AD_____
(Leave blank)

Award Number: DAMD170210440

TITLE:

Novel Image Analysis to Link Sub-Nuclear Distribution of Proteins with Cell Phenotype in Mammary Cancer

PRINCIPAL INVESTIGATOR:
(Enter the name and degree of Principal Investigator and any Associates)
David W. Knowles, Ph.D., PI, LBNL
Sophie A. Lelièvre, D.V.M., Ph.D., Collaborator, Purdue University, Basic Medical Sciences
Fuhui Long, PhD., Post Doctoral Fellow, Collaborator, LBNL
Jennifer Gong, Under Graduate Student, Collaborator, LBNL
Sunil Badve, MBBS, MD(Path), FRCPath, Collaborator, Indiana University Cancer Center
Gunther Weber, Ph.D., Collaborator,University California, Davis, Visualization and Graphics Research Group
Bernd Hamann, Ph.D., Consultant, University California, Davis, Visualization and Graphics Research Group
Damir Sudar, M.Sc., Consultant, LBNL
Mina J. Bissell, Ph.D., Consultant, LBNL

CONTRACTING ORGANIZATION:
(Enter the Name, City, State and Zip Code of the Contracting Organization)

Lawrence Berkeley National Laboratory, Berkeley, CA 94720

REPORT DATE:
(Enter month and year, i.e., January 2001)
May 2006

TYPE OF REPORT:
(Enter type of report, i.e., annual, midterm, annual summary, final)
Final Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. Agency Use Only (Leave blank)	2. Report Date	3. Report Type and Period Covered: Final Report
	20 May 2006	23rd April 2002 - 22 April 2006

4. Title and Subtitle
Novel Image Analysis to Link Sub-Nuclear Distribution of Proteins with Cell Phenotype in Mammary Cancer

5. Award Number
DAMD170210440

6. Author(s)
David W. Knowles

7. Performing Organization Name (Include Name, City, State, Zip Code and Email for Principal Investigator)
Lawrence Berkeley National Laboratory,
84R0171, 1 Cyclotron Road, Berkeley, CA 94720

E-**Mail: dwknowles@lbl.gov**

8. Performing Organization Report Number (Leave Blank)

9. Sponsoring/Monitoring Agency Name and Address
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. Sponsoring/Monitoring Agency Report Number (Leave Blank)

11. Supplementary Notes (i.e., report contains color photos, report contains appendix in non-print form, etc.)

12a. Distribution/Availability Statement (check one)
☐ Approved for public release; distribution unlimited
☒ Distribution limited to U.S. Government agencies only - report contains proprietary and unpublished information.

12b. Distribution Code
(Leave Blank)

13. Abstract ***(Maximum 200 Words)*** *(abstract should contain no proprietary or confidential information)*

The goal of this project is to develop novel optical imaging / image analysis techniques to allow automated, quantitative screening to distinguish malignant, pre-malignant, and non-malignant mammary tissue. Our hypothesis is that cellular and tissue phenotype is reflected by the organization of components within the nucleus. Using cultured and biopsied human breast tissue, we have developed imaging methods to quantify the distribution of fluorescently-stained nuclear proteins in different mammary phenotypes. Automated image segmentation of nuclei was generated to isolate thousands of nuclei from three-dimensional confocal images. Prominent features of fluorescently-stained nuclear proteins were detected and analyzed using a novel local bright feature analysis techniques. Cluster analysis techniques were developed to classify individual cells. Visualization techniques were developed to map results onto the tissue context for viewing. The results revealed marked changes in the distribution of NuMA and H4K20m bright features as non-neoplastic cells underwent phenotypically normal acinar morphogenesis. The analysis also discriminated proliferating non-neoplastic cells from proliferating malignant cells, suggesting that these imaging methods are capable of identifying alterations linked not only to the proliferation status but also to the malignant character of cells. We believe that this quantitative analysis will have additional applications for classifying normal and pathological tissues.

14. Subject Terms (keywords previously assigned to proposal abstract or terms which apply to this award)

15. Number of Pages (count all pages including appendices)

16. Price Code (Leave Blank)

17. Security Classification of Report	18. Security Classification of this Page	19. Security Classification of Abstract	20. Limitation of Abstract
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

Novel Image Analysis to Link Sub-Nuclear Distribution of Proteins with Cell Phenotype in Mammary Cancer

Investigators: **David W. Knowles** (Collaborating laboratories: Sudar, Lelièvre, Bissell, Badve & Hamann)

INTRODUCTION:

The goal of this project is to develop novel optical imaging / image analysis techniques that will allow automated, quantitative screening to distinguish malignant, pre-malignant, and non-malignant mammary tissue. Our hypothesis is that cellular and tissue phenotype is reflected by the organization of components within the nucleus. By developing imaging-based methods to quantify the spatial distribution of these proteins, the work will provide understanding of how the nuclear distribution of various chromatin-associated proteins correspond with the phenotype of the cells and tissue. We believe that this quantitative analysis will have applications for classifying normal and pathological tissues.

BODY:

A. Summary of work:

This work focuses on the distribution of chromatin-related nuclear proteins in a progression series of cultured HMT-3522 human mammary epithelial cells (HMECs) that mimic early stages of cancer development. Nonneoplastic S1 HMT-3522 cells recapitulate differentiation into phenotypically normal breast glandular structures in 10 day of three-dimensional (3D) culture. 3D culture is preformed by placing cells in contact with an exogenous extracellular matrix enriched in basement membrane components (Matrigel), and supplying cells with essential growth factors and hormones. Glandular structure formation encompasses cell proliferation (until day 6), growth-arrest and deposition of a continuous endogenous basement membrane around the glandular structure. Malignant T4-2 HMT-3522 cells mimic tumor growth, with the formation of disorganized multicellular structures in which cells keep proliferating, when cultured in the same conditions. Nuclear mitotic apparatus, NuMA, protein has been previously found to undergo remarkable changes in its nuclear organization during glandular structure formation[6]. During the proliferating stage the distribution patterns of NuMA appear homogenous throughout the cell nucleus and was seem similar to what is observed in tumor cells. Whereas, upon differentiation, NuMA is reorganized into foci within the cell nucleus and ring-like patterns at its periphery. Culturing of the cells and fluorescence labeling of the NuMA protein was be carried out, under subcontract, at the Lelièvre laboratory at Purdue University. Image collection, the development of novel image analysis techniques and the image analysis has be performed at our facility at LBNL.

Year 1. Key outcomes of the first year of work (Months 1 - 12) were the acquisition of images of NuMA organization in cells cultured between 4 and 12 days and the development of a local bright feature (LBF) analysis method. Previously we had only studies proliferating cells at 4 to 5 days of culture, but after 10 days of culture, nonmalignant cells have differentiated into acini and malignant cells have proliferated into large unorganized clusters. Analyzing images from these cultures has strengthened our initial hypothesis about the reorganization of NuMA in nonmalignant cells and this has been key in the further development of our image analysis techniques. The LBF analysis method we developed allows bright foci within a nucleus to be isolated from the diffuse NuMA staining.

Year 2. In the second year of work (Months 13 - 24) we have maintained our primary focus on quantifying the distribution of NuMA and extended our image analysis capability to allow the distribution of NuMA foci to be measured within an individual nucleus and thus on a per nucleus basis. Briefly, this has been achieved by several key developments. 1) Our nuclear segmentation ability has been extended to allow automated

segmentation of individual nuclei from a cluster of cells. 2) Our local bright feature (LBF) analysis has been modified to allow both local bright and local dark features within a nucleus to be isolated for further analysis. 3) Image analysis techniques were developed to allow each nuclear volume to be subdivided into a set of concentric terraces. These refinements have allowed us to calculate the relative density of NuMA foci in each nucleus, on a per terrace basis and evaluated as an effective radial distribution from the perimeter of the nucleus to its center.

Year 3. In year 3, (Months 25-36) significant development was done on an automated segmentation analysis which isolated the position and extent of nuclei from 3D images. The local bright feature analysis was refined to extract local dark as well as local bright features from within images of nuclei fluorescently stained for specific nuclear proteins. Compartmentalization of individual nuclei into a set of concentric terraces allowed the quantification of the radial distribution of specific labeled proteins in individual nuclei. Application of these automated imaging techniques to reveal striking differences in the organization of NuMA between proliferating non-malignant cells and proliferating malignant cells. These methods have allowed us to representing the distribution of NuMA bright features associated with different mammary phenotypes as a simple graph, hence enabling an easy interpretation of the spatial distribution of the protein. The performance of the LBF analysis was greatly enhanced by the development of an automated segmentation of the nuclear volume that enabled us to analyze thousands of nuclei in a short period of time. Using this novel image analysis technique we measured the striking reorganization of NuMA during acinar morphogenesis, while no such reorganization occurred during tumor formation. Most importantly, the LBF analysis permitted a clear discrimination between proliferating non-malignant cells and proliferating malignant cells, which was not achieved so far using other evaluation methods.

Year 4. In year 4, (Months 37-48) significant progress was achieved.
1) Based on our cluster analysis results for the distribution of NuMA in nonneoplastic (S1) and malignant (T4) cells, we have developed a classification method which is able to detect the phenotype of cells at the single cell level (SectionC.4). This work is the basis of a manuscript in preparation[42].
2) We have measured the radial distribution of NuMA in premalignant (S2) cells which growth arrest into spheroids which comprise vastly different numbers of cells.
3) We then applied our clustering analysis technique to the distribution results of NuMA in premalignant (S2) cells, which reveals at least two separate spheroid phenotypes. Interesting the two phenotypes were size dependent.
4) In this past year we have started developing visualization methods in three dimensions. Through a collaboration with the Visualization and Graphics Research Group at the University California, Davis, we have developed novel visualization tools to: i) render protein distribution onto the segmentation mask (see Section C.6.1), ii) render analysis results onto the tissue context (see Section C.6.2) and iii) render analysis results onto section of normal human mammary tissue (see Section C.6.3).
5) This work was selected for platform presentation at the U.S. Department of Defense Breast Cancer Research Program funded Era of Hope meeting, June 8-11 2006 in Philadelphia.
6) The project has had a major publication in the **Proceedings of the National Academy of Sciences**[24]. See appendices.
7) The publication in PNAS[24] resulted in a press release and an article describing the work in the monthly magazine *The Berkeley Lab View*
(see **http://dwknowles.lbl.gov/060306.Lelievre.fluo.html**)
(see **http://www.lbl.gov/Publications/Currents/Archive/Apr-21-2006.html**).
8) Knowles has submitted and resubmitted an R33 proposal to continue this work to the NIH/NCI. In the first round, the proposal received a score or 168 and in the second round, a score of 153. The proposal is currently in programmatic review. As a new investigator, I was encouraged by, and grateful for, the comments and concerns

raised by the reviewers. All have agreed that the proposed work was innovative and all had been highly supportive.

> Reviewer 1 said "***The proposed research will yield a unique bridge between structural expression of nuclear proteins and cancer classification phenotype.***"
>
> Reviewer 2 said (Histopathology) "***is laborious, time-consuming, and importantly, susceptible to examiner subjectivity. Consequently, means to make the process quantitative and automatable is highly desirable and of great significance.***"
>
> Reviewer 3 said "***This is careful, deliberative work. The investigators have mapped out a good plan that promises to yield a significant new technique for cancer treatment.***"
>
> Reviewer 4 said "***This has required the development of some fairly sophisticated nuclear segmentation and image analysis tools, for which they should be applauded. The next steps, which are to move to mixed cultures and eventually to frozen sections and tissue biopsies, are exactly the right steps.***"

B. Background

Under an exploratory Department of Defense Breast Cancer Research Program Award (see: http://cdmrp.army.mil/cgi-bin/search/get_abstract.pl?id=5467&log=BC011187) which has supported a multidisciplinary collaboration between Sophie Lelièvre, a cancer biologist at Purdue University and myself, we have developed image analysis methods to quantitatively describe the organization of specific nuclear chromatin-associated proteins. By applying our novel Local Bright Feature (LBF) methods to three-dimensional culture models that mimic normal and malignant breast epithelial tissue we have demonstrated that the distribution of Nuclear Mitotic Apparatus protein (NuMA) and heterochromatin related protein histone 4 methylated on lysine 20 (H4-K20m) are biomarkers capable of distinguishing non-neoplastic and malignant human mammary epithelial cells.

The broad goal of this project is to develop technologies capable of turning high resolution fluorescence images of human mammary epithelial tissue into tissue-maps which report the probable nonneoplastic, premalignant and malignant phenotype at cellular resolution. The translational application of these methods is to aid the treatment decision process of breast cancer patients by providing pathologists with a *phenotype tissue-map*, based on nuclear protein organization, to aid and support the histological classification of biopsied breast tissue.

Our working hypothesis is that the distribution of chromatin-related proteins will permit a novel imaging-based phenotype screening of individual nuclei and the recognition of subtle differences in tissue morphology and behavior, which would enable better detection of benign and malignant lesions. Our rationale is that chromatin organization and associated redistribution of chromatin-related proteins reflect the changes in gene expression that accompany alterations in cell phenotype. Thus, a wide range of distinct distributions of chromatin-related proteins characteristic of different stages of breast cancer and/or degrees of cell malignancy should be recognized. The goal of this project will be achieved in three specific aims.

