (mitosis) in eukaryotes (animals, plants, fungi). The spindle is a tubulin-based protein structure, formed during mitosis, that is critical in chromosome segregation, which is the process of splitting duplicated genetic material between two developing daughter cells. A simple illustration of the spindle is shown in Fig. 1. During cell division, spindle microtubules attach to the centers (centromere) of the chromosomes. The chromosomes then split at this point of attachment, forming sister chromatids which are pulled apart by the spindle. The spindle ends, toward which the sister chromatids are pulled, are referred to as the spindle poles. The spindle poles, which are located at opposite ends

of the whole cell, are in fact organelles known as centrioles.

Biologists are particularly interested in the mitotic spindle for several reasons. Abnormal function or disruption of the spindle will lead to cell death, abnormal cell growth, and/or genomic instability because of inaccurate chromosome segregation [5]. Inaccurate chromosome segregation can lead to genomic instability, which is a characteristic common to all cancer cells. Furthermore, many anti-cancer agents (e.g., Taxol or paclitaxel) are actually tubulin poisons which attack the spindle structure to prevent cell division, which proceeds rampantly in cancer cells [6]. Finally, many proteins interact with the spindle in unknown or poorly understood ways. The principal motivation for the image analysis work we present here is to develop tools that will aid cell biologists, in understanding the roles of spindle-associated proteins, and also pharmaceutical companies, in searching for novel tubulin poisons. An improved understanding of these protein roles may also help uncover future anti-cancer drugs and drug targets.

To date, there has been very little work regarding quantitative three-dimensional (3D) or four-dimensional (4D, or 3D plus time-lapse imaging of living cells) imaging of the spindle. Electron tomography [7] has been employed to find the 3D structure of the early spindle in hopes of uncovering structure-function relationships in mutated yeast. Although this approach provides excellent spatial resolution, the noisy images would be challenging to analyze automatically and the required manual efforts limit potential high-throughput capabilities. Courtesy of advances in microscopy [8] and imaging techniques [9], the need for quantitative 3D (and 4D) analysis of the spindle has been clearly demonstrated. In fact, 4D imaging of mitosis has shown that different modes of cell division are linked to 3D rotations and asymmetric orientation of the spindle body in the fruit fly [10], the mouse [11], and in plant cells [12]. 3D live-cell imaging has also been employed to study the effects of mutations on the mitotic spindle in hopes of uncovering protein interactions [13]. Although some previous research has employed quantitative measures [10, 11, 14], more extensive and descriptive measures, such as those we present in this paper, would significantly improve current methodology. Furthermore, there has been no

quantitative, image-based research comparing spindles from different genetic strains and/or under different environmental factors. Additional research [12, 13] has been primarily qualitative.

The remainder of this paper is organized as follows. In Section 2, we briefly discuss the cell biology research that motivates our image analysis work. We then turn our attention to the analysis of the images in Section 3, where we describe image-based features we have developed to characterize the spindle. In Section 4, we describe the results of our approach on several data sets of interest and then we conclude in Section 5 with some summary remarks and notes about our future research goals.

2 Biology Background

As noted in Section 1, there are many proteins associated with the spindle whose roles are poorly understood or completely unknown. One way that biologists investigate the roles of a protein in cells is to produce and study gene-knockout (KO) organisms, where "knockout" means that a gene is removed from the organism. Since specific proteins are encoded by corresponding genes, the role of a protein can be elucidated by examining the phenotype of the organism deficient in that protein due to gene-KO.

The mouse has been employed as the model organism in this paper for several reasons. The mouse is the best-characterized experimental mammal and has the most well-defined genetics of any non-human vertebrate. Extensive homologies between mouse and human genomes have been established, thus permitting the mouse to be used as a model system for many human genetic diseases. Finally, both the genome and the embryo of the mouse can be manipulated with relative ease, allowing for an expansion of the genetic resources necessary for functional interpretation of genes and DNA sequences.

The cells used for all imaging studies are known as mouse embryonic fibroblasts, or MEFs. *Primary* and *immortalized* MEFs derived from wild-type (normal) or one of two different knock-

out mice are employed. Primary MEFs can only divide in cell culture for a limited period of time. Immortalized cells, on the other hand, have undergone a random mutation (or mutations) and hence gain the capability to continue to divide in culture indefinitely. The two knockout or mutant MEFs employed herein are deficient for either the *p53* gene or the *Vparp*[2] gene. P53 is a tumor suppressor and its function has been linked to a variety of cellular processes, such as cell cycle control, cell apoptosis (programmed cell death), and transcription activation [15, 16, 17]. *p53*-deficient MEFs also exhibit a significantly higher probability of becoming immortalized relative to wild-type MEFs. VPARP – vault poly(ADP-ribose) polymerase – was originally identified as a minor protein component of the vault ribonucleoprotein (RNP) particle [18], whose function may be involved in molecular assembly or subcellular transport [19, 20, 21]. In addition to its association with cytoplasmic vault particles, sub-populations of VPARP localize to the nucleus and the mitotic spindle, indicating VPARP may have other cellular functions [18, 22]. In this study, we are interested to determine if *Vparp* deficiency in mammalian cells can lead to changes in spindle structure.

As a positive control for abnormal spindle characteristics, we treated some immortalized cells (see Section 4) with the tubulin poison Taxol (also known as paclitaxel), which is a microtubule depolymerization inhibitor. To function properly, microtubules must be in a dynamic state of polymerization and depolymerization [6]. Since the mitotic spindle is composed of microtubules, treatment with Taxol will certainly affect structural changes. Taxol is also widely used as an anti-cancer drug because of its effect on the spindle.

For imaging, fixed-cell (i.e., non-living) MEFs are employed. These fixed cells were probed with fluorescent, anti-tubulin antibodies that bind to the tubulin structure of the MEF spindles. The cells were then imaged with a confocal laser scanning microscope to produce 3D image-stacks. An example image, projected to 2D by summing along the optical axis, is shown in Fig. 2.

[2]When italicized, *p53* and *Vparp* refer to genes. When given with capital letters, P53 and VPARP refer to the proteins encoded by those genes.

3 Image Analysis

In this section, we describe 49 features that are computed for each spindle image stack. These features are used for comparing spindles to one another, as we will describe in Section 4. The first few features, described in Section 3.1, are independent of the spindle shape and are related purely to the distribution of gray-level values in each stack. The remaining features, described in 3.2, are all related to the spindle morphology.

We note that oftentimes a single image-stack will contain the spindle in which we are interested as well as microtubules from other cells nearby. To alleviate the contribution of this background, we have manually selected a polygonal region-of-interest (ROI) for each image stack and all processing done below refers only to the voxels in the ROI. At present, with respect to the time required to prepare cells and acquire images, this manual processing is not a limitation. As we intend to move toward high-throughput in the future, automatically detecting and isolating the spindle may become a future topic of research.

3.1 Gray-level Features

The gray-levels in our data are proportional to the fluorescent signal intensity in a voxel and therefore represent the concentration of tubulin in that voxel. Since our images have been captured under controlled and consistent conditions, basic gray-level features may indicate variations in tubulin levels of distribution that are of interest to the biologist. Of the 49 features in our complete set, we compute eight such gray level features.

F01. ***Mean.*** The mean gray value, μ, of all voxels. A higher mean value may indicate a higher concentration of tubulin in the mitotic spindle.

F02. ***Standard deviation.*** The standard deviation, σ, of all voxels. A high standard deviation may indicate a very irregular distribution of tubulin in the mitotic spindle.

F03. ***Histogram skewness.*** The skewness, ψ, is a measure of the asymmetry of a distribu-

tion and is given by

$$\psi = \frac{m_3}{m_2^{3/2}}, \tag{1}$$

where m_q is the qth central moment. For our discrete distributions, where we have bins given by $b(i)$ and bin masses or probabilities given by $p(i)$, the central moments are computed according to

$$m_q = \sum_n \big(b(i) - \mu\big)^q p(i) \tag{2}$$

where μ represents the distribution mean and is given simply by $\mu = \sum_i b(i)p(i)$. For the histogram, $b(i) \in [0, 255]$ represents the 8-bit gray value and $p(i)$ is the number of voxels of gray level $b(i)$ divided by the total number of voxels. Characteristics of the histogram, as measured by this feature and F04-F08, may indicate more subtle variations in tubulin distribution than captured by features F01 and F02 alone.

F04. ***Histogram kurtosis excess.*** The kurtosis excess is a measure of how peaked a distribution is and is given by

$$\kappa = \frac{m_4}{m_2^2} - 3. \tag{3}$$

With this definition, the normal (Gaussian) distribution has $\kappa = 0$ while an exponential distribution has $\kappa = 6$.

F05-08. ***Histogram PCA coefficients.*** We represent each histogram as a 256-point vector and then perform principal component analysis (PCA) [23] on the collection of these vectors. We keep as features the PCA coefficients corresponding to the first four principal components, which represent the four most significant shape variations in the histograms. The first four PCA coefficients represent "how much" of the corresponding component is in a given histogram. In Fig. 3 we show the mean histogram and the first two principal components.

3.2 Morphology Dependent Features

We now turn our attention to features that are related to the mitotic spindle morphology. Each of the features described in this section relate to the shape of the spindle as well as the 3D tubulin distribution in the spindle. Since the spindles are in fact 3D objects, and we cannot guarantee that they are all imaged in the same pose, we must first establish an internal coordinate system for each spindle stack. We begin by letting each voxel in the image stack represent a point mass with mass equal to the voxel's value. For each spindle image-stack, we compute the center of mass and the principal axes of inertia [24, 25] to serve as the origin and axes, respectively, of a spindle-centric coordinate system.

The principal axes of inertia represent mutually orthogonal axes about which a rigid body can be dynamically balanced, meaning its angular velocity is parallel to its angular momentum. A rigid body that is imparted with an angular velocity in the direction of (i.e., made to spin about) a principal axis will continue to spin. The principal axes are given by the eigenvectors of the moment of inertia matrix

$$\mathrm{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{pmatrix}, \tag{4}$$

where the elements are computed as follows. Let $v(x_o, y_o, z_o)$ represent a voxel value in the image stack, (x_o, y_o, z_o) the physical coordinates in the original coordinate system, $(\bar{x}, \bar{y}, \bar{z})$ the center of mass relative to the original coordinate system, and $\mathcal{M}$ the total spindle mass, then the elements of I are given by

$$I_{xx} = \Big(\sum_{x_o}\sum_{y_o}\sum_{z_o} v(x_o, y_o, z_o)\big(y_o^2 + z_o^2\big)\Big) - \mathcal{M}(\bar{y}^2 + \bar{z}^2), \tag{5}$$

$$I_{yy} = \Big(\sum_{x_o}\sum_{y_o}\sum_{z_o} v(x_o, y_o, z_o)\big(x_o^2 + z_o^2\big)\Big) - \mathcal{M}(\bar{x}^2 + \bar{z}^2), \tag{6}$$

$$I_{zz} = \Big(\sum_{x_o}\sum_{y_o}\sum_{z_o} v(x_o, y_o, z_o)\big(x_o^2 + y_o^2\big)\Big) - \mathcal{M}(\bar{x}^2 + \bar{y}^2), \tag{7}$$

$$I_{xy} = -\Big(\sum_{x_o}\sum_{y_o}\sum_{z_o} x_o y_o\, v(x_o, y_o, z_o)\Big) + \mathcal{M}\bar{x}\bar{y}, \tag{8}$$

$$I_{xz} = -\Big(\sum_{x_o}\sum_{y_o}\sum_{z_o} x_o z_o\, v(x_o, y_o, z_o)\Big) + \mathcal{M}\bar{x}\bar{z}, \quad \text{and} \tag{9}$$

$$I_{yz} = -\Big(\sum_{x_o}\sum_{y_o}\sum_{z_o} y_o z_o\, v(x_o, y_o, z_o)\Big) + \mathcal{M}\bar{y}\bar{z}. \tag{10}$$

We then compute the eigenvalues and eigenvectors of I and construct a new reference frame whose origin is at the center of mass and whose three axes are in the direction of the eigenvectors (i.e., the principal axes). The eigenvalues are known as the principal moments of inertia and we sort the axes in the new reference frame, (x, y, z), in order of descending eigenvalues. Hence, a position (x_o, y_o, z_o) in the original coordinate system is transformed into (x, y, z) in the spindle-centric coordinate system by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{pmatrix} \begin{pmatrix} x_o - \bar{x} \\ y_o - \bar{y} \\ z_o - \bar{z} \end{pmatrix} \tag{11}$$

where the 3-vector $\mathbf{a}_i$ represents principal axes i, given as coordinates in the original reference frame, sorted by descending eigenvalue. In Fig. 4, we show an example spindle with its spindle-centric coordinate system as determined by the principal axes. For each spindle, we transform each voxel's coordinates into this spindle-centric coordinate system and then compute the shape-based features described below.

F09-F12. ***Axis 1 mass percentiles.*** These features are intervals, in microns, along the first principal axis, that enclose a certain percentage of the spindle mass and are computed as follows. Define N bins, X_k, along the first principal axis (i.e., Axis 1) with centers at

$$x_k = x_{min} + \frac{k-1}{N-1}\big(x_{max} - x_{min}\big), \tag{12}$$

for $k = 1, \ldots, N$ and with bin width

$$\epsilon = \frac{x_{max} - x_{min}}{N - 1}. \tag{13}$$

The kth bin, X_k, is given by the interval $X_k = (x_k - \epsilon/2, x_k + \epsilon/2]$. In our experiments, described later in Section 4, we use $N = 200$ and x_{min} and x_{max} are set to $-10\mu m$ and $+10\mu m$, respectively. We now compute the projection of spindle mass onto Axis 1, $p_1(x_k)$, as follows:

$$p_1(x_k) = \frac{1}{\mathcal{M}} \sum_{x \in X_k} \sum_{y} \sum_{z} v(x, y, z), \tag{14}$$

where $\mathcal{M}$ represents the total spindle mass. An example of $p_1(x_k)$, computed for the spindle of Fig. 4, is shown in Fig. 5. An Axis 1 mass percentile, $\Delta_1(f)$, is the minimum-width interval about $x = 0$ in the spindle-centric coordinate system that contains the fraction f of the total mass and is computed as follows:

$$\Delta_1(f) \;=\; \arg\min_{\Delta} \Big(\sum_{x_k \in [-\frac{\Delta}{2}, \frac{\Delta}{2}]} p_1(x_k) \Big) > f. \tag{15}$$

We compute and store $\Delta_1(f)$ for $f \in \{0.25, 0.50, 0.75, 0.95\}$. These features represent the spread of the spindle mass along the first principal axis.

F13-F16. *Axis 1 mass PCA coefficients.* We represent the mass projected onto Axis 1, $p_1(x_k)$ as N-vectors and perform PCA on the collection of these vectors. We keep as features the four coefficients corresponding to the first four principal components. In Fig. 6 we show the mean projection of mass onto Axis 1 and the first two principal components.

F17. *Axis 1 mass skewness.* The skewness, ψ_1 of the mass projected onto Axis 1, ψ_1, using Eqs. (1) and (2) with $p_1(x_k)$ representing the distribution and x_k the bins.

F18. *Axis 1 mass kurtosis excess.* The kurtosis excess, κ_1, of the mass projected onto

Axis 1, $p_1(x_k)$, using Eqs. (3) and (2).

F19-22. ***Axis 2 mass percentiles.*** These features are computed exactly the same as the Axis 1 mass percentiles (F09-F12), except that we now project along the second principal axis represented as the y direction in the spindle-centric coordinate system. An example of the spindle mass projected onto Axis 2, $p_2(y_k)$, is shown in Fig. 7. We keep $\Delta_2(f)$, for $f \in \{0.25, 0.50, 0.75, 0.95\}$. These features represent the spread of the spindle mass along the second principal axis.

F23-26. ***Axis 2 mass PCA coefficients.*** First four coefficients from PCA on the collection of $p_2(y_k)$. In Fig. 8 we show the mean projection of mass onto Axis 2 and the first two principal components.

F27. ***Axis 2 mass skewness.*** We compute the skewness, ψ_2, of the mass projected onto Axis 2, $p_2(y_k)$, using Eqs. (1) and (2).

F28. ***Axis 2 mass kurtosis excess.*** We compute the kurtosis excess, κ_2, of the mass projected onto Axis 2, $p_2(y_k)$, using Eqs. (3) and (2).

F29-32. ***Axis 3 mass percentiles.*** These features are computed exactly the same as the Axis 1 mass percentiles (F09-F12), except that we now project along the third principal axis which is represented as z in the spindle-centric coordinate system. An example of the spindle mass projected onto Axis 3, $p_3(z_k)$, is shown in Fig. 9. We keep $\Delta_3(f)$, for $f \in \{0.25, 0.50, 0.75, 0.95\}$. These features represent the spread of the spindle mass along the third principal axis.

Note that Axis 3, which corresponds to the lowest principal moment of inertia, is oriented in the pole-to-pole direction of the spindle. Since tubulin concentrations tend to be higher at the poles, we can see in Fig. 9 that $p_3(z_k)$ tends to be bimodal, where the two peaks correspond to the poles. Since we are interested in variations in spindle morphology that may be asymmetric, we orient all the Axis 3 mass projections, $p_3(z_k)$, so that the larger peak is to the left (i.e., the positive side of Axis 3 points towards the lower concentration peak). Because of this bimodality, we compute kurtosis excess only for the entire curve, but skewness and kurtosis excess for each

side of the curve individually.

F33-36. ***Axis 3 mass PCA coefficients.*** First four coefficients from PCA on the collection of $p_3(z_k)$. In Fig. 10 we show the mean projection of mass onto Axis 3 and the first two components.

F37. ***Axis 3 mass kurtosis excess.*** The kurtosis excess, κ_3, of the mass projected onto Axis 3, $p_3(z_k)$, using Eqs. (3) and (2).

F38. ***Axis 3 mass left-side skewness.*** The skewness, ψ_{3l}, of $p_3(z_k)$ for $k = 1, \ldots, N/2$.

F39. ***Axis 3 mass left-side kurtosis excess.*** The kurtosis excess, κ_{3l}, of $p_3(z_k)$ for $k = 1, \ldots, N/2$.