B1. The Problem

Breast tumors are detected by self exam, clinical exam and mammogram. Suspicious results are evaluated with diagnostic mammography, ultra-sound and magnetic resonance imaging, and needle and surgical biopsy followed by histological and cytological analysis. The histological classification of these biopsies plays a determining role in the treatment decision. Unfortunately, the risk associated with benign or pre-invasive

disease, the appropriate treatment or adjuvant treatment, and the risk of reoccurrence are poorly understood. As a consequence patient treatment is based on epidemiological findings rather than individual needs. Non-invasive tissue imaging using X-rays, ultra-sound and magnetic resonance provide information about the presence of breast tumors but the goal of early cancer detection by routine non-invasive screening is simply not currently feasible or realistic. However, once a tumor has been detected and biopsied, there is urgent need for novel methods that would rapidly quantify subtle epigenetic abnormalities within the tumor cells and aid the identification of lesions of poor prognosis within the categories of benign, premalignant and malignant diseases.

B2. Existing Knowledge

One key epigenetic marker of cell phenotype is the organization of nuclear proteins which direct and reflect nuclear functions necessary to orchestrate cell proliferation and differentiation[1-3]. It is becoming increasingly apparent that the organization of the cell nucleus plays a central role in directing nuclear functions necessary to orchestrate cell proliferation and differentiation[1-3]. Nuclear functions are carried out by proteins that display specific compartmentalization (e.g., nucleolus, nuclear domains, chromatin) and distribution (e.g., diffuse, foci-like) characteristics. In addition, there is growing evidence to show that the organization of nuclear proteins is dynamic, as illustrated during the cell cycle[4-5], upon the switch between proliferation and growth arrest[6], and following cell differentiation[7-8]. The use of 3-D models of cell culture has revealed that the distribution of nuclear proteins is also dependent on tissue morphogenesis[9,6].

3D cell culture models have been used to compare the organization of major nuclear proteins in non-differentiated and differentiated mammary epithelial cells to better apprehend the role of these proteins in the establishment and maintenance of mammary acinar differentiation. Remarkable differences have been found in the distribution of nuclear proteins including Rb, splicing factor SRm160, and NuMA between the early stage and the completion of acinar morphogenesis by non-neoplastic HMT-3522 S1 human mammary epithelial cells (HMECs)[6]. Notably, the distribution of NuMA displays a gradual modification during the 10-day *in vitro* morphogenesis process. NuMA staining is diffusely distributed within the nuclei of proliferating cells, but aggregates into foci of increasing size as cells arrest proliferation and complete acinar morphogenesis. Importantly, the distribution of NuMA in acinar S1 cells is similar to that observed on biopsies of normal breast tissue, indicating that the 3D model of acinar morphogenesis reproduces physiologically relevant characteristics of NuMA organization. Although the distribution of NuMA is classically reported as diffuse in cells cultured in non-differentiating condition (i.e., cultured on plastic surface producing a flat monolayer of cells), the reorganization of NuMA observed during mammary epithelial morphogenesis in 3D culture has been steadily reinforced by reports indicating that this protein has distinct distribution patterns in different tissues[10] and is observed as a single aggregate in the center of the cell nucleus in different cell types undergoing apoptosis[11-12]. Moreover, a more punctate distribution of NuMA has been associated with a higher susceptibility to apoptosis induction in lymphocytes[13]. Thus, NuMA distribution appears highly dependent on cell and tissue phenotypes and, as such, it has been proposed to constitute a good indicator of cell behavior[6;11;13].

Notably, the involvement of NuMA in the chromatin compartment and the importance of NuMA organization for the maintenance of chromatin structure[14], suggest that NuMA distribution may be highly sensitive to changes in gene expression profile associated with alterations in cell phenotype. Similarly we have observed dramatic alterations in the distribution of chromatin markers during breast differentiation such as heterochromatin markers histone 3 methylated on lysine 9 (H3K9m)[9] and H4-K20m, and telomere-associated protein TIN2[15]. In contrast, structural proteins like lamin B or PML did not show a striking reorganization upon acinar differentiation[6].

C. Results

C1. Development of the Local Bright Feature (LBF) image analysis

It has been understood for decades that cancer cells display alterations in nuclear size and general chromatin organization as measured by density and texture[16], but there is very little information available regarding alterations in specific nuclear components in neoplasias. With the advent of confocal imaging technology, three-dimensional high-resolution imaging has become a powerful method for obtaining such organizational data within the nucleus. However, the availability and application of robust image analysis tools in biology remain in their infancy[17] and groups who have relied on three-dimensional images to understand underlying mechanisms in biology have had to develop their own image analysis tools[18-21;8;22-23]. As described below, we have developed methods which first quantify the nuclear distribution of specific proteins from 3D fluorescence images of cells and tissue, at cellular resolution. Then, using the distribution results from each cell as a feature that describes that cell, we are able to group cells with similar features and study their arrangement within the natural tissue context. Our current methods are described.

C.1.1. Automated segmentation from three-dimensional images

To analyze the internal nuclear organization of thousands of nuclei on a per nucleus bases, we have developed an automated nuclear segmentation algorithm that is able to define the location and extent of all nuclei from 3D fluorescence images of total DNA[24]. The method is based on the assumption that nuclei of epithelial cells are of simple geometry in that they comprise a single spherical core. An adaptive threshold was first applied to the DAPI-stained image to produce a binary mask of the nuclei. The technique uses a difference-of-Gaussians filter[25], followed by a morphological closing filter and a flood-fill algorithm to produce the binary segmentation mask[26-28]. While this technique accurately separates nuclei from their background, it does not completely isolate neighboring nuclei when they are tightly clustered. To separate nuclei that are connected by the binary mask, their central nuclear cores were located using standard template matching techniques, and dilated into the rest of the nucleus using standard region-growing techniques[26-28]. Briefly, a template, with dimensions that approximate the average spherical core of nuclei, was convolved with the binary nuclear mask. This convolution produced a map that indicated the percentage of the template that fit within the binary mask at each point in the image. Then, the template was stamped into the binary mask at locations where there were corresponding local maxima in the map that exceeded 70%. The templates were stamped at the center-of-mass of the local maximum, in an order ranked by their percentage, starting from the highest. A template was not stamped if the local maximum was less than 70%, if it overlapped a previously stamped template by more than 70%, or if the local maximum was at the boundary of the binary mask. Once all the nuclear cores were located, each template was dilated in a semi-intelligent fashion into the binary mask. The template dilation was done independently in the positive and negative X, Y, and Z directions. Dilation along any direction was halted when 60% of the dilating template boundary reached the boundary of the binary mask. This prevented a dilating template from squeezing through narrow regions in the binary mask that connected two adjacent nuclei. In addition, dilation was stopped in all directions if the volume of the dilated template exceeded nine times its original volume. This phenomenon occurred if nuclei were clustered so closely that the initial segmentation mask failed to separate them adequately. The resulting object was reported as an under-segmentation error. Our novel approach of nuclear segmentation is similar in part to that of Irinopoulou and colleagues[19]. However, in their work, they used a global threshold, followed by a distance transform and watershed method to segment nuclei on a per image-slice basis. Then they implemented a rule-set for correctly joining nuclei in adjacent slices. Instead, our technique uses an adaptive threshold, which works in three dimensions, using a template matching method followed by intelligent region growing of the templates.

C.1.2. LBF algorithm used to isolate bright features of immunostaining

Following immunostaining, image acquisition by confocal microscopy, and segmentation, bright staining features/foci are detected and isolated by the LBF analysis technique. This technique was originally developed

with NuMA staining but has wide application. In this technique, pixel brightness in the raw NuMA images is normalized by the local average brightness using an extension of the difference-of-Gaussians technique[25]. First the raw NuMA image is masked by the binarized segmentation result derived from the DAPI image as described above (Figure 1). Then, image brightness within each nucleus is rescaled by dividing the brightness at each point by the local average brightness in a region surrounding that point. The local average brightness is calculated over a region with a dimension half that of the dimension of the nuclear core. This is important because the LBF technique sensitively resolves light or dark features that are smaller than the region but ignores features that are larger. Thus, this technique allows the bright features and dark regions of interest within the nucleus to be resolved and normalized for the low-frequency brightness variations due to the geometry of the nucleus and the axial resolution of the microscope. In the resulting LBF images, bright image features have values above unity while dark image features have values below unity.

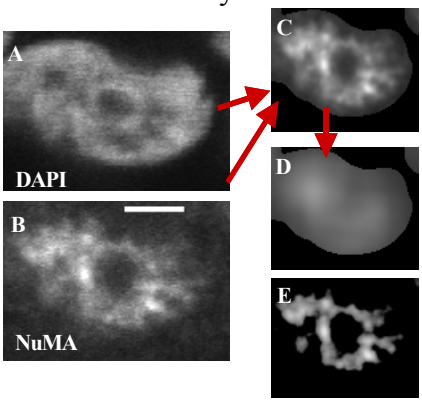


Figure 1. Isolation of bright features of NuMA staining using the LBF algorithm. (**A**) Single central optical section from a 3D image of an nucleus stained for total DNA. Total DNA images are used to create the the nuclear segmentation masks. (**B**) The corresponding optical section showing fluorescence immunostained NuMA. (**C**) The corresponding slice of the NuMA image, multiplied by the segmentation mask and convolved with a narrow Gaussian filter which removes shot or detector noise. (**D**) The similar process as in C but using a wider Gaussian filter. The result is the local average brightness of the image. (**E**) The corresponding slice shows the local bright feature (LBF) image, generated by taking the positive potion of the subtraction of images shown in C and D. This process reveals local bright features and is termed the LBF image.

A major concern of using three-dimensional confocal images for quantitative image analysis is the inherent properties of the images, which are linked to the nature of their acquisition. Confocal images are more highly resolved in directions perpendicular than parallel to the optical axis. This is due to the spatial asymmetry of both the point-spread function of the excitation illumination and the microscope's "pinhole" spatial filter. In addition, the collection efficiency and hence the brightness of confocal images decreases with depth into the object. This penetrative loss is the result of the physical properties of the object, its mounting medium and the objective lens. Standard image analysis techniques are often based on ideal imaging assumptions that neglect the inherent properties of confocal images. In such cases, images must be preprocessed using restoration techniques to remove the confocal artifacts before a quantitative analysis proceeds. In contrast, our approach was to design a technique that takes the inherent properties of confocal images into account, and still allow results from images acquired in different ways to be quantitatively compared. The aim of the LBF analysis was to isolate local bright and local dark features within an image using an adaptive approach. To do so, at each point in the image a kernel of neighboring imaging pixels is defined around the point of interest. The LBF analysis then uses the relative brightness of the neighboring pixels in the kernel to classify each pixel. These types of non-linear techniques are powerful because they mimic human visual perception, especially the ability to isolate rare events like small numbers of foci in a diffuse background. Furthermore, the size of the kernel sets a spatial sensitivity limit to the LBF technique and its relative dimensions can be easily adjusted to match the spatial sampling asymmetry of the microscope. Consequently, the LBF technique is not affected by the absolute brightness of an image or long scale brightness variations, like penetrative loss. Restoration techniques like background subtraction, attenuation correction and image interpolation are simply not necessary.

C.2. LBF analysis of NuMA measures different distributions depending on the mammary tissue phenotype.

The radial-LBF analysis calculates the normalized distribution of bright features within fluorescence images at single cell resolution. The application of this technique for quantifying the nuclear organization of NuMA is shown in Figure 2.