F40. ***Axis 3 mass right-side skewness.*** The skewness, ψ_{3r}, of $p_3(z_k)$ for $k = N/2 + 1, \ldots, N$.

F41. ***Axis 3 mass right-side kurtosis.*** The kurtosis excess, κ_{3r}, of $p_3(z_k)$ for $k = N/2 + 1, \ldots, N$.

F42-45. ***Radial mass percentiles about Axis 3.*** With the final two sets of features (F42-45 and F46-49), we consider the radial distributions of spindle mass about Axis 3. To achieve rotational invariance about Axis 3, we represent $v(x, y, z)$ as $v(r, z)$ where $r = \sqrt{x^2 + y^2}$. We construct P radial bins, R_i for $i = 1, \ldots, P$, centered on

$$r_i = \frac{i-1}{P-1} r_{max} \tag{16}$$

and defined by $R_i = [r_i, r_{i+1})$. For features F42-45, we use $P = 200$ and $r_{max} = 10\mu m$. Similar to the computation for mass projected onto Axis 1 (recall F09-12), we can then compute the radial distribution of the mass (about Axis 3), for the entire spindle, $p_r(r_i)$, as follows:

$$p_r(r_i) = \frac{1}{\mathcal{M}} \sum_{r \in R_i} \sum_{z} v(r, z), \tag{17}$$

where again $\mathcal{M}$ represents the total spindle mass. The radial mass percentiles, $\Delta_r(f)$, are then given by

$$\Delta_r(f) \;=\; \arg\min_{\Delta} \Big(\sum_{r_i \leq \Delta} p_r(r_i) > f \Big), \tag{18}$$

In other words, $\Delta_r(f)$ is just the minimal radius enclosing the fraction f of the entire mass. We compute and store $\Delta_r(f)$ for $f \in \{0.25, 0.50, 0.75, 0.95\}$

F46-49. ***Radial-Axis 3 mass PCA coefficients.*** In F42-45, we computed the radial distribution of the total spindle mass about Axis 3. With features F46-49, we consider a similar radial distribution, but in "slices" that are parallel to Axes 1 and 2 and placed along Axis 3. We define radial bins again just as we did for F42-45 in Eq. (16). We also define bins, Z_j, along the third principal axis that are centered on

$$z_j = z_{min} + \frac{j-1}{Q-1}\big(z_{max} - z_{min}\big), \tag{19}$$

with bin width

$$\delta = \frac{z_{max} - z_{min}}{Q-1} \tag{20}$$

and given by $Z_j = (z_j - \delta/2, z_j + \delta/2]$. With these definitions, we define the radial-Axis 3 mass distribution to be

$$p_{r3}(r_i, z_j) = \frac{1}{\mathcal{M}} \sum_{r_i \in R_i} \sum_{z_j \in Z_j} v(r, z). \tag{21}$$

Here we use $P = 50$, $Q = 50$, $r_{max} = 10\mu m$, $z_{min} = -10\mu m$, and $z_{max} = +10\mu m$. This implies that the mass distribution $p_{r3}(r_i, z_j)$ is a two-dimensional, 50×50 array. In Fig. 11 we show an example of p_{r3} as a surface plot. The two peaks correspond to the spindle poles. We raster scan the 50×50 array making a 2500-point vector and then perform PCA on the complete data from all spindles, keeping as features the first four PCA coefficients.

4 Experimental Results

In this section, we present and discuss results of comparing the features from Section 3 for several data sets. The acquired experimental data comprised 44 image stacks from six cell types – the first three types are primary MEFs and the last three are immortalized MEFs – as described below.

1. Wild-type (WT). Primary MEFs from normal (non-mutant) mice. Eight image stacks.
2. VPARP-knockout (VPARP-KO). Primary MEFs, VPARP encoding gene is removed. Eight image stacks.
3. P53-knockout (P53-KO). Primary MEFs, P53 encoding gene is removed. Four image stacks.
4. Immortalized (IM). Immortalized MEFs derived from primary P53-KO MEFs. Eight image stacks.
5. Immortalized, Taxol treated at low concentration (IM-TaxLo). Immortalized MEFs derived from primary P53-KO MEFs, treated with 0.01μM Taxol for 24 hours. This serves as a positive control since the anti-cancer drug Taxol is a tubulin poison and is known to give rise to structural effects on the spindle. Eight image stacks.
6. Immortalized, Taxol treated at a higher concentration (IM-TaxHi). Immortalized MEFs derived from primary P53-KO MEFs, treated with 0.1μM Taxol for 24 hours. Eight image stacks.

The image stacks contained voxels of approximately $0.03\mu m \times 0.03\mu m \times 0.08\mu m$ in size. Each image was 1024×1024 pixels and the stacks ranged from about 40 images deep to 100 images deep, depending on the size and/or pose of the imaged spindle. The 49 features described in Section 3 were computed for each of the 40 spindle image stacks.

We then performed six different pair-wise comparisons of the data sets, including both positive control and negative control comparisons. For the comparisons, we used the two-sample t statistic [26] to determine if the individual feature means differed significantly between the data sets. We note as significant any features whose means were statistically different at $P < 0.03$. For the t statistic, the P-value represents the likelihood that the differences in the two means would be observed purely by chance. In other words, any observed differences in the feature means at $P < 0.03$ would occur less than 3% of the time randomly. In the context of the 49 features that we compare, we expect that one or two features will almost always appear different at $P < 0.03$ since $0.03 \times 49 = 1.47$. We discuss the four pair-wise comparisons in Sections 4.1 4.4 below.

4.1 Comparison 1: IM vs. IM-TaxLo (Positive Control)

Here we compared immortalized (IM) MEFs to IM MEFs that were treated with 0.01μM Taxol for 24 hours. We found that 10 of the 49 features differed at $P < 0.03$. We note that this comparison represents positive control data since the treatment with tubulin poison (Taxol) is known to affect the spindle structure, often in ways that are visually apparent to a biologist. We therefore expect to see statistically significant differences in several of our features. Since we found 10 such features, which is significantly more than the one or two we would expect randomly, we conclude that some of the differences in the treated vs. untreated spindles are indeed captured by the features we have implemented.

The 10 features that differed were F05, F28-34, F46, and F48. In Fig. 12, we show a scatter plot of features F28 and F30 for this comparison. F28 is the kurtosis excess of the Axis 2 projected mass and F30 is the interval that contains 50% of the total spindle mass projected onto Axis 3 (see Section 3.2). As we can see in Fig. 12, F28 for the IM-TaxLo data tends to be lower, indicating less peakedness than the IM data. This implies that the tubulin concentration along Axis 2 is slightly lower near the spindle center in the Taxol treated cells. This indeed can

be seen in Fig. 13, where we have plotted the mean Axis 2 mass projections for both the IM and the IM-TaxLo data. Regarding F30 ($\Delta_3(0.5)$), it is evident in Fig. 12 that it tends to be lower for the Taxol treated cells, indicating that these spindles are shorter along Axis 3 (i.e., more concentrated toward the center) than the untreated cells. This can be seen in Fig. 14, where we have plotted the mean Axis 3 mass projections for both the IM and the IM-TaxLo data. The shorter spindles reflect the consequences of the microtubules' inability to undergo polymerization/depolymerization due to the Taxol treatment.

4.2 Comparison 2: IM-TaxLo vs. IM-TaxHi

In this comparison, we were interested to see if the higher concentration of Taxol produced any significant spindle changes relative to the cells treated with a lower concentration of Taxol. Here we found only one feature that was different at $P < 0.03$, indicating that, according to our feature set, there are no statistically significant differences between the cells treated with low or high concentration of Taxol. In other words, at the Taxol concentrations tested, approximately equivalent changes were made to the spindle.

4.3 Comparison 3: WT vs. IM

Here we compared normal MEFs to immortalized p53-KO MEFs. It has been shown that the absence of the p53 tumor suppressor gene can lead to spindles with multiple poles (as opposed to the normal two), which is a defect that is readily apparent to a human observer in spindle imagery (and can lead to cancer) [15]. Although multi-polar spindles were observed in the preparations, only bipolar IM spindle images were acquired for comparison. In other words, all of the IM spindle images we have used appear visually normal to a biologist.

In WT vs. IM comparison, we found 26 features that differed at $P < 0.03$. These features were F01-F05, F09-F14, F22, F24, F28-F33, F37-F41, F46, F47. Obviously the number of differing

features seems to indicate significant differences between the normal (WT) and the IM cells. We note that this result is important since abnormalities in normal-appearing, bipolar immortalized spindles have not been observed or reported to date. Several features related to the projection onto Axis 3 are significantly different (F29-33, F37-41). We show in Fig. 15 the mean projections of mass onto Axis 3 for the wild-type and IM cells. We can see from Fig. 15 that the poles of the IM spindles – represented by the two peaks in each curve – tend to be further apart and the IM spindles tend to have less tubulin, relative to the poles, in their center.