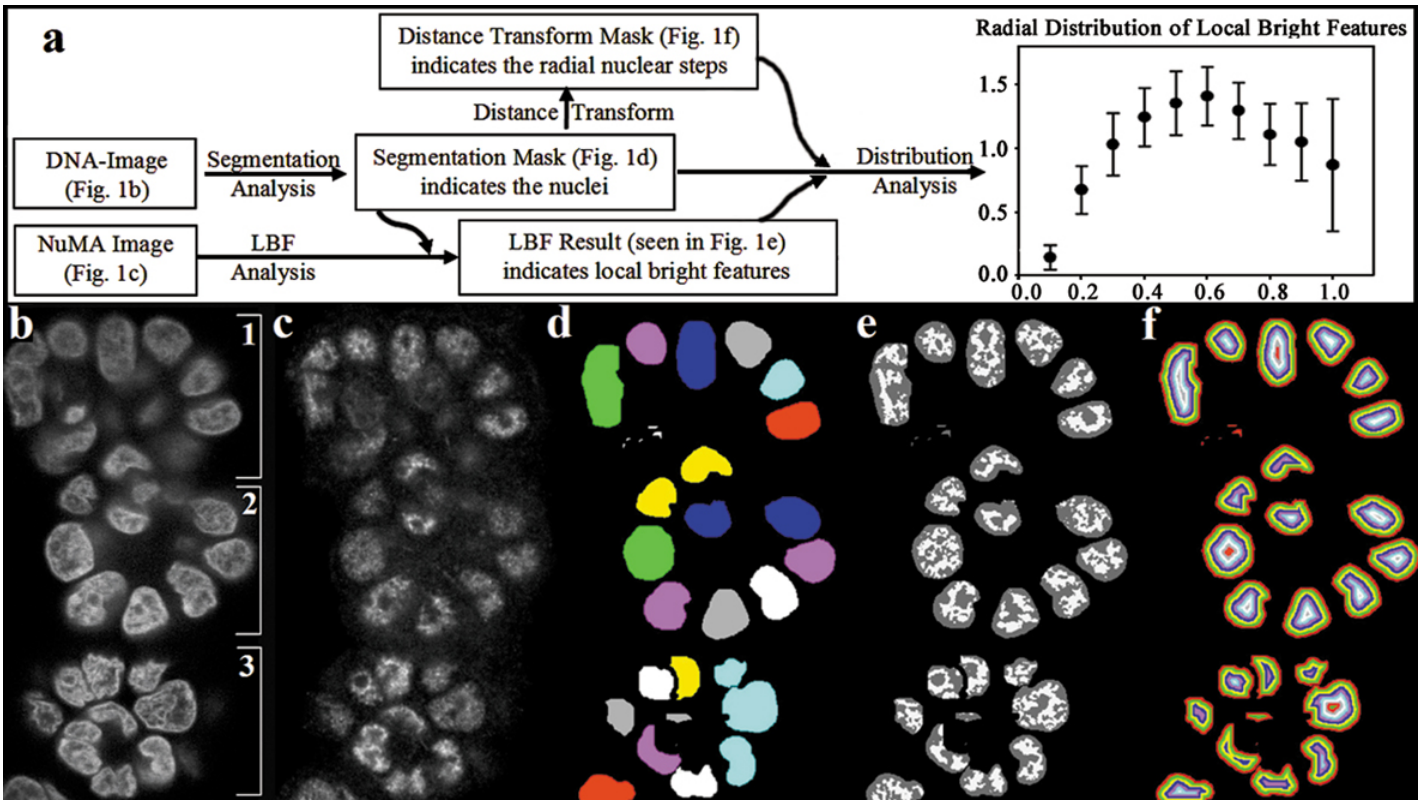


Figure 2. LBF analysis of NuMA distribution from 3D images. a. Flowchart of the imaging processing steps including a graph of the relative density of local bright features of NuMA in 77 nuclei from the three acini depicted in Fig 1b. The radial distributions of local bright features within each nucleus in the NuMA image are calculated by first generating a segmentation mask from the image of DAPI-stained DNA. The segmentation mask not only defines the extent of each nucleus, but it is also used to define a set of radial steps within each nuclear volume. The graph shows the mean and standard deviation of the relative density of NuMA bright features extracted by LBF analysis (ordinate) as a function of the relative distance from the perimeter (0.0) to the center (1.0) of the nuclei (abscissa). **b-h**. Distribution density of the bright features of NuMA in acinar cells. S1 HMECs were cultured in 3D to induce acinar morphogenesis. Each panel corresponds to the application of the different steps of distribution analysis starting from the same original image. **b.** Fluorescence micrograph of DAPI-stained nuclei from a single optical section containing three acini [1;2;3]. **c.** Fluorescence micrograph of Texas-red immunolabeled NuMA from the optical section corresponding to the DAPI image shown in **a**. **d.** Segmentation mask derived from the DAPI-stained image showing a single slice of individually enumerated nuclei. **e.** Composite view of the local bright features (light gray) extracted by the LBF analysis overlaid with the segmentation mask (dark gray). **f.** Concentric terraces resulting from the application of the distance transform on the segmentation mask.

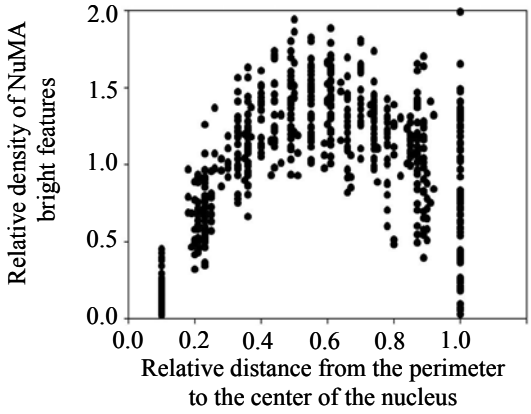


Figure 3. Relative density of local bright features of NuMA in 77 nuclei from a single image of three acini. Multi-overlay plots of the radial-LBF analysis of HMECs cultured in 3D to induce acinar morphogenesis shows the relative density of NuMA bright features as a function of the relative distance from the perimeter to the center of each nucleus.

3D fluorescence image are acquired for DAPI-stained total DNA (Figure 2A) and Texas-Red immunostained-NuMA (Figure 2B). The segmentation mask (Figure 2C) is derived from the total-DNA image as described in section C.1. For visualization purposes, the LBF image (Figure 2D) calculated from the NuMA image is overlaid with the segmentation mask. This clearly indicates that the density of NuMA bright features is low at the periphery of the nucleus and varies with the depth, radially, into the nucleus. In order to calculate the radial distribution of the local bright features, a distance transform[26-28] was applied to the nuclear segmentation mask. The transform calculates the shortest distance of each point within a nucleus to the nuclear boundary and in doing so, divides each nucleus into a set of concentric terraces of equal thickness (Figure 2E). The LBF image is then used in conjunction with the nuclear segmentation mask and the distance transform to compute the density of local bright features in each terrace of each nucleus. To reveal the relative distribution of the density of bright features within each nucleus, the density per terrace was normalized so that the average density of bright features is unity for each nucleus. The distances defined by the distance transform were also normalized so that the distance at the nuclear perimeter was 0 and the distance at the center of the nucleus was 1.0. This normalization was done to account for variations in the number of terraces per nucleus due to variations in nucleus size and shape. Finally, the normalized density of bright features was plotted against normalized distance from the perimeter of the nucleus to its center. To demonstrate the consistency of this method, the radial distribution was plotted for 77 nuclei of acinar S1 cells analyzed within a single image (Figure 3). As the visual representation indicated (see Figure 2D), the density of NuMA bright features was below average at the perimeter of the nuclei. And as the radial depth into the nucleus increased, the density of bright features increased and reached a peak, above the average density, at a radial distance located around 0.55+-0.05. Then the density decreased to a value close to the average density as the center of the nucleus was reached.

Previous visual analysis of the distribution of NuMA in S1 HMECs during the proliferation stage (day 3 of 3-D culture) and upon acinar differentiation (day 10 of 3-D culture), suggested that there was an increase in the foci-like aggregation of NuMA upon completion of acinar morphogenesis[6]. This analysis was based on visual estimation and the manual measurement of the size of NuMA foci on images of NuMA staining. Thus, this painstaking and quasi-quantitative analysis only allowed us to get a gross comparison of the redistribution of NuMA during morphogenesis.

To assess the efficiency of the radial-LBF analysis in measuring reproducibly the changes in NuMA distribution along the morphogenesis process, we analyzed images of S1 HMECs cultured in 3D for 3, 5, 10 and 12 days. The radial-LBF analysis revealed changes in the distribution of the relative density of bright features of NuMA immunostaining during acinar morphogenesis. For S1 HMECs cultured for 3 days, the density of NuMA bright features was highest in a region towards the perimeter (at 0.35R), and maintained an average density elsewhere within the nucleus. After 12 days of culture there was a clear reorganization of NuMA away from the perimeter and towards the center of the nucleus (Figure 4). This was accompanied by a significant decrease of NuMA

bright features at the perimeter of the nucleus. Images of NuMA distribution in acinar cells indeed showed that this protein was usually absent from regions located at the periphery of the nucleus (Figure 5). These results confirmed our initial observation that acinar morphogenesis is accompanied by reorganization of NuMA[6]. Similar density distributions were obtained with two different antibodies directed against NuMA.

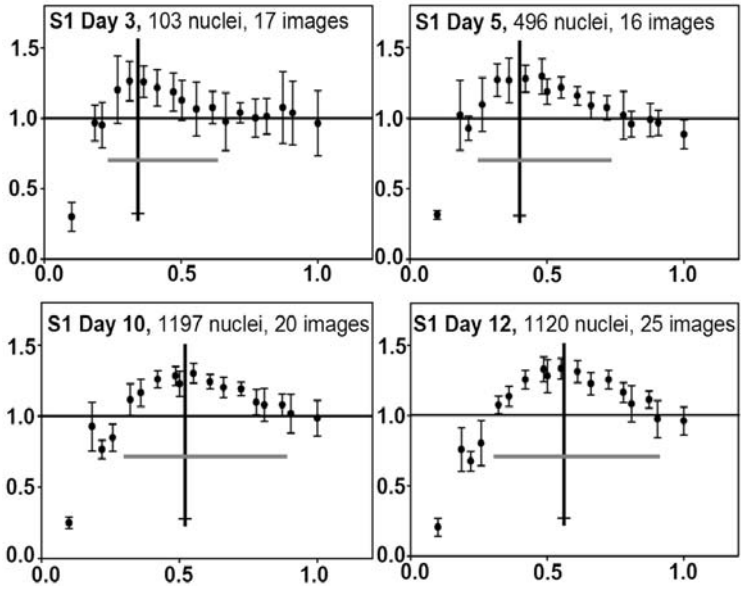


Figure 4. Average relative density of local bright features of NuMA during acinar morphogenesis. S1 cells were cultured in 3D for 3, 5, 10 and 12 days. Plots represent the relative density of NuMA bright features extracted by LBF analysis (ordinate) of a population of nuclei as a function of the relative distance from the perimeter (0.0) to the center (1.0) of the nuclei (abscissa) for each time point. Vertical lines (black) represent the location of the peak of bright feature density in the nucleus. Horizontal lines (gray) represent the extent of nuclear volume with densities of bright features above the average. The number of days the cells were in culture and the number of nuclei analyzed are indicated above each corresponding graph. Bars represent the standard deviations of the relative density of NuMA bright features calculated from multicellular units of the same phenotype, on a per image basis.

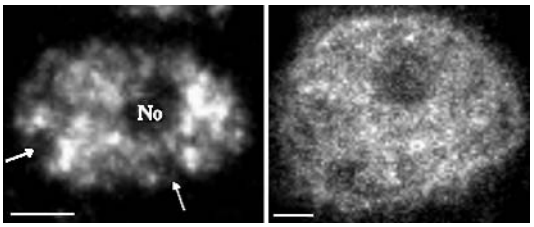


Figure 5. **Immunostaining of NuMA in acinar S1 and malignant T4-2 HMECs.** Image slices of an S1 nucleus (left) and T4 (right) nuclei fluorescently stained with anti-NuMA antibodies from cells cultured in 3D for 10 days. Each image shows a single confocal section through the approximate center of the nucleus. Arrows indicate typical empty areas in NuMA staining at the periphery of the nucleus of acinar cells. The bar indicates 2 microns. No=nucleolus

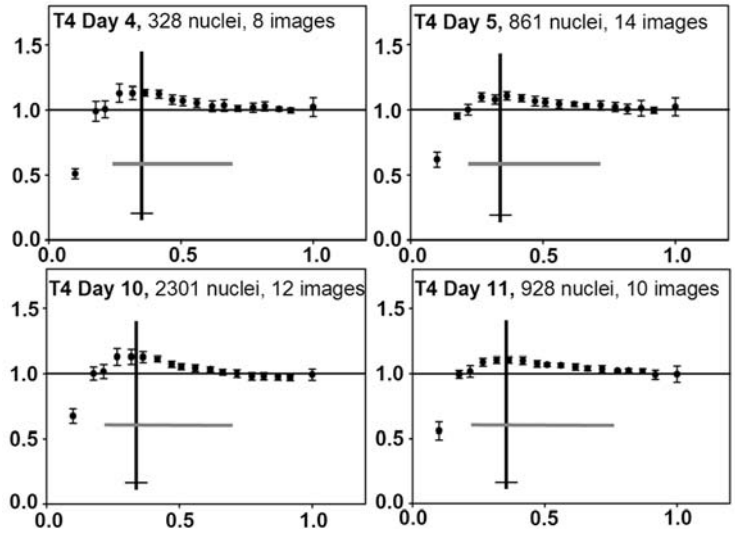


Figure 6. Average relative density of local bright features of NuMA during tumor development. T4-2 cells were cultured in 3D for 4, 5, 10 and 11 days. Plots represent the relative density of NuMA bright features extracted by LBF analysis (ordinate) of a population of nuclei as a function of the relative distance from the perimeter (0.0) to the center (1.0) of the nuclei (abscissa) for each time point. Vertical lines (black) represent the location of the peak of bright feature density in the nucleus. Horizontal lines (gray) represent the extent of nuclear volume with densities of bright features above the average. The number of days the cells were in culture and the number of nuclei analyzed are indicated above each corresponding graph.