4.4 Comparison 4: WT vs. VPARP-KO

In this comparison, we were interested to see if the computed features indicated any differences between normal (WT) spindles and VPARP-KO spindles in hopes of determining if VPARP has a role at the spindle. For the WT vs. VPARP-KO comparison, we found six features that differed at $P < 0.03$; these were F01, F02, F06, F07, F18, and F28. Since six features is more than the one or two we would expect to see purely by chance, this comparison seems to indicate that the VPARP-KO spindles are indeed different from the WT cells, but that the differences are subtle. The difference of six features, however, is indeed significant when we consider the negative control experiment discussed below in Section 4.6. We note that one of the differing features is F28, which is the kurtosis excess of the mass projected onto Axis 2. In Fig. 16, we show the mean Axis 2 projected mass of the wild-type cells and the VPARP knockout cells. It is apparent that the VPARP-KO cells are indeed "flatter" about the center of mass. The results of this comparison may indicate a structural variation in the spindle, related to VPARP deficiency, that warrants further biological experimentation.

4.5 Comparison 5: WT vs. p53-KO

Here we compared wild-type cells to p53-KO primary (i.e., not immortalized) cells. We compared the four p53-KO image stacks to every possible combination of four WT samples selected from the set of eight. For each of the 70 possible combinations, we performed the pair-wise feature comparison as discussed above and recorded the number of features differing with $P < 0.03$. The results indicated that the mean number of differing features was 16.46 and the median number of differing features was 16. A histogram of the number of differing features is shown in Fig. 17, where we see the fewest number of differing features was 13 in four of the 70 cases. Twelve features differed at $P < 0.03$ in all 70 combinations; these were F20-F23, F26, F29-F33, F45, and F46. Many of these features are associated with the projections onto Axis 2 and Axis 3. We show in Fig. 18 and Fig. 19 the means of the projections onto Axis 2 and Axis 3, respectively, for the WT and p53-KO cells. It is evident that the p53-KO spindles appear to spread across a larger area.

4.6 Comparison 6: WT vs. WT (Negative Control)

In this final, negative control comparison, we compared wild-type MEFs to wild-type MEFs. Since the cell types are the same, we expect there to be no significant differences in the features. We selected all possible subdivisions of the eight WT image stacks into two sets of four image stacks each; this results in 35 possible subdivisions. For each of the 35 possible combinations, we performed the pair-wise feature comparison as discussed above and recorded the number of differing features. The results indicated that the mean number of differing features was 1.2 and the median number of differing features was in fact zero, as predicted. These results also indicate that the differences we have detected in the other comparisons are indeed significant, especially in the case of the WT vs. VPARP-KO comparison where only six features were found to differ with $P < 0.03$.

5 Conclusions

In this paper, we have described an image analysis approach to detect quantitative differences in mitotic spindles that have been imaged, in 3D, using confocal microscopy. This approach is based upon a set of features or numerical descriptors that are computed from each 3D image stack representing a spindle. Using these features, we compared several data sets that were characterized by different biological (genotype) and/or environmental (drug treatment) conditions. The experimental results on positive- and negative-control data (i.e., where we knew there should be significant differences or no differences) indicate that the proposed approach is indeed effective at detecting differences in spindle properties. Furthermore, in two of the experimental comparisons, where mutant MEFs were compared to wild-type (normal) MEFs, the results indicate spindle differences that were previously unobserved by biologists.

Finally, we note some directions of future research. The work presented in this paper was applied only to fixed-cell (i.e., non-living) preparations. The mitotic spindle, however, is a dynamic structure. In future efforts, we will be imaging and analyzing spindles in living cells, where dynamic variations may also be uncovered. We also intend to apply the image analysis approach to additional gene-knockout models that are deficient in other spindle-associated proteins of interest. Finally, as a long-term goal, we plan to incorporate the methods developed here into a software tool to be used by cell biologists in their own laboratories and to provide for data exchange between different research groups.

Acknowledgments

The authors would like to acknowledge Dr. Lea Harrington and Bryan Snow, who directed the generation of a VPARP-deficient mouse model, and Dr. Atsushi Hirao for providing *p53* deficient MEFs. We would also like to thank Drs. Mark Ungrin, Philip Bingham, and Ken Tobin for their

thorough reviews of this manuscript. Research sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory (ORNL), managed by UT-Battelle, LLC for the U. S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

[1] J. H. Hoh and P. K. Hansma, "Atomic force microscopy for high-resolution imaging in cell biology," *Trends in Cell Biology* **2**, pp. 208–213, July 1992.

[2] J. G. McNally, T. Karpova, J. Cooper, and J. A. Conchello, "Three-dimensional imaging by deconvolution microscopy," *Methods* **19**(3), pp. 373–385, 1999.

[3] G. Kino and T. Corle, *Confocal Scanning Optical Microscopy and Related Imaging Systems*, Academic Press, 1996.

[4] A. Diaspro, F. Beltrame, M. Fato, and P. Ramoino, "Characterizing biostructures and cellular events in 2D/3D," *IEEE Engineering in Medicine and Biology* **15**(1), pp. 92–100, 1996.

[5] A. Kramer, K. Neben, and A. Ho, "Centrosome replication, genomic instability and cancer," *Leukemia* **16**(5), pp. 767–775, 2002.

[6] K. Wood, W. Cornwell, and J. Jackson, "Past and future of the mitotic spindle as an oncology target," *Current Opinion in Pharmacology* **1**, pp. 370–377, 2001.

[7] E. O'Toole, M. Winey, and J. McIntosh, "High-voltage electron tomography of spindle pole bodies and early mitotic spindles in the yeast Saccharomyces cerevisiae," *Molecular Biology of the Cell* **20**, pp. 2017–2031, 1999.

[8] D. Stephens and V. Allan, "Light microscopy techniques for live cell imaging," *Science* **300**(5616), pp. 82–86, 2003.

[9] C. Rieder and A. Khodjakov, "Mitosis through the microscope: advances in seeing inside live dividing cells," *Science* **300**(5616), pp. 91–96, 2003.

[10] J. Kaltschmidt, C. Davidson, N. Brown, and A. Brand, "Rotation and asymmetry of the mitotic spindle direct asymmetric cell division in the developing central nervous system," *Nature Cell Biology* **2**(1), pp. 7–12, 2000.

[11] T. Haydar, E. Ang, and P. Rakic, "Mitotic spindle rotation and mode of cell division in the developing telencephalon," *Proceedings of the National Academy of Sciences* **100**(5), pp. 2890–2895, 2003.

[12] S. Cutler and D. Erhart, "Polarized cytokinesis in vacuolate cell of Arabidopsis," *Proceedings of the National Academy of Sciences* **99**(5), pp. 2812–2817, 2002.

[13] J. Paluh, E. Nogales, B. Oakley, K. McDonald, A. Pidoux, and W. Cande, "A mutation in gamma-tubulin alters microtubule dynamics and organization and is synthetically lethal with the kinesin-like protein pkl1p," *Molecular Biology of the Cell* **11**(4), pp. 1225–1239, 2000.

[14] J. Swedlow, K. Hu, P. Andrews, D. Roos, and J. Murray, "Measuring tubulin content in Toxoplasma gondii: a comparison of laser-scanning confocal and wide-field fluorescence microscopy," *Proceedings of the National Academy of Sciences* **99**(4), pp. 2014–2019, 2002.

[15] K. Fukasawa, T. Choi, R. Kuriyama, S. Rulong, and G. F. Vande Woude, "Abnormal centrosome amplification in the absence of p53," *Science* **271**, pp. 1744–1747, March 1996.

[16] J. Fridman and S. Lowe, "Control of apoptosis by p53," *Oncogene* **22**(56), pp. 9030–9040, 2003.

[17] K. Vermeulen, Z. Berneman, and D. Van Bockstaele, "Cell cycle and apoptosis," *Cell Proliferation* **36**(3), pp. 165–175, 2003.

[18] V. Kickhoefer, A. Siva, N. Kedersha, E. Inman, C. Ruland, M. Streuli, and L. Rome, "The 193-kD vault protein, VPARP, is a novel poly(ADP-ribose) polymerase," *Journal of Cell Biology* **146**(5), pp. 917–928, 1999.

[19] M. Mossink, A. van Zon, R. Scheper, P. Sonneveld, and E. Wiemer, "Vault: a ribonucleoprotein particle involved in drug resistance?," *Oncogene* **22**(47), pp. 7458–7467, 2003.

[20] K. Suprenant, "Vault ribonucleoprotein particles: sarcophagi, gondolas, or safety deposit boxes?," *Biochemistry* **41**(49), pp. 14447–14454, 2002.

[21] A. van Zon, M. Mossink, R. Scheper, P. Sonneveld, and E. Wiemer, "The vault complex," *Cellular and Molecular Life Sciences* **60**(9), pp. 1828–1837, 2003.

[22] Y. Liu, B. Snow, V. Kickhoefer, N. Erdmann, W. Zhou, A. Wakeham, M. Gomez, L. Rome, and L. Harrington, "A vault protein, VPARP, is associated with mammalian telomerase and is dispensable in telomerase function and vault structure in vivo," *Molecular and Cell Biology* , 2004. in press.

[23] K. Fukunaga, *Statistical Pattern Recognition*, Morgan Kaufmann, 1990.

[24] W. Flygare, *Molecular Structure and Dynamics*, Prentice-Hall, 1978.

[25] J. Foote and A. Raman, "A relation between the principal axed of inertia and ligand binding," *Proceedings of the National Academy of Sciences* **97**, pp. 978–983, February 2000.