Since in non-differentiated proliferating S1 cells NuMA distribution is more diffuse than in growth-arrested (day 5 of 3-D culture) and fully differentiated (day 10 of 3-D culture) acinar cells, we asked whether the diffuse distribution of NuMA was a characteristic of a cell population actively proliferating, regardless of the origin, non-neoplastic or malignant, of the cells. Indeed, the overall distribution of NuMA in proliferating tumor cells did not seem visually different from that observed in proliferating non-neoplastic S1 cells. To quantitatively assess our visual observation, we applied the LBF analysis to 3D cultures of T4-2 cells at days 4, 5, 10 and 11.

During this culture period T4 cells formed disorganized tumor-like nodules of increasing sizes. In contrast to non-neoplastic S1 cells, the radial-LBF analysis showed a fairly flat distribution of NuMA bright features in malignant cells, regardless of the number of days in 3-D culture (Figure 6). Thus, in agreement with the fact that there were no significant alterations in the phenotype of tumor nodules, except for their size, during 11 days of 3D culture, there was no apparent change in the density of NuMA bright features in the nuclei of tumor cells during this culture period.

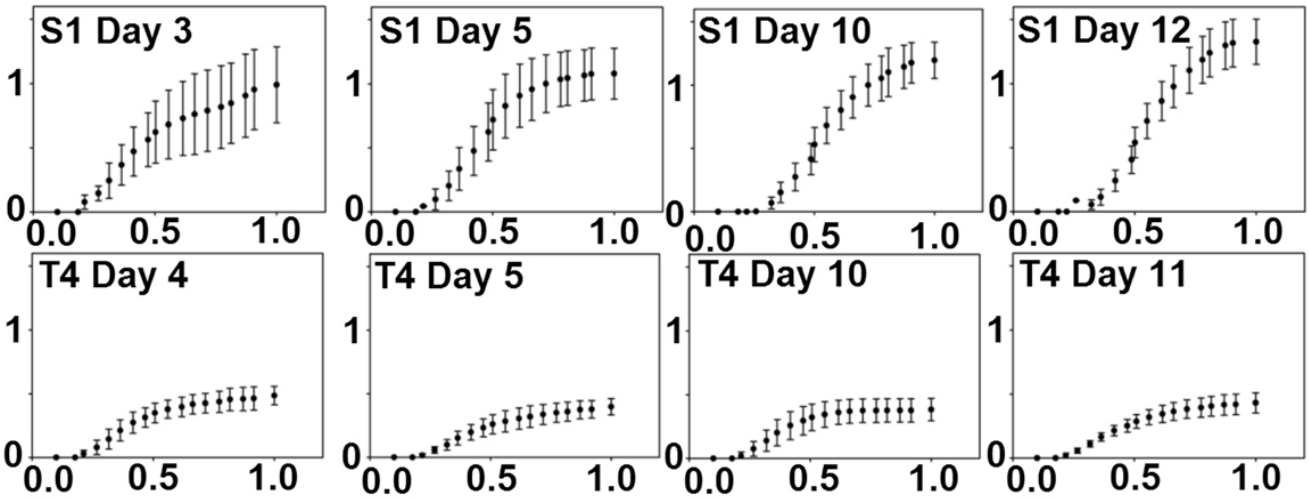


Figure 7. The relative density of local bright features of NuMA is different between non-neoplastic cells and malignant cells regardless of the differentiation stage. Cumulative plots of the relative density of NuMA bright features above unity (ordinate) at different time points of 3D culture of S1 and T4-2 cells as a function of the relative distance from the perimeter (0.0) to the center (1.0) of the nucleus (abscissa). Cumulative plots for S1 and T4-2 cells were prepared from the relative density data shown in figures 2 and 4a, respectively. Bars represent the standard deviations of the relative density of NuMA bright features calculated from multicellular units of the same phenotype, on a per image basis.

The distribution curves of the density of bright features of NuMA in T4 cell nuclei did not clearly show a peak at any of the time points, suggesting that there was a difference in NuMA distribution not only between malignant T4 cells and acinar S1 cells, but also between malignant T4 cells and proliferating S1 cells. To better visualize the differences in the distribution of the bright features of NuMA for the different phenotypes and time points described above, we plotted the cumulative density of NuMA bright features as a function of the distance from the nuclear boundary. The cumulative plots unambiguously show that the distribution of the bright features of NuMA is consistently similar for the different culture time points of the malignant T4 cells and that such a distribution is remarkably different from any of the stages, including proliferation, of acinar morphogenesis (Figure 7).

C.3. LBF analysis of H4-K20m measures different distributions depending on the mammary tissue phenotype.

Encouraged by the success of the difficult analysis of NuMA distribution, we went on to analyze the distribution of H4-K20m, for which we had observed a difference in the staining pattern when comparing S1 acinar cells and T4 malignant cells (Figure 8).

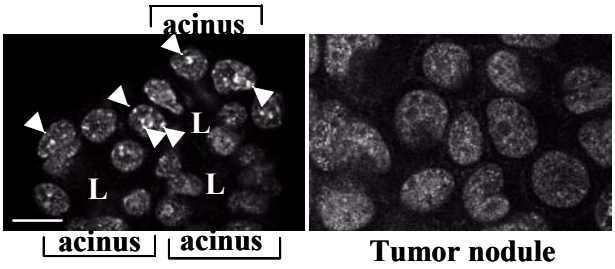


Figure 8. **H4-K20m distribution depends on breast phenotype.** Left image: immunostaining for H4K20m in acini formed in 3D culture (3 acini are shown). Arrows point to large foci of H4K20m. L= lumen. Right image: immunostaining for H4K20m in tumor nodules formed in 3D culture (a portion on a nodule is shown). Size bar= 10 microns.

Notably, acinar cells seemed to possess larger H4-K20m staining foci compared to malignant cells. To quantify the nuclear distribution of this protein we have developed a foci-counting algorithm which is based on our LBF technique. In this foci-counting-LBF analysis, a nuclear segmentation mask is first generated from the total-DNA image. Bright foci were then detected and isolated by the LBF analysis from the H4-K20m image and scored based on their relative brightness and relative size. Briefly, the score was generated from the LBF image convolved with a narrow Gaussian function.

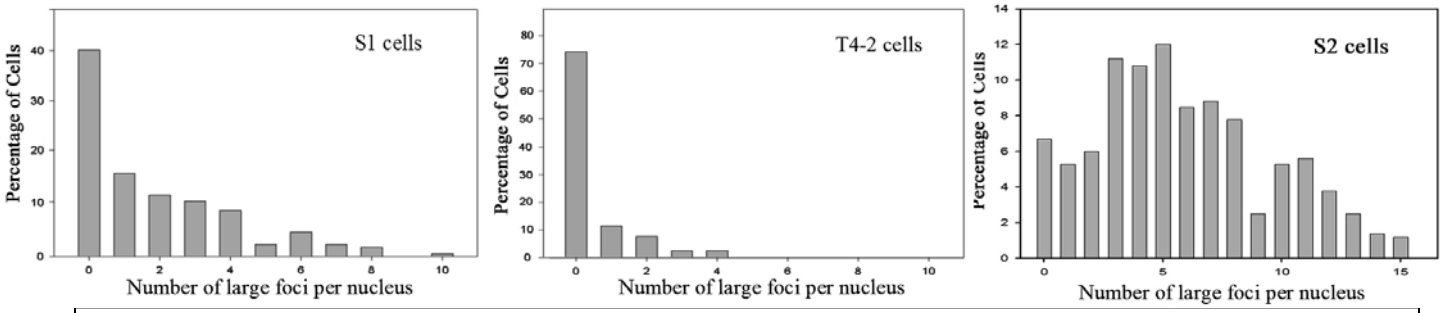


Figure 9. **Histogram of the number of large H4-K20m foci per nucleus in the S1 acinar cell population, T4 malignant cell population and S2 premalignant cell population.** S1, S2 and T4 cells were cultured in 3D for 10 days and immunostained for H4-K20m. The number of foci per nucleus was determined by our foci-counting-LBF analysis. Histograms of the number of foci per nucleus clearly show that large H4-K20m foci are much more abundant in nonmalignant S1 cells than malignant T4 cells. In this revised submission we show data for premalignant S2 cells which are characterized by a high number of large foci.

This manipulation reduced the foci brightness as an inverse function of the foci size. Thus, the local brightness maximum in the blurred-LBF image is a combined measure of the foci brightness and size. The histogram of foci "score" for all nuclei within an image, revealed a large number of small foci but also larger foci at much lower frequency. This result was in agreement with visual analysis of the images. Large foci were defined as having scores larger than two standard deviations away from the histogram mean. To refine the analysis of bright feature distribution, the number of large foci in each nucleus was counted (Figure 9). For a total of 167 nuclei of S1 cells, a total of 311 large foci were counted. Interestingly, 40% of the S1 nuclei were devoid of large foci. In the nuclei that did have large foci there were on average 3.1 large foci per nucleus. For 77 nuclei of T4-2 cells, 35 large foci were counted. 75% of T4 nuclei were devoid of large foci. In the remaining nuclei there were an average of 1.8 large foci per nucleus. The *p* value of the two probability distributions (S1 and T4-2 cells) obtained from the H4K20 analysis was calculated using a Z-test, which compares the significance of the difference between the population means (in this case it is the mean number of large foci per nucleus). For the total number of nuclei the mean number of large foci was 1.86+/-2 for S1 cells and 0.45+/-0.8 for T4 (Z>7 and

p<0.001). Another way to assess if the two distributions are significantly different is to consider the significance of finding a single nucleus with three large foci within the same cell population. The p-value for n=3 in the T4 distribution was p=0.05 (for a single nucleus with 3 or more foci), whereas the p-value for n=3 in the S1 distribution was p=0.4 (not significant), as should be the case. The analysis was on per nucleus basis and nuclei with and without large foci could be easily individually isolated.

We then went on to analyze and show that the nuclear organization of H4-K20m in premalignant S2 cells is markedly different from that in either nonmalignant S1 cells or malignant T4 cells. Figure 9, right panel, shows the percentage histograms of the number of large H4-K20 foci per nucleus for 568 nuclei imaged in multiple S2 spheroids. The data shows that only a small percentage (<10%) of S2 nuclei are without large foci and that on average S2 nuclei have 6.3+-3 large foci. These new data support our hypothesis by showing that the nuclear distribution of H4-K20m is an excellent marker of non-neoplastic, pre-malignant and malignant phenotypes.

C.4 Clustering the radial distribution feature (radial-LBF) of Non-neoplastic S1 and Malignant T4 cells

To this point the results of our LBF analysis have been reported on an average basis. Although this is a powerful way to understand the link between nuclear organization and cell phenotype, it does not allow analysis on a per cell basis. To achieve this we have developed a spectral clustering[29-31] method to group individual nuclei based on the results of their nuclear organization. To demonstrate this capability, we have grouped thousands of non-neoplastic S1 nuclei, based on the radial-LBF analysis results of their NuMA distribution (shown in Figure 4), and analyzed the statistical link between these clusters and the cell phenotype.

The method represents the radial-LBF analysis results or each nucleus in a high-dimensional space and computes the Euclidian pair-wise separation for all nuclei in that space. The inverse of these distances is used to calculate the normalized "similarity scores" to describe how similar two nuclei are based on their NuMA distribution. The spectral clustering analysis then uses a hierarchical bipartition scheme to group nuclei into similar clusters based on the "similarity scores". In each iteration, the spectral clustering approach finds an optimal partition that can minimize the similarity between clusters and at the same time maximize the similarity within each cluster. The iteration stops when a metric combining these two criteria converges. Note that this entire procedure is solely based on the NuMA distribution and does not use prior knowledge about the cell phenotype.

Initially we applied this cluster analysis to group 2673 non-neoplastic S1 cells cultured for 3, 5,10 and 12 days into four clusters. The overlaid plot of the NuMA distributions for each cluster is shown by the graphs in the first row of the table (Figure 10). To analyze the statistical link between these clusters and the cell phenotype, we computed the percentage of nuclei falling into each cluster for S1 cells cultured for 3, 5, 10, and 12 days, respectively. The result is shown in the second to fifth row of the table. The results clearly show that: 1) the fraction of nuclei in cluster 1 decreases with the number of days in culture while the fraction of nuclei in cluster 3 and 4 increases; 2) the maximum fraction of nuclei at Day 3 (61%), Day 5 (46%), Day 10 (27%), and Day 12 (34%) fall into cluster 1, 1, 3, 4 respectively; 3) most of the nuclei (over 90%) at Day 3 (corresponding to the proliferation stage) belong to clusters 1 and 2, while most of the nuclei (about 66%) at Day 12 (corresponding to full differentiation into acini) belong to clusters 3 and 4; 4) at Day 12 the fraction of nuclei within each cluster increases from cluster 1 to 4 while at Day 3 the fraction decreases. Plotting the percentage of nuclei falling into each cluster as a function of clusters, we generated a histogram indicating how nuclei are distributed across clusters for each day, as shown in the last column of the table (Figure 10). The results clearly show that there is a statistical correlation between the histogram patterns and the phenotype of cells. In fact, the many clear trends revealed by the clusters generated from spectral clustering method would not have been obtained unless the LBF analysis of NuMA distribution worked correctly and constituted a good marker of cell phenotype.