[26] D. S. Moore and G. P. McCabe, *Introduction to the Practice of Statistics*, Freeman, 3rd ed., 1999.

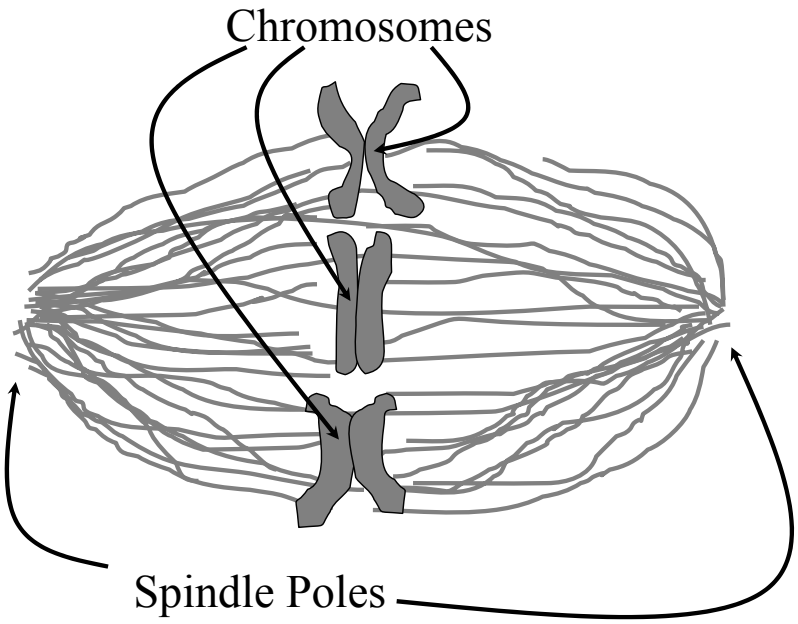

Figure 1: A simple illustration of the mitotic spindle. The spindle emanates from the spindle poles and facilitate chromosome segregation. During mitosis (cell division) the separated chromosomes are pulled by the spindle toward the poles into two forming daughter cells.

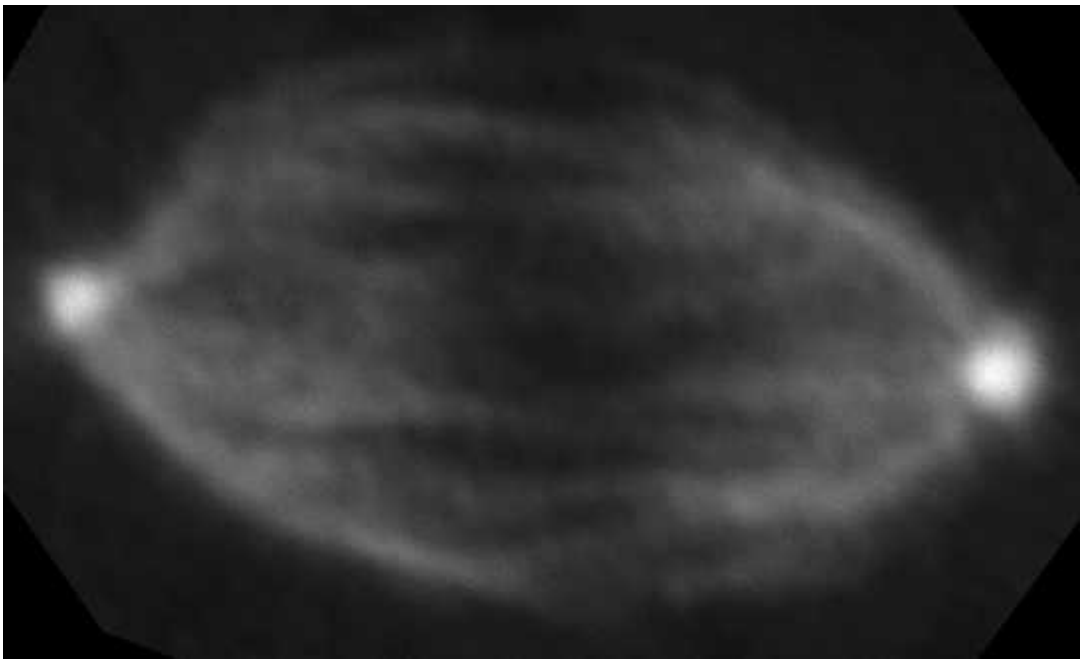

Figure 2: An example spindle image. This image was produced by summing along images in the stack along the z_o direction and has been contrast-enhanced for print reproduction. Note that only the fluorescent-stained tubulin is visible.

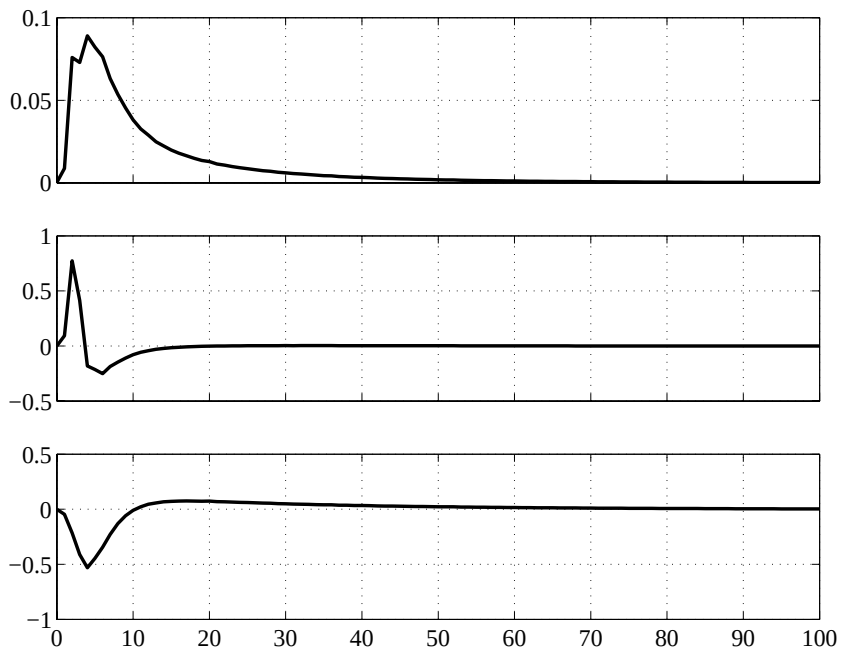

Figure 3: Principal component analysis of the histograms. From top to bottom, the curves represent the mean histogram and first two principal components, respectively.

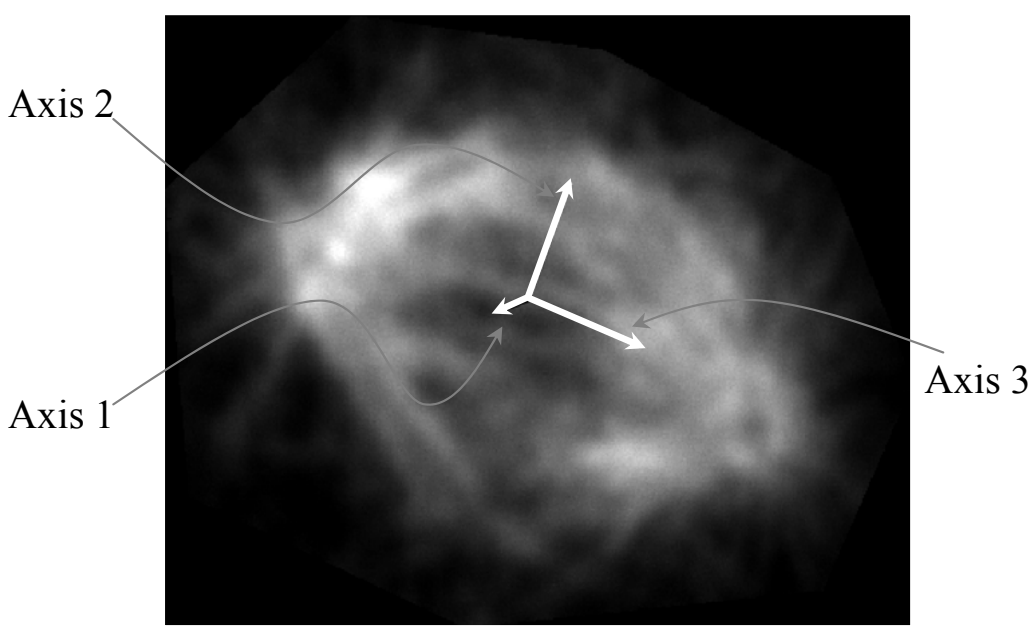

Figure 4: An example spindle shown with it principal axes. Again, this image was produced by projecting along the z_o direction and has been contrast-enhanced for print reproduction.