We then extended this analysis by including all the malignant T4-2 cells. Cluster analysis was performed on the 2673 nonneoplastic S1 cells along with 4418 malignant T4-2 cells. Figure 10 upper right, shows the results of clustering the radial distribution of NuMA for these 7091 cells into 20 clusters. The clusters have been arbitrarily ordered. Figure 10 lower-panel shows the corresponding "cluster histograms" which link the nuclei within each cluster to their phenotype. As expected the "cluster histograms" show clear differences between S1 cells cultured for different days (which correspond to different stages in the differentiation process including proliferation, growth arrest and progressive polarization).

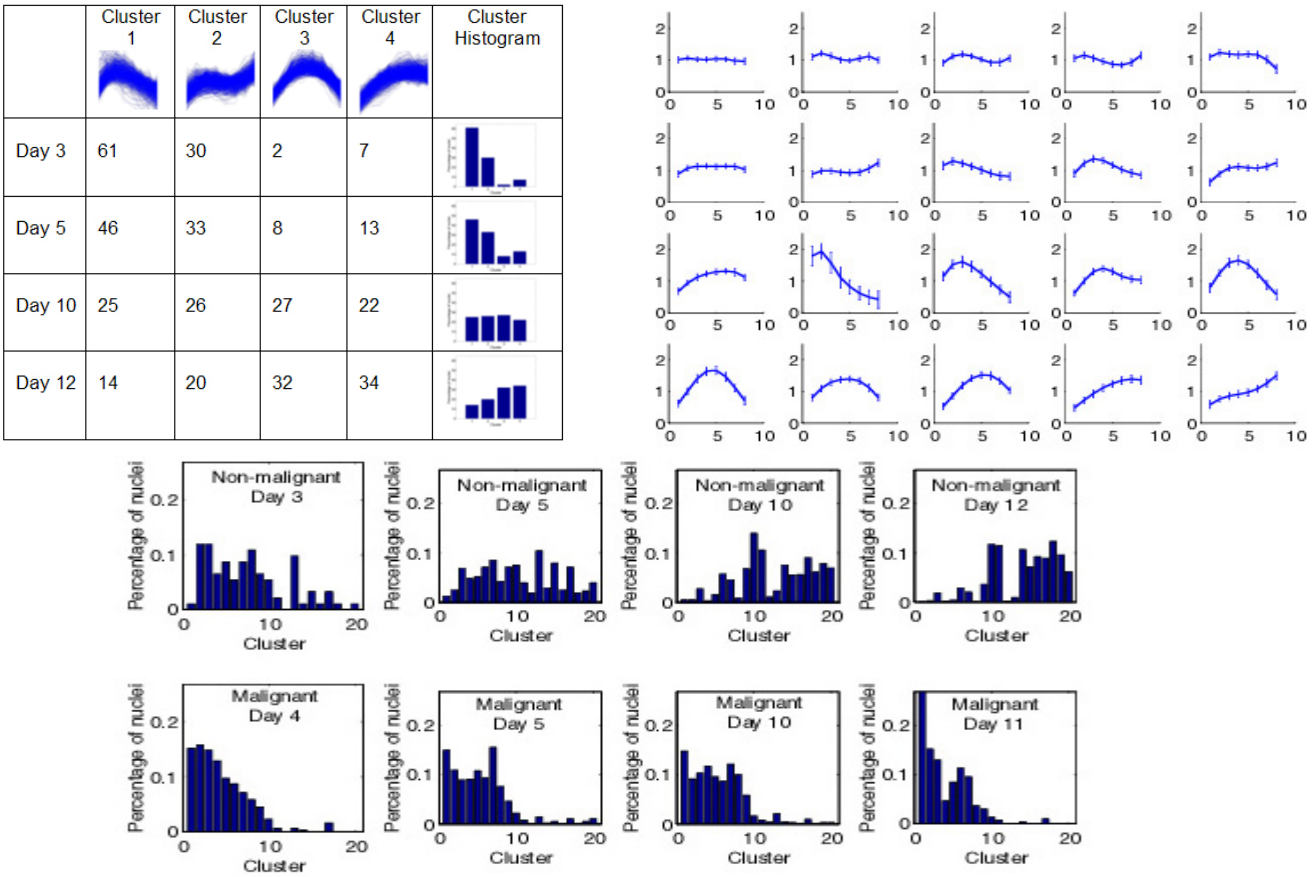


Figure 10. **Clustering Analysis and Classification Results**. The table shows results of the cluster analysis for non-neoplastic S1 cells cultured for 3, 5, 10 & 12 days as the cells differentiate into acini. Cluster analysis grouped the radial-LBF analysis of NuMA from 2673 nuclei into 4 distinct clusters (Table,Top Row). To understand the link between the clusters and the phenotype, the percentage of nuclei from each day which fall into each cluster is calculated (Rows 2-5) and plotted as a histogram (Table,Right Column). The results show a clear link between clusters and cell phenotype. We then extended our cluster analysis to group the radial distribution of NuMA in 2673 nonneoplastic S1 cells in combination with 4418 malignant T4-2 cells into 20 clusters (upper right). Each cluster shows the mean and standard deviation of the relative density of NuMA bright features (ordinate) plotted against the normalized distance from the perimeter of the nucleus to its center (abscissa). We then create the "cluster histograms" for the 7091 cells. Each graph shows the percentage of nuclei, of a particular cell type and number of days in culture, that belong to each of the 20 clusters (lower panel). The top four histograms show nonneoplastic cells cultured for 3, 5, 10 and 12 days and the bottom 4 histograms show malignant T4-2 cells cultured for 4, 5, 10 and 11 days. These cluster histograms clearly reveal the reorganization of NuMA as nonneoplastic S1 cells differentiate, and show no significant reorganization of NuMA as malignant T4-2 cells proliferate. Using this scheme, we were able to predict the number of days in culture of nonneoplastic S1 cells with an accuracy of 74±2%. We were able to predict if nuclei were nonneoplastic or malignant with an accuracy of 95±3%.

Importantly these results also show that there is little change in the distribution of NuMA in T4-2 cells with the number of days in culture. In the case of T4-2 cells, the number of days in culture (from day 1 to 11 here) correspond to increasing sizes of tumor nodules but it is not accompanied with significant changes in the phenotype or cell behavior. These new results support our overall goal of developing an image-based technology capable of turning high resolution fluorescence images of human mammary epithelial tissue into tissue-maps which report the probable nonneoplastic, premalignant and malignant phenotype at cellular resolution. To test the sensitivity of this approach, we developed a classification method which was trained on a subset of the results chosen randomly from the 7091 nuclei. We then classified the remaining nuclei, by considering the organization of NuMA in the nucleus in question along with its nearest neighbors, and correlated the classification result with the known phenotype of each nucleus. By using this scheme, we were able to predict the number of days in culture (which correspond to specific stages of differentiation) of nonneoplastic S1 cells with an accuracy of 74±2%. We were able to predict if nuclei were nonneoplastic or malignant with an accuracy of 95±3%.

C.5: Cluster Analysis Reveals Sensitivity on a per Nucleus Basis

One of our next challenges was regarding the sensitivity of this approach for identifying phenotypically different nuclei on a per cell basis. Since the proposal submission we have extended our cluster analysis and classification methods to better understand the sensitivity of our technique. Specifically, **we wanted to know, with what significance could we classify <u>individual</u> nuclei into their known phenotypes, based on the radial distribution of NuMA?** Using the cluster analysis results presented in Figure 10, we have developed a classification method which was trained on a subset of the results chosen randomly from 7334 non-neoplastic and malignant nuclei. We then classified the remaining nuclei, not used as part of the training set. Nuclei were classified by considering the organization of NuMA in the nucleus in question along with its nearest neighbors within the tissue context, and the classification result was correlated with the known phenotype of each nucleus.

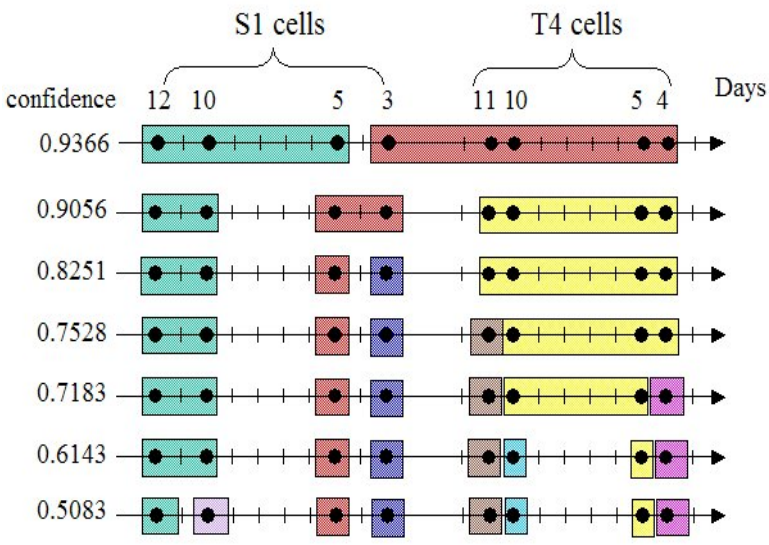


Figure 11. Hierarchical representation of the probability of classifying individual cells by phenotype based on nuclear organization. This figure shows a hierarchical tree structure indicating the confidence of predicting the known phenotype of an individual cell and its nearest neighbors based on the nuclear radial distribution of NuMA. The left column shows the fractional probability of being able to classify nuclei into the groups, indicated by different colors along each row. For example, the 2nd row shows that with 90.56% confidence we are able to correctly classify individual nuclei as corresponding to either, a differentiated non-neoplastic S1 cell (Days 10 and 12), a proliferating S1 cell (Days 5 and 3) or a malignant T4-2 cell.

The results (Figure 11) show that with more than 93% accuracy we can distinguish individual non-neoplastic S1 cells at 5 days of culture (corresponding to proliferation stage), and 10 and 12 days of culture (corresponding to differentiated stages) from all malignant T4-2 cells. At greater than 90% accuracy we can distinguish differentiated S1 cells (10 and 12 days in culture), proliferating S1 cells (3 and 5 days in culture) and T4-2 cells. The results also show that we cannot distinguish the malignant T4-2 cells cultured for different numbers of days

with high certainty. This was expected from the results presented in Figure 6 of the proposal, which shows little change in the nuclear organization of proliferating T4-2 cells with the number of days the cells were in culture. These results indicate that the phenotype of individual epithelial cells can be predicted with high sensitivity based on the organization of the nuclear protein NuMA.

This work is included in a manuscript in preparation[42].

C.6 The distribution of NuMA in S2 spheroids reveals at least two separate phenotypes.

One of the goals of this work is developing the capability of subdividing a tissue into different phenotypes based on the nuclear organization. Since the proposal was submitted, we have used cluster analysis of the distribution of NuMA to show that there are at least two phenotypically different groups of nuclei in premalignant S2 spheroids which correlate with spheroid size. To do this, the radial distribution of NuMA was quantified for multiple premalignant S2 spheroids of varying size. The radial-LBF distributions of all nuclei were then randomized and cluster analysis performed to group the distributions into an arbitrary number of clusters. In Figure 12 we show the results for partitioning all the radial-LBF distributions into eight clusters. The graphs show the mean and standard deviation of the distribution in each cluster. By grouping nuclei according to the spheroid they came from, we are able to create cluster histograms for each spheroid (Figure 12, lower panel). We have called these 'cluster histograms' the "**protein distribution profiles**" of each spheroid. They clearly show that the distribution of NuMA in the smaller spheroids is different than that in larger spheroids.

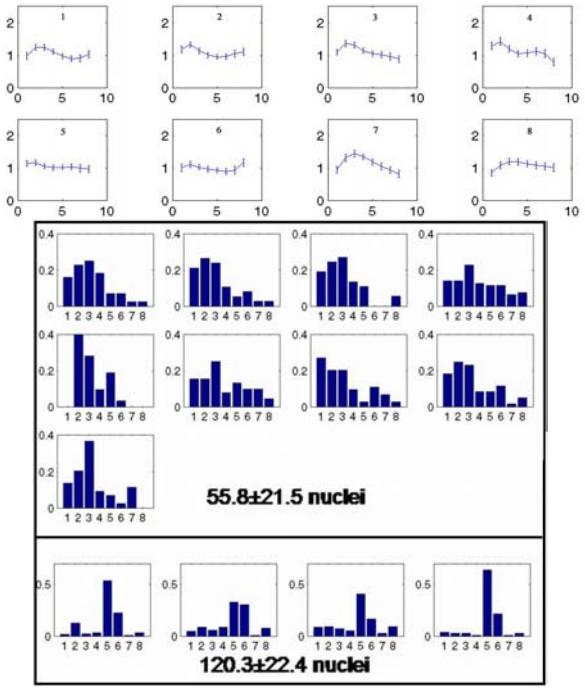


Figure 12. Distribution of NuMA in S2 spheroids. For the cluster analysis of distribution of NuMA in S2 spheroids, the radial-LBF distributions of all nuclei were randomized and grouped into an arbitrary number of clusters. Here we show the results for eight clusters (upper). Each cluster shows the mean and standard deviation of the relative density of NuMA bright features (ordinate) plotted against the normalized distance from the perimeter of the nucleus to its center (abscissa). The lower panel shows the resulting cluster histograms of 13 spheroids. The cluster histograms reveal two distinct phenotypes which correlate to the spheroid size (as measured by the number of nuclei in that spheroid). The top nine spheroids contained 55.8±21.5 nuclei and the bottom four spheriods contained 120.3±22.4 nuclei. Each graph shows the percentage of S2 nuclei that belong to each of the eight clusters.

C.7 Novel Visualization Tools

C.7.1 Novel Visualization Tools for Rendering Nuclear Distribution and Tissue Morphology in Three Dimensions

Finally, in collaboration with Gunther Weber and Bernd Hamann of the Visualization and Graphics Research Group, at the University California, Davis (see: http://graphics.cs.ucdavis.edu/ and http://graphics.cipic.ucdavis.edu/~hamann/index.shtml), significant progress has been made towards creating visualization tools needed to display the *phenotype tissue-map*. Indeed, we have developed tools to visualize the nuclear distribution of proteins with the tissue morphology in three dimensions. Figure 13 shows a 3D view of the three differentiated acini from Figure 2. The tool allows simple three dimensional rendering of the positions and volumes of the nuclei, with the image of the protein of interest (Figure 13, left panel). The tool allows the user to visualize the image at any rotation and to step through the image at any angle. By doing this, the nuclear distribution within each nucleus is clearly revealed (Figure 13, middle) and this view can also be rotated to any angle (Figure 13, right panel). Our goal is to link this visualization tool directly to the distribution and clustering analysis and to allow the subsequent results to be mapped back onto the tissue morphology. Tools will also be developed to allow the user to visually display nuclei within the same cluster-group, to subdivide nuclei by cluster-group, to identify nuclei with the same or similar distributions as a chosen nuclei, or to display nuclei in terms of a continuous variable formed in multiple ways from the multidimensional feature space in which the nuclear distributions are represented.

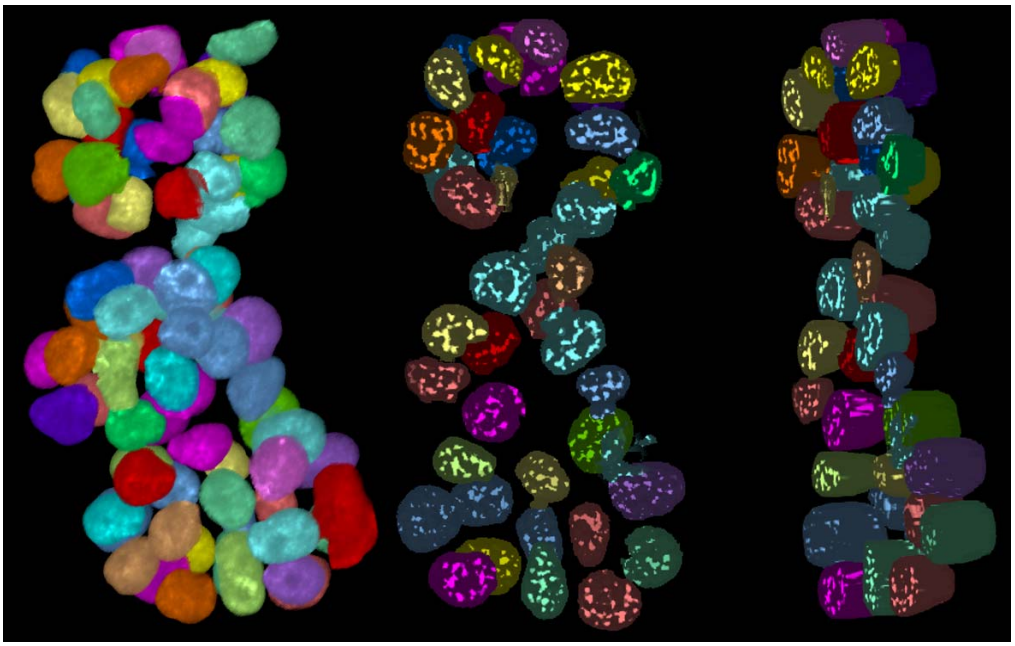


Figure 13. Visualization of Tissue Morphology and Nuclear Distribution in Three Dimensions. This figure shows views of the visualization tool we have developed to render nuclear positions and volumes with the nuclear protein of interest in 3D (left panel). The tools allow visualization at any plane within the nuclei (middle panel) and rotations of the 3D image at any angle (right panel).

C7.2 3D Visualization Tool Renders Distributional Analysis Results onto the Tissue Morphology

We have extended our three dimensional (3D) visualization tool. We have developed the capability of rendering analysis results on the nuclear segmentation mask. Figure 14, shows different views of a single spheroid of premalignant S2 cells. Nuclei shown in red have greater than 10 bright foci of heterochromatin marker H4K20m and nuclei shown in green have less than 10 bright foci of H4K20m, as determined by our foci-counting LBF-analysis. By rendering these analysis results onto the tissue morphology, one can quickly see that the two groups of nuclei are not randomly distributed but rather tend to be grouped together, within the spheroid.

The ability to render analysis results onto the tissue morphology, gives a clear indication of the arrangement of nuclei with different internal organization within the tissue. This allows a completely novel view of a tissue and is a step towards our longer term goal of developing the proposed ***phenotype tissue-map***.

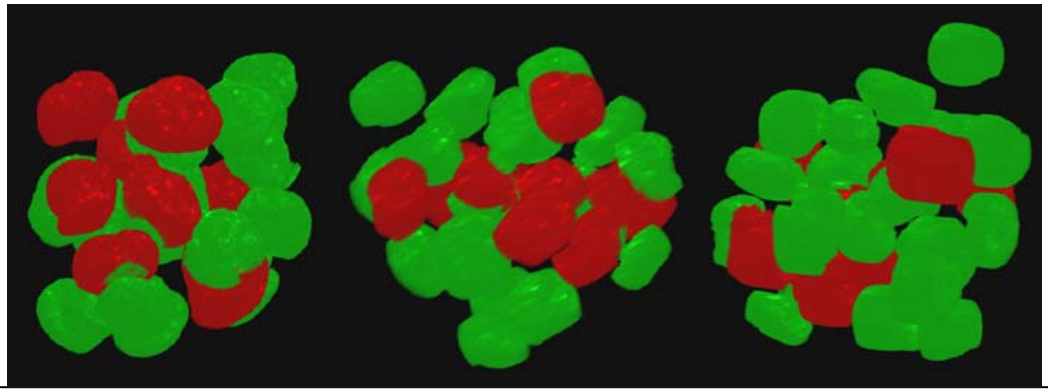


Figure 14**. Rendering analysis results onto the tissue morphology give a completely new view of cellular arrangement**. This figure shows three views of a single spheroid of premalignant S2 cells rendered in 3D to show nuclei with greater than 10 bright foci of H4K20m (red) and those with less than 10 bright foci (green). The spheriod has been rotated in 3D to show different aspects. The ability to render analysis results onto the tissue morphology gives a completely new view of a tissue. In this case one can quickly see that the two groups of nuclei, identified by quantitative analysis of the distribution of a nuclear protein, are grouped together within the spheroid.

C7.3. Analysis of the Radial Distribution of Fluorescently-Stained NuMA in Deparaffinized Sections of Normal Human Mammary Tissue.

One important area of progress was with the use of fixed, paraffin-embedded tissue biopsy sections. In the past year we have started working with normal, premalignant and malignant human breast tissue with a collaboration with Sunil Badve MBBS, MD(Path), FRCPath, who is Director of Research Immunohistochemistry and Assistant Professor in Pathology at Indiana University. Initially we had worked with freshly frozen tissue because it is most amenable to aqueous-based fluorescent staining procedures we use. However, in preparation for our image-based techniques to be more widely available, we needed to circumvent the problem of working with paraffin-embedded tissues. Paraffin embedding followed by hematoxylin and eosin (H&E) staining are standard histological techniques but paraffin has a large autofluorescent signal and needs to be removed to allow standard aqueous-based techniques of fluorescence immunochemistry.

Our current technical analysis shows that a combination of biotin and tyramide gives the brightest specific signal for NuMA and H4K20m on paraffin sections (Figure 15). In the previous years report we not only established our ability to perform fluorescent immunostaining for NuMA on deparaffinized tissue sections but also to extract bright features from the nuclei in the resulting images. In this last year, we have extended our analysis techniques by establishing that they work with these tissue sections. In figure 16, we show images of fluorescently-stained total DNA (left panel) and NuMA (middle panel) which has been overlaid with an outline of the nuclear segmentation mask. The radial distribution of NuMA was calculated for each of the nuclei within the structure. The majority of the nuclei had a NuMA distribution similar to that of differentiated cultured S1 cells (see figure 4); however some nuclei had a different distribution of NuMA. By rendering the analysis results onto the segmentation mask, we are quickly able to see which nuclei have distributions of NuMA similar to differentiated S1 cells (shown in red) and which nuclei have other types of distribution, shown in green (right panel). The lower graph shows the average and standard deviation of the radial distribution of NuMA for the nuclei labeled in red.

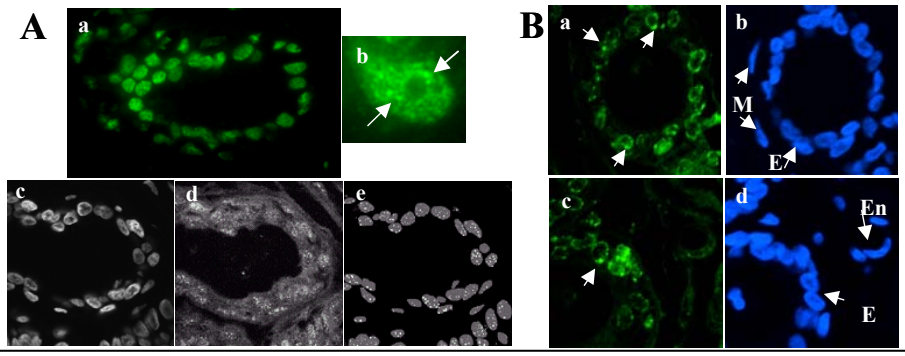
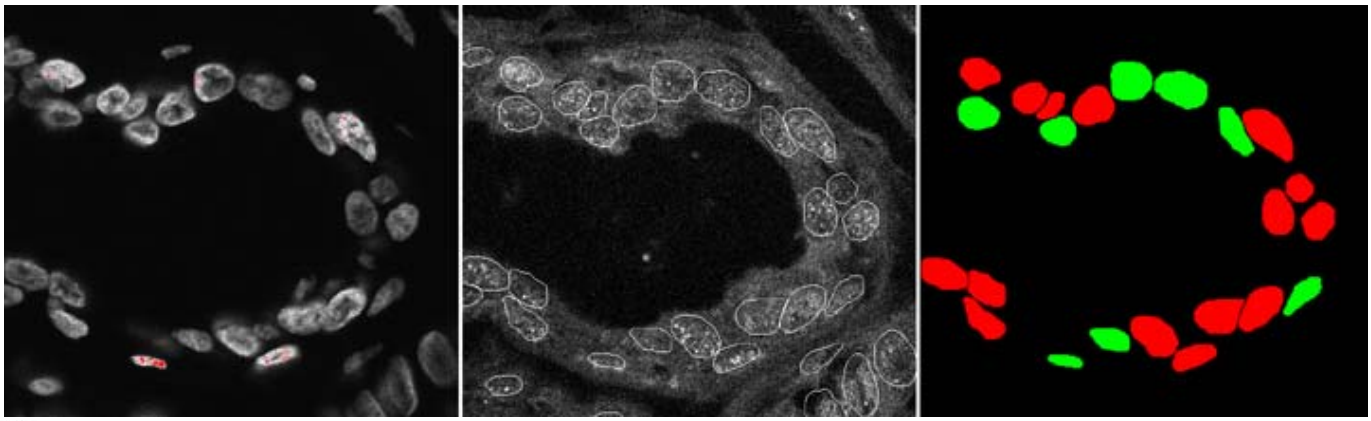


Figure 15. **Fluorescence staining for NuMA and H4K20m on paraffin section of normal breast tissue.** Samples were immunostained for NuMA and H4K20m using streptavidin-biotin-tyramide amplification method. **A**. Low resolution images of NuMA immunostaining (a and d). Arrows indicate the concentration of NuMA bright staining features in mid nucleus (b). DAPI image (c) corresponding to NuMA staining (d). Overlay of the LBF image of NuMA and the segmentation mask (for c and d) is shown in (e). **B.** Low resolution images of H4K20m immunostaining (a and c) and DNA counterstain (b and d). Arrows indicate large H4K20m foci in a and c. E= luminal epithelium; M= myopepithelial cells; En= endothelial cells.