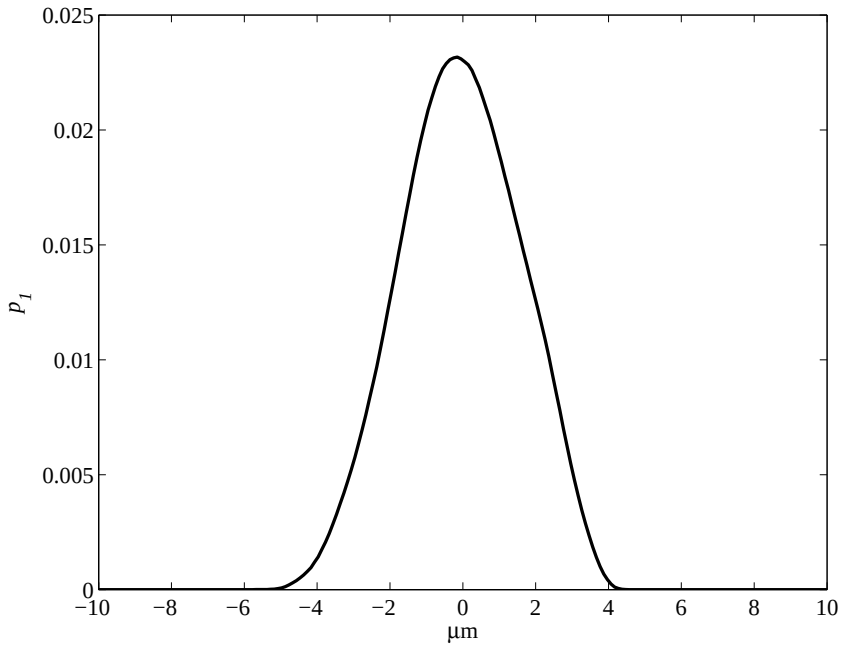

Figure 5: Mass projected onto the first principal axes, $p_1(x_k)$, for the spindle of Fig. 4.

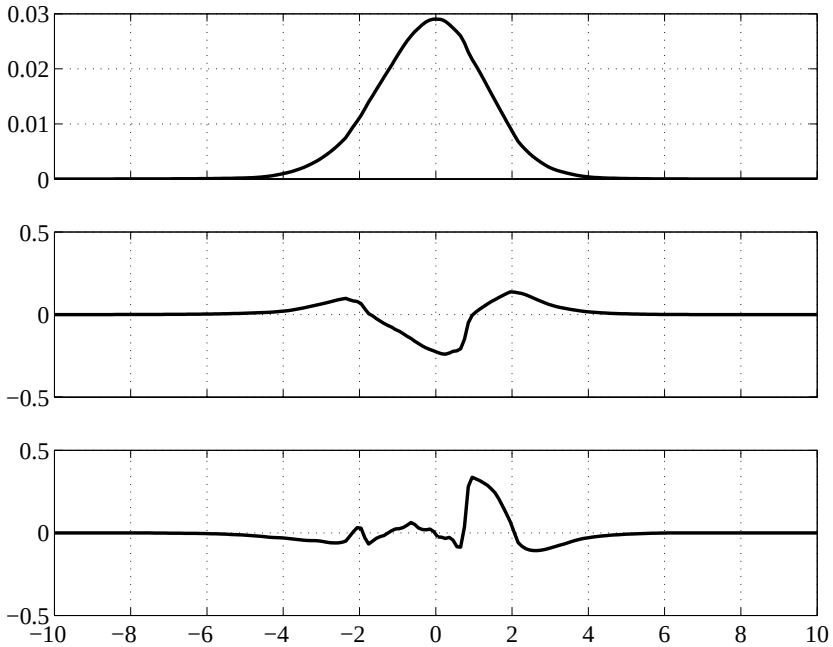

Figure 6: Mean and first two principal components, from top-to-bottom respectively, of mass projected onto the first principal axis.

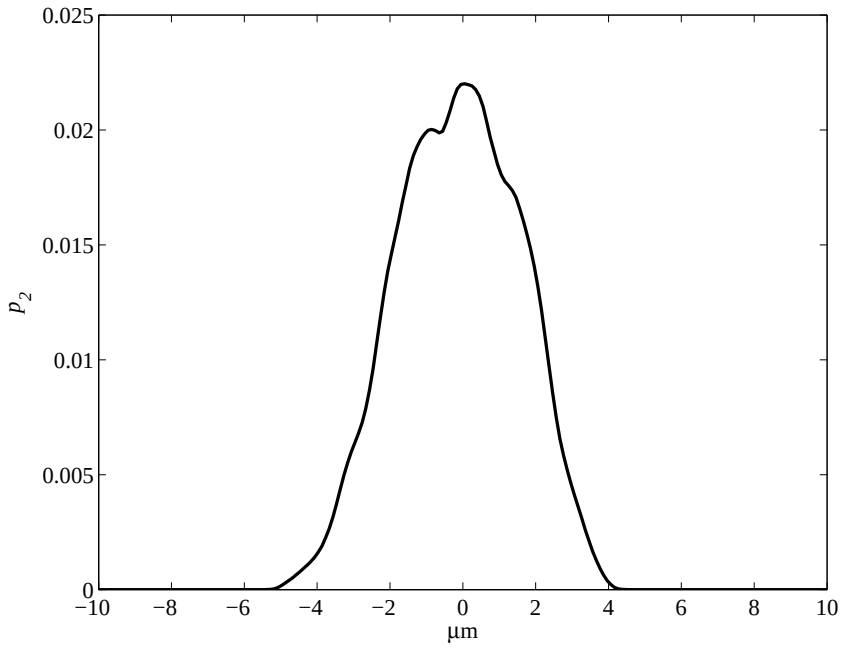

Figure 7: Mass projected onto the second principal axes, $p_2(y_k)$, for the spindle of Fig. 4.

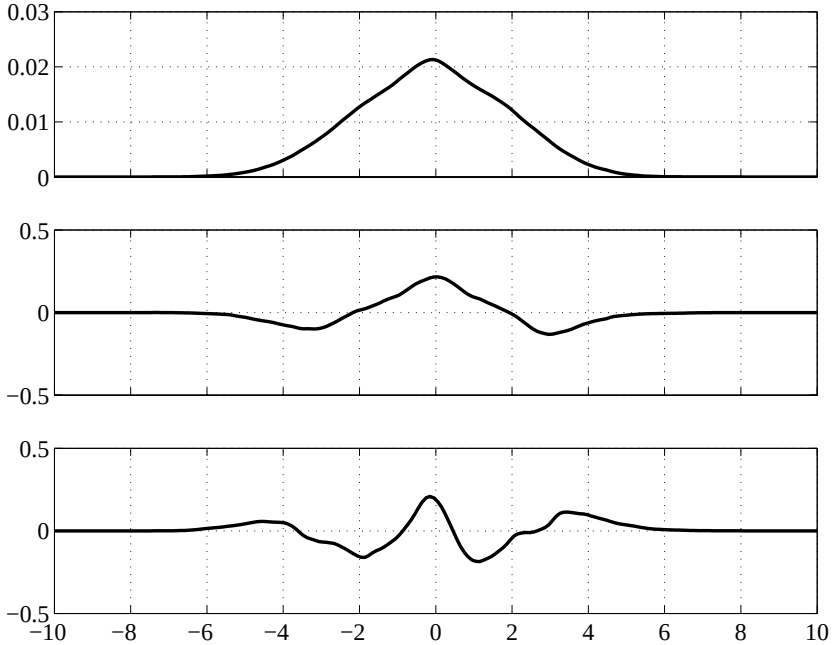

Figure 8: Mean and first two principal components, from top-to-bottom respectively, of mass projected onto the second principal axis.

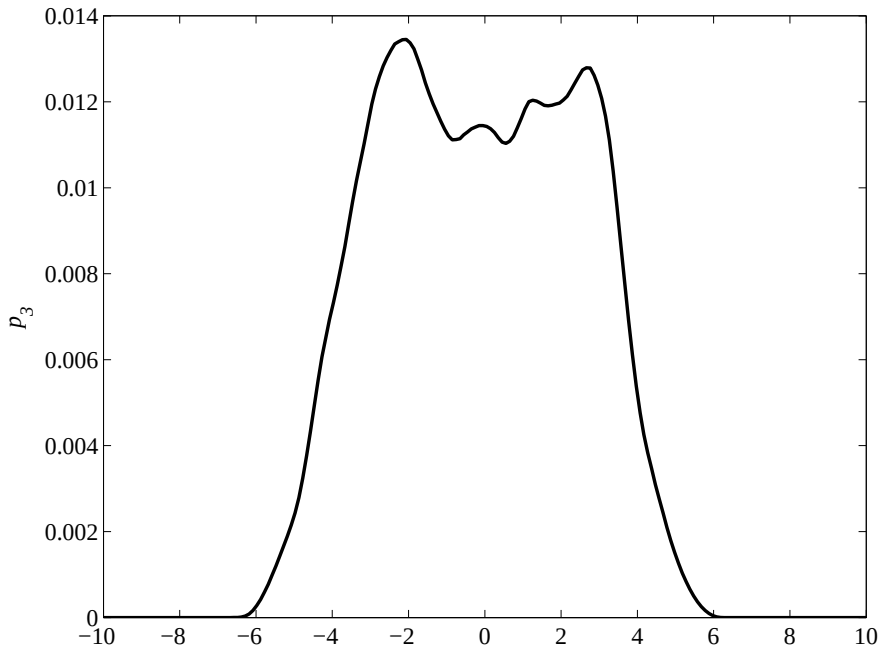

Figure 9: Mass projected onto the third principal axes, $p_3(x_k)$, for the spindle of Fig. 4.