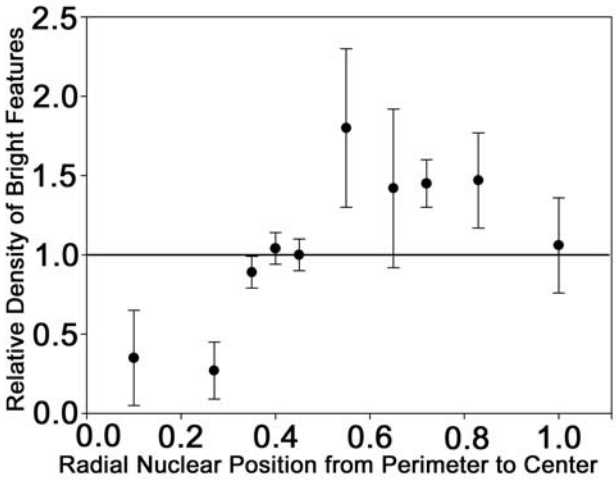


Figure16**. Radial LBF analysis of NuMA in deparaffinized sections of normal human mammary tissue**. A section from a biopsy of normal looking mammary tissue was deparaffinized and the total DNA (top left) and NuMA (top middle) were fluorescently stained. The nuclear segmentation mask produced from the DNA image has been overlaid on the NuMA image. The radial-LBF analysis was performed for epithelial cells within the ductal-like structure and the results applied onto the segmentation mask (top right). Nuclei labeled red have a NuMA distribution similar to that of differentiated cultured S1 cells, and nuclei with labeled in green have a different NuMA distribution. The graph (left) shows the average and the standard deviation of the radial distribution of NuMA for the nuclei in red.

These new results support our overall goal of developing an image-based technology capable of turning high resolution fluorescence images of human mammary epithelial tissue into tissue-maps which report the probable non-neoplastic, premalignant and malignant phenotype at cellular resolution.

D. Future Aims

By understanding the links between nuclear protein organization and the functioning and phenotype of cells, it is the future goal of this research to develop image analysis technology to produce a method capable of turning high resolution fluorescence images of human mammary epithelial tissue into tissue-maps which report the probable non-neoplastic, premalignant and malignant phenotype at cellular resolution. The aim is to aid the treatment decision process of breast cancer patients by providing pathologists with a *phenotype tissue-map*, based on nuclear protein organization, to aid and support the histological classification of biopsied breast tissue. We believe that such image analysis capability will have an enormous impact in aiding the work of both cancer biologists and oncologists who study and classify cells by epigenetic means. We identify the following as the next critical steps in the future of this work.

Aim 1: To develop and assess the capability of our Local Bright Feature (LBF) analysis methods to identify mixed populations of phenotypically different human mammary epithelia. Currently, our LBF analysis methods have been developed using phenotipically homogenous populations of both nonneoplastic and malignant cells. Our future work will demonstrate that we can detect phenotypically different cells within a heterogeneous population. We will expand our previous work by using cluster analysis on the measured distributions of NuMA and H4-K20m to identify and group phenotypically similar cells from heterogeneous populations of premalignant cultured cells, mixed populations of premalignant and malignant cultured cells and premalignant human mammary tissue biopsies.

Aim 2: To develop and assess the capability of our Local Bright Feature (LBF) analysis methods to automatically analyze heterogeneous populations of human mammary epithelia. Currently our image analysis techniques have been developed to automatically analyze the nuclear organization in homogeneous populations of epithelia with nuclei of similar volume and "well behaved" shape. One major challenge of working with heterogeneous populations of epithelia is identifying those cells belonging to common tissue-structures and working with nuclei with variations in size and shape. In this phase of future work, spatial statistical methods will be developed to identify neighbouring cells comprising a common tissue structure and novel improvements will be made to our LBF analysis to maintain automation and analysis accuracy when dealing with morphologically heterogeneous populations of epithelia.

Aim 3: To develop and assess the capability of an image-based classification system, that uses the nuclear organization of specific proteins to define new sub-classes of various graded lesions. In this phase of the future work we will use the cluster analysis results from non-neoplastic, premalignant and malignant cells to define a set of features that characterize these cell phenotypes. Using these we will develop a classification system which will assign the probable tissue phenotype at cellular resolution. This *phenotype tissue-map* technology could be tested on needle-core biopsies of a variety of premalignant tumors with the aim of defining sub-classes of graded lesions. The results will be correlated with the histopathology of the initial needle-core and the follow-up surgical biopsies with the hope of predicting more aggressive phenotypes missed by the initial screen.

KEY RESEARCH ACCOMPLISHMENTS:

During the course of this project significant accomplishments have been made.

1) Novel automated image-analysis software has been produce which is able to quantify the nuclear distribution of labeled proteins on a per nucleus basis from three dimensional fluorescence images of cells and tissue.

2) The image analysis has been used to quantify the nuclear distribution of several nuclear proteins and demonstrate that the nuclear organization of specific chromatin-related proteins is a marker for the phenotype of the cells and tissue. This results is potentially translatable to the clinic.

3) We have developed cluster analysis methods to group cells based on the nuclear distribution of specific proteins, and to provide a "cluster histogram" which links the distribution analysis to the phenotype of the cells.

4) Through a collaboration with the UC Davis Visualization Group, novel visualization tools have been developed to view cells and tissue in their three dimensional context and to map and view the distributional analysis results in their tissue context. This ability is providing a new view of the tissue at the cellular level.

5) Through a collaboration with the Indiana University Cancer Center we have started to apply our image analysis methods to biopsied human breast tissue.

6) We have established protocols for deparaffinizing embedded tissue, fluorescently labeling nuclear proteins and have established that our image analysis techniques work in biopsied paraffin-embedded human breast tissue.

REPORTABLE OUTCOMES:

Manuscripts:

1) **Drs. Knowles** and Lelièvre have published part of the work from this project in PNAS[24].

2) Drs. Luengo Hendriks and **Knowles** have a letter *in press* at the Journal of Microscopy[43].

3) Drs. Long and **Knowles** have a paper in preparation[42].

Invited Presentations:

1) **Dr. Knowles** was invited speaker at the LBNL Friends Tour, 30th July 2005

2) **Dr. Knowles** was invited and paid speaker the IEEE Computational Systems Bioinformatics Conference August 8-11 2005, Stanford University.
David W. Knowles 2005 *Novel Visualization and Quantitative Analysis Methods in BioImaging*

3) **Dr. Knowles** was invited, paid speaker at the Basic Medical Sciences Departmental Seminar Series, Purdue University, April 2006

4) **Dr Knowles** is invited Program Committee Member for the 2006 Workshop on Multiscale Biological Imaging, Data Mining and Informatics, Santa Barbara, CA, USA Sept 7-8 2006

Abstracts / Presentations:

1) Fuhui Long, Hanchuan Peng, Damir Sudar, Sophie Lelièvre, **David W. Knowles** 2005
Cell Phenotype Classification Based on 3D Cell Image Analysis
csbw, p. 374, 2005 IEEE Computational Systems Bioinformatics Conference - Workshops (CSBW'05), 2005

2) **David W. Knowles**, Sophie A. Lelièvre, Carlos Ortiz de Solórzano, Stephen J. Lockett, Mina J. Bissell, Damir Sudar 2001
Quantitative model-based image analysis of NuMA distribution links nuclear organization with cell phenotype
Microsc. Microanal. 7:578-579

3) **David W. Knowles**, Sophie A. Lelievre, William S. Chou, Aaron Lee, Wanling Wen, Carlos Ortiz de Solorzano, Mina J. Bissell, Stephen J. Lockett 2000

Quantitative model-based image analysis of sub-visual changes in NuMA distribution links nuclear organization with cell phenotype Biophys. J. 78:250A 2000

Licenses: Software:
LBNL under their DOE contract has licensed the software developed during this project.

Funding Applied for:
1) Drs. Knowles has a pending resubmitted NCI R33 proposal which received a score of 153 and is currently being consider by programmatic management.
2) Drs. Knowles has submitted an internal LBNL research proposals in April of this year to fund the visualization tool development for this project.
3) Dr. Knowles applied for a 2005 DOD BCRP Idea Award which was not funded.
4) Dr. Fuhui Long applied for a 2005 DOD BCRP Postdoctoral Fellow Award which was not funded.
5) Dr. Fuhui Long applied for a 2005 California BCRP.
5) Dr Knowles applied for a 2005 NIH/NCI R33 award which scored 168 but was not funded.
6) Dr. Knowles resubmitted the NIH/NCI R33 proposal in 2006. The proposal scored 153 and is currently in programmatic consideration.
7) Dr. Knowles has applied for an internal LBNL, LDRD award to fund the visualization component of this project. We have succeed to make the second round of this competition. The outcome will be known in October 2006.
8) Drs. Knowles and Lelièvre applied to the NIH for a 2004 SPORE Award as part of a collaborative group centered at Indiana University.
9) Drs. Knowles and Lelièvre applied for 2003 Friends You Can Count On Award.
10) Dr Lelièvre applied for a 2005 DOD BCRP Scholars Award
11) Dr. Denise Munoz has applied for a 2006 DOD BCRP Era of Hope, Postdoctoral Fellow Award, with Dr. Knowles as a collaborator.

Employment Opportunities:
Dr Fuhui Long applied for and accepted a **Staff Scientist** position with Prof. Gene Myers at the Janelia Farm Research Campus, Howard Hughes Medical Institute.

CONCLUSIONS:
The ability to quantify the spatial distribution of fluorescent bright cellular features has many biological applications ranging from the study of gene expression and protein movement in live cells, and the exploration of the structural aspects of cell division, to the investigation of the role of nuclear alterations in pathologies[19,20,23,44-48]. We believe that the LBF analysis, which isolates local bright features, and the radial-LBF analysis, which quantifies the distribution of the bright features, are examples of powerful tools capable of measuring differences in the complex distribution of endogenously expressed nuclear proteins from 3D images acquired following simple immunostaining procedures. Radial-LBF analysis has led to findings that strongly support the concept that specific cell and tissue phenotypes are reflected by the organization of nuclear components. These findings underline the importance of reorganization within the nucleus during the differentiation process and the alterations in nuclear organization that may be associated with tumor behaviour.

We are currently developing methods to turn the ability to quantify nuclear protein distributions into a robust classification technique to define the probability that individual nuclei belong to a specific cell phenotype. We are adding new feature-extraction algorithms and focusing on the distribution of several nuclear proteins. We have also started using the LBF-based imaging techniques to investigate the remodeling of protein distributions in normal and diseased cells in clinically-derived tissue samples.

Our goal is to create a new quantitative 3D "view" of cells and tissues, based on the remodeling of nuclear proteins that helps understand the organization of the nucleus and aids in the classification of pathological samples.

Dr. Knowles is deeply grateful for the opportunities provided by this award and the privilege to conduct the research that this award has funded.