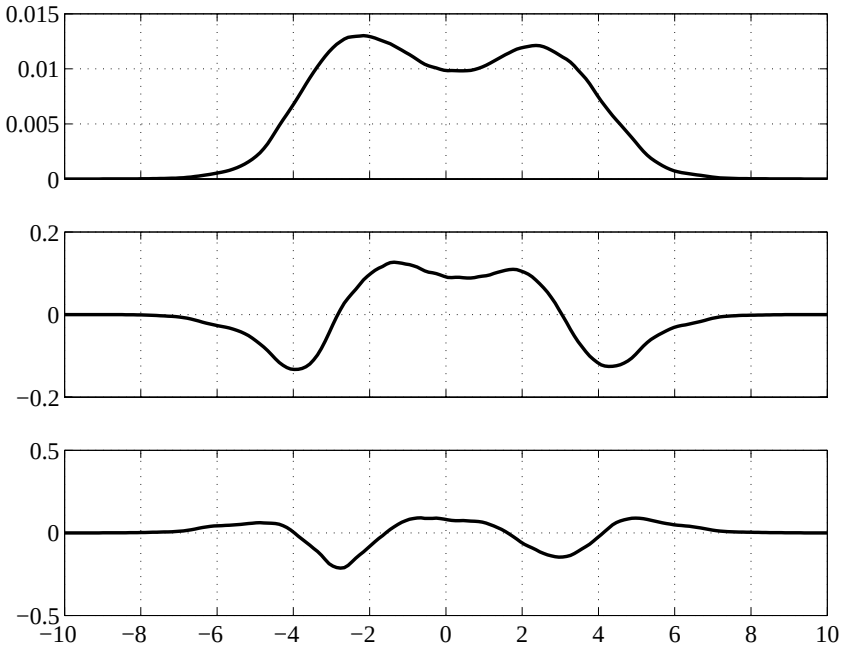

Figure 10: Mean and first two principal components, from top-to-bottom respectively, of mass projected onto the third principal axis.

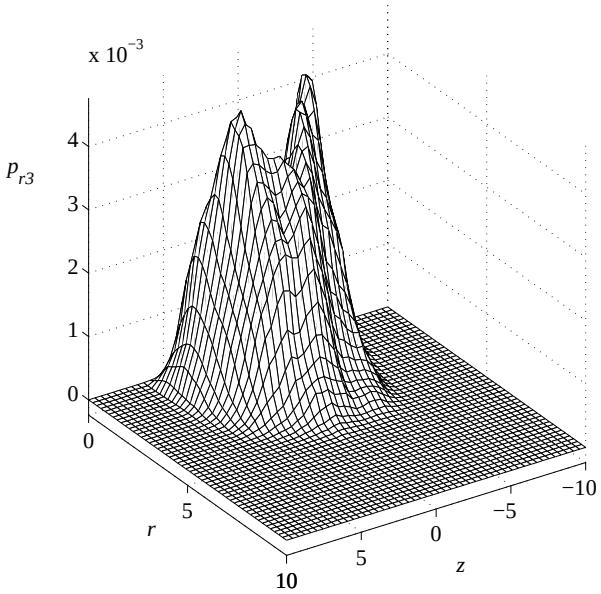

Figure 11: Radial-Axis 3 mass distribution, p_{r3}, of the spindle in Fig. 4, shown as a surface plot.

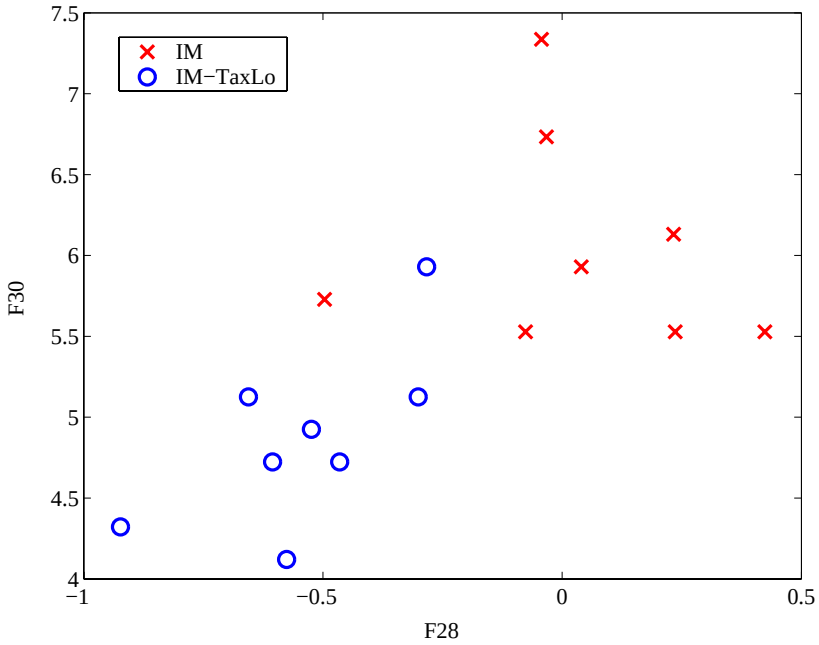

Figure 12: Scatter plot of features F28 and F30 for the IM vs. IM-TaxLo comparison of Section 4.1.

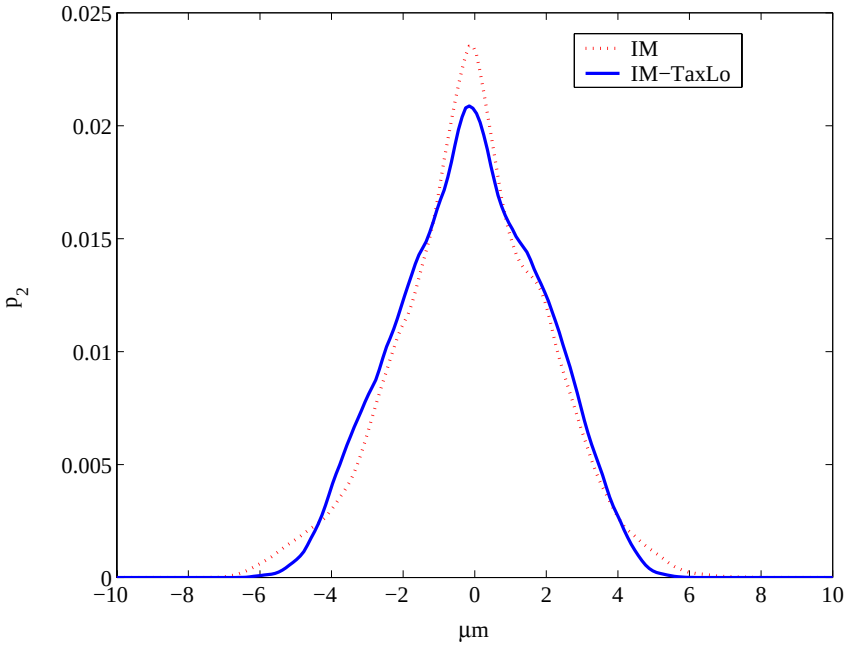

Figure 13: Mean of Axis 2 mass projections for the IM vs. IM-TaxLo comparison of Section 4.1.

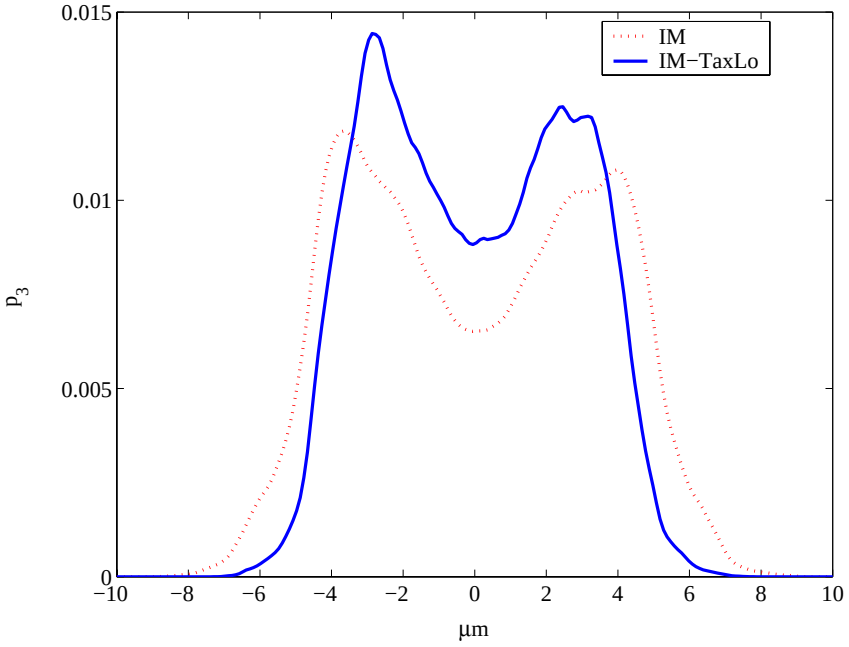

Figure 14: Mean of Axis 3 mass projections for the IM vs. IM-TaxLo comparison of Section 4.1.

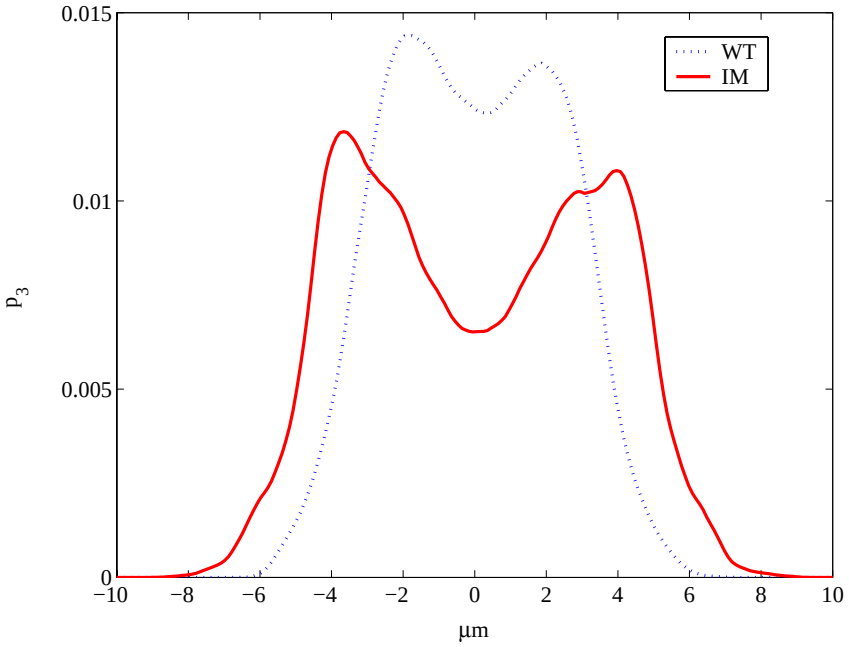

Figure 15: Mean of the Axis 3 mass projections 3 for the WT vs. IM comparison of Section 4.3.

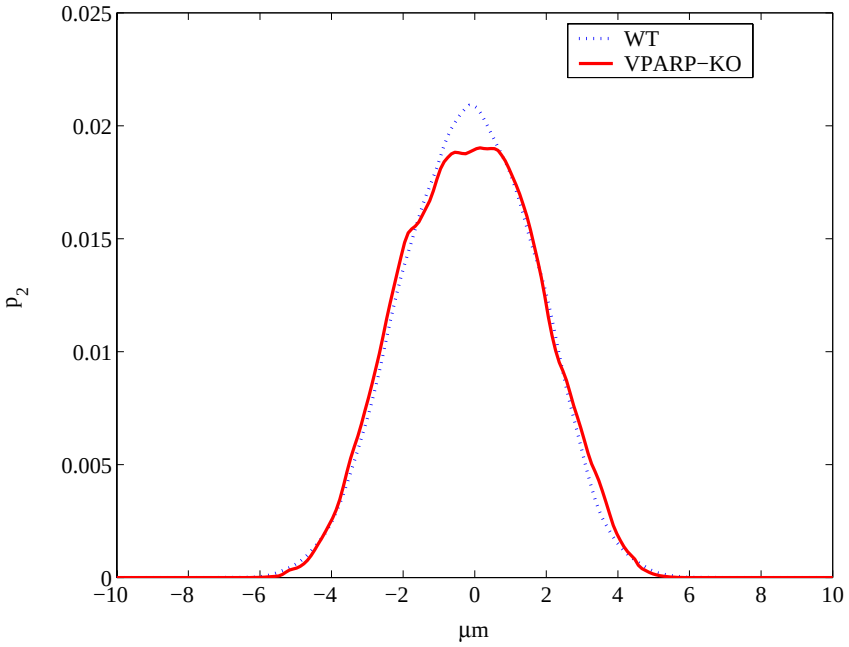

Figure 16: Mean of the Axis 2 mass projections for the WT vs. VPARP-KO comparison of Section 4.4.

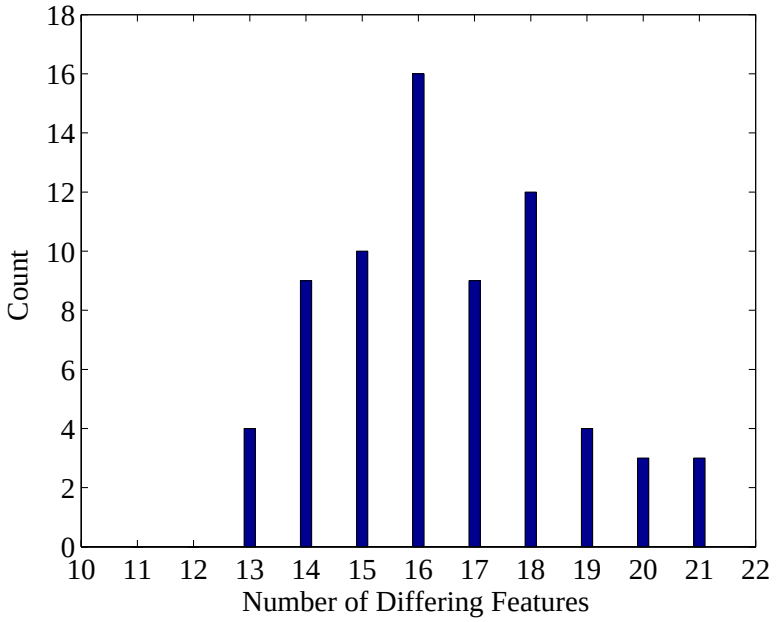

Figure 17: Number of differing features for the 70 comparisons of WT vs. p53-KO from Section 4.5. The mean number of differing features is 16.46 and the median is 16.

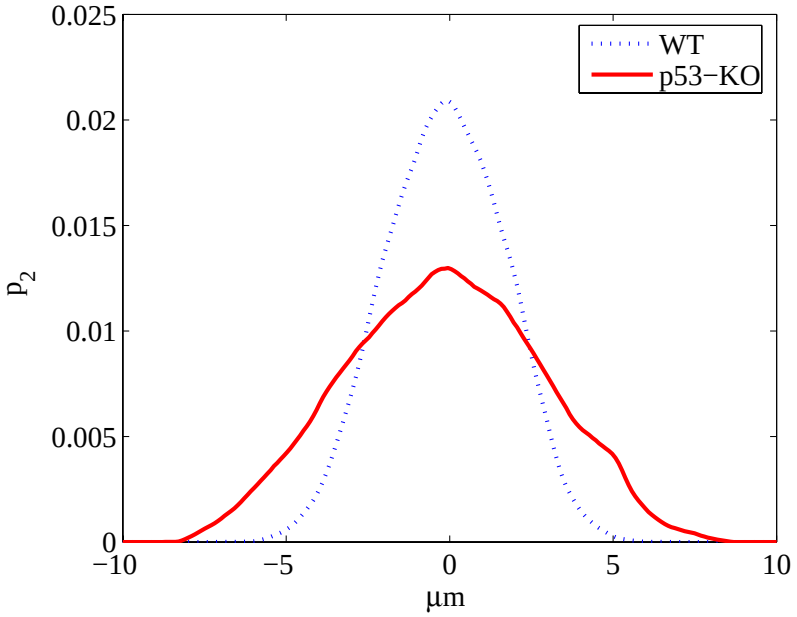

Figure 18: Mean of the Axis 2 mass projections for the WT vs. p53-KO comparison of Section 4.5.

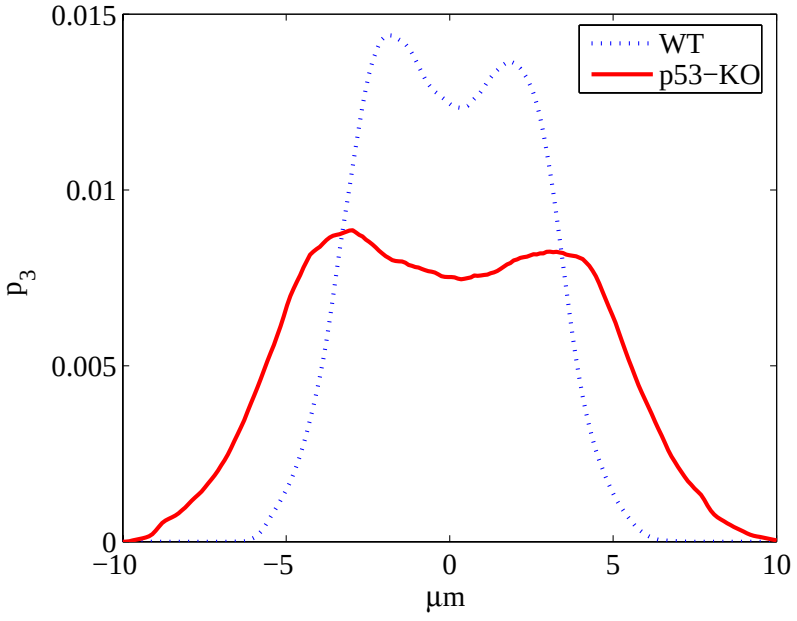

Figure 19: Mean of the Axis 3 mass projections for the WT vs. p53-KO comparison of Section 4.5.