REFERENCES:

1. Zink, D., A.H. Fischer, and J.A. Nickerson. 2004. Nuclear structure in cancer cells. *Nat Rev Cancer*. 4:677-87.
2. Lelièvre, S.A., M.J. Bissell, and P. Pujuguet. 2000. Cell nucleus in context. *Crit Rev Eukaryot Gene Expr*. 10:13-20.
3. Dillon, N., and R. Festenstein. 2002. Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Genet*. 18:252-8.
4. Mancini, M.A., B. Shan, J.A. Nickerson, S. Penman, and W.H. Lee. 1994. The retinoblastoma gene product is a cell cycle-dependent, nuclear matrix-associated protein. *Proc Natl Acad Sci U S A*. 91:418-22.
5. Gerlich, D., J. Beaudouin, M. Gebhard, J. Ellenberg, and R. Eils. 2001. Four-dimensional imaging and quantitative reconstruction to analyse complex spatiotemporal processes in live cells. *Nat Cell Biol*. 3:852-5.
6. Lelièvre, S.A., V.M. Weaver, J.A. Nickerson, C.A. Larabell, A. Bhaumik, O.W. Petersen, and M.J. Bissell. 1998. Tissue phenotype depends on reciprocal interactions between the extracellular matrix and the structural organization of the nucleus. *Proc Natl Acad Sci U S A*. 95:14711-6.
7. Antoniou, M., M. Carmo-Fonseca, J. Ferreira, and A.I. Lamond. 1993. Nuclear organization of splicing snRNPs during differentiation of murine erythroleukemia cells in vitro. *J Cell Biol*. 123:1055-68.
8. Beil, M., D. Durschmied, S. Paschke, B. Schreiner, U. Nolte, A. Bruel, and T. Irinopoulou. 2002. Spatial distribution patterns of interphase centromeres during retinoic acid-induced differentiation of promyelocytic leukemia cells. *Cytometry*. 47:217-25.
9. Plachot, C., and S.A. Lelièvre. 2004. DNA methylation control of tissue polarity and cellular differentiation in the mammary epithelium. *Exp Cell Res*. 298:122-32.
10. Gribbon, C., R. Dahm, A.R. Prescott, and R.A. Quinlan. 2002. Association of the nuclear matrix component NuMA with the Cajal body and nuclear speckle compartments during transitions in transcriptional activity in lens cell differentiation. *Eur J Cell Biol*. 81:557-66.
11. Weaver, V.M., C.E. Carson, P.R. Walker, N. Chaly, B. Lach, Y. Raymond, D.L. Brown, and M. Sikorska. 1996. Degradation of nuclear matrix and DNA cleavage in apoptotic thymocytes. *J Cell Sci*. 109:45-56.
12. Zweyer, M., B.M. Riederer, R.L. Ochs, F.O. Fackelmayer, T. Kohwi-Shigematsu, R. Bareggi, P. Narducci, and A.M. Martelli. 1997. Association of nuclear matrix proteins with granular and threaded nuclear bodies in cell lines undergoing apoptosis. *Exp Cell Res*. 230:325-36.
13. Sodja, C., P.R. Walker, D.L. Brown, and N. Chaly. 1997. Unique behaviour of NuMA during heat-induced apoptosis of lymphocytes. *Biochem Cell Biol*. 75:399-414.
14. Abad, P.C., Mian, I.S., Plachot, C., Nelpurackal, A., Bator-Kelly, C. and SA. Lelièvre. 2004. The C-terminus of the nuclear protein NuMA: phylogenetic distribution and structure". *Protein Sci.* (In press; Oct 1, vol. 13).
15. Kaminker, P., Plachot, C., Kim, S.H., Chung, P., Crippen, D., Petersen, O.W., Bissell, M.J., Campisi, J. and S.A. Lelièvre. "Higher order nuclear organization in growth arrest of human mammary epithelial cells: A novel role for telomere-associated protein TIN2". (Submitted).
16. Gil, J., and H.S. Wu. 2003. Applications of image analysis to anatomic pathology: realities and promises. *Cancer Invest*. 21:950-9.
17. Ecker, R.C. and G.E. Steiner. 2004. Microscopy-based multicolor tissue cytometry at the single-cell level. *Cytometry* 59:182-90.

18. Ancin, H., Roysam, B., Dufresne, T.E., Chestnut, M.M., Ridder, G.M., Szarowski, D.H. and J.N. Turner. 1996. Advances in automated 3-D image analyses of cell populations imaged by confocal microscopy *Cytometry* 25:221-34.
19. Irinopoulou, T., Vassy, J., Beil, M., Nicolopoulou, P., Encaoua, D. and J.P. Rigaut. 1997. Three-dimensional DNA image cytometry by confocal scanning laser microscopy in thick tissue blocks of prostatic lesions. *Cytometry* 27:99-105.
20. Lieb, J.D., de Solorzano, C.O., Rodriguez, E.G., Jones, A., Angelo, M., Lockett, S., and B.J. Meyer. 2000. The Caenorhabditis elegans dosage compensation machinery is recruited to X chromosome DNA attached to an autosome. *Genetics* 156:1603-21.
21. Schupp, S., Elmoataz, A., Herlin, P. and D. Bloyet. 2001. Mathematical morphologic segmentation dedicated to quantitative immunohistochemistry. *Anal Quant Cytol Histol.* 23:257-67.
22. Eils, R., and C. Athale. 2003. Computational imaging in cell biology. *J Cell Biol.* 161:477-81.
23. Voss, T.C., Demarco, I.A., Booker, C.F., Day, R.N. 2004. Computer-assisted image analysis protocol that quantitatively measures subnuclear protein organization in cell populations. *Biotechniques*. 36:240-7.
24. Knowles, D.W., Sudar, D., Bator-Kelly, C,. Mina J. Bissell, M.J., Lelièvre, S.A. 2006 Automated local bright feature image analysis of nuclear protein distribution identifies changes in tissue phenotype. *Proc. Natl. Acad. Sci. USA* **103**, 4445-4450
25. Marr, D. 1982.*In* Vision, W.H.Freeman & Company.
26. Castleman K. R. 1996. *In* Digital Image Processing, Prentice-Hall.
27. Russ J.C. 1992. *In* The Image Processing Handbook, CRC Press.
28. Soille P. 2003. *In* Morphological Image Analysis, 2nd ed. Springer.
29. Shi, J.B., and J. Malik. 2000 Normalized cuts and image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8): 888-905.
30. Peng,H.C., X. He, and F. Long. 2004. Automatic content extraction of filled form images based on clustering component block projection vectors, *Proc. IS&T/SPIE 16th Annual Symp. of Electronic Imaging,Conf. on Document Recognition and Retrieval XI*, San Jose, CA, USA, pp.204-212.
31. Ding, C., X. He, H. Zha, M.Gu, and H. Simon. 2001. A min-max cut algorithm for graph partitioning and data clustering, *Proc. of 1st IEEE Int'l Conf. Data Mining*. San Jose, CA, pp. 107-114.
32. Briand, P., O.W. Petersen, and B. Van Deurs. 1987. A new diploid nontumorigenic human breast epithelial cell line isolated and propagated in chemically defined medium. *In Vitro Cell Dev Biol*. 23:181-8.
33. Briand, P., and A.E. Lykkesfeldt. 2001. An in vitro model of human breast carcinogenesis: epigenetic aspects. *Breast Cancer Res Treat*. 65:179-87.
34. Gudjonsson, T., L. Ronnov-Jessen, R. Villadsen, M.J. Bissell, and O.W. Petersen. 2003. To create the correct microenvironment: three-dimensional heterotypic collagen assays for human breast epithelial morphogenesis and neoplasia. *Methods*. 30:247-55.
35. Weaver, V.M., O.W. Petersen, F. Wang, C.A. Larabell, P. Briand, C. Damsky, and M.J. Bissell. 1997. Reversion of the malignant phenotype of human breast cells in three-dimensional culture and in vivo by integrin blocking antibodies. *J Cell Biol*. 137:231-45.
36. Muthuswamy, S.K., D. Li, S. Lelievre, M.J. Bissell, and J.S. Brugge. 2001. ErbB2, but not ErbB1, reinitiates proliferation and induces luminal repopulation in epithelial acini. *Nat Cell Biol*. 3:785-92.
37. Lelièvre, S.A. and M.J. Bissell. Three dimensional cell culture: The importance of context in regulation of function. *In* Encyclopedia of Molecular Cell Biology and Molecular Medicine (EMCBMM). (Submitted).
38. Lelièvre, S.A. and M.J. Bissell. Utilization of nuclear structural proteins for targeted therapy and detection of proliferative and differentiation disorders. US patent 6,287,790 B1, Sep.11, 2001.
39. Peng, H.C. and F. Long 2005 Feature selection based on mutual information: criteria of maximal dependency, maximal relevance, and minimal redundancy. *IEEE Transaction on Pattern Analysis and Machine Intelligence* (in press)
40. Khan SA and Badve S. 2001. Phyllodes tumors of the breast. *Curr. Treat. Options Oncol.* 2: 139-147.

41. Petersen, O.W., L. Ronnov-Jessen, A.R. Howlett, and M.J. Bissell. 1992. Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells [published erratum appears in Proc Natl Acad Sci U S A 1993 Mar 15;90(6):2556]. *Proc Natl Acad Sci U S A*. 89:9064-8.
42. Fuhui Long, Hanchuan Peng, Damir Sudar, Sophie Lelièvre and David Knowles 2006. Predicting the Phenotype of Mammalian Epithelial Cells using Cluster Analysis of Nuclear Protein Distribution. *In preparation***.**
43. C. L. Luengo Hendriks and D. W. Knowles 2006 Comments on the paper "A novel 3D wavelet-based filter for visualizing features in noisy biological data", by Moss et al. *Journal of Microscopy in press*
44. Tvarusko, W., Bentele, M., Misteli, T., Rudolf, R., Kaether, C., Spector, D. L., Gerdes, H. H., & Eils, R. (1999) *Proc. Natl. Acad. Sci.* USA **96**, 7950-7955.
45. Eils, R., Gerlich, D., Tvarusko, W., Spector, D. L., & Misteli, T. (2000) *Mol. Biol. Cell.* 11, 413-418.
46. Platani, M., Goldberg, I., Swedlow, J. R., & Lamond, A. I. (2000) *J. Cell Biol.* 151, 1561-1574.
47. Kyan, M. J, Guan, L., Arnison, M. R., & Cogswell, C. J. (2001) *IEEE Trans. Biomed. Eng.* 48, 1306-1318.
48. Chin, K., Ortiz de Solorzano, C., Knowles, D., Jones, A., Chou, W., Rodriguez, E. G., Kuo, W. L., Ljung, B. M., Chew, K., Myambo, K., *et al.* (2004) *Nat. Genet.* 36, 984-988.

APPENDICES:

Knowles, D.W., Sudar, D., Bator-Kelly, C,. Mina J. Bissell, M.J., Lelièvre, S.A. **2006**
Automated local bright feature image analysis of nuclear protein distribution identifies changes in tissue phenotype.
Proc. Natl. Acad. Sci. USA **103**, 4445-4450

Automated local bright feature image analysis of nuclear protein distribution identifies changes in tissue phenotype

David W. Knowles*†, Damir Sudar*, Carol Bator-Kelly‡, Mina J. Bissell*, and Sophie A. Lelièvre‡§

*Biophysics and Cancer Biology Departments, Life Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720; and ‡Department of Basic Medical Sciences and Cancer Center, Purdue University, 625 Harrison Street, West Lafayette, IN 47907-2026

Communicated by Carlos J. Bustamante, University of California, Berkeley, CA, November 17, 2005 (received for review March 3, 2005)

The organization of nuclear proteins is linked to cell and tissue phenotypes. When cells arrest proliferation, undergo apoptosis, or differentiate, distribution of nuclear proteins changes. Conversely, forced alteration of the distribution of nuclear proteins modifies cell phenotype. Immunostaining and fluorescence microscopy have been critical for such findings. However, there is increasing need for quantitative analysis of nuclear protein distribution to decipher epigenetic relationships between nuclear structure and cell phenotype and to unravel the mechanisms linking nuclear structure and function. We have developed imaging methods to quantify the distribution of fluorescently stained nuclear protein NuMA in different mammary phenotypes obtained using 3D cell culture. Automated image segmentation of DAPI-stained nuclei was generated to isolate thousands of nuclei from 3D confocal images. Prominent features of fluorescently stained NuMA were detected by using a previously undescribed local bright feature analysis technique, and their normalized spatial density was calculated as a function of the distance from the nuclear perimeter to its center. The results revealed marked changes in the distribution of the density of NuMA bright features when nonneoplastic cells underwent phenotypically normal acinar morphogenesis. Conversely, we did not detect any reorganization of NuMA during formation of tumor nodules by malignant cells. Importantly, the analysis also discriminated proliferating nonneoplastic from proliferating malignant cells, suggesting that these imaging methods are capable of identifying alterations linked not only to the proliferation status but also to the malignant character of cells. We believe that this quantitative analysis will have additional applications for classifying normal and pathological tissues.

3D automated nuclear segmentation | breast cancer | nuclear organization | NuMA | quantitative imaging

The organization of proteins within the cell nucleus appears to play a central role in directing nuclear functions necessary for cell proliferation and differentiation (1, 2). Several nuclear proteins have been reported to display a specific compartmentalization (e.g., within the nucleolus, nuclear domains, or chromatin), and distribution (e.g., diffuse or aggregated), which change during the cell cycle (3, 4), upon a switch between proliferation and growth arrest (5), or after cell differentiation (6–9). Studies using three-dimensional (3D) culture of breast epithelial cells in the presence of laminin-rich extracellular matrix have revealed that the distribution of certain nuclear proteins depends also on tissue morphogenesis (5, 10). These culture models mimic the formation of specific tissue structures, where cells display both the function and the spatial arrangement typically found in a given organ (11, 12). As a consequence, 3D cell culture models are being recognized as the systems of choice for unraveling critical cellular events involved in the development of pathologies such as cancer (13).

The HMT-3522 cancer progression series of human mammary epithelial cells (HMECs), cultured in 3D, constitutes a physiologically relevant model for studying the relationship between cellular organization and gene expression in normal and malignant cells (5, 14–17). In such cultures, nonneoplastic HMT-3522 S1 HMECs (18) reproduce the formation of phenotypically normal, tissue-like glandular structures referred to as acini (17). Acinar morphogenesis proceeds by stepwise events including a proliferation stage from days 1 to 6 of culture, followed by growth arrest and the formation of the baso-apical polarity axis. Upon completion of acinar morphogenesis at day 10, S1 cells are organized into spherical and hollow structures delineated by a basement membrane at their basal pole and a lumen at their apical pole. Conversely, malignant T4-2 cells, which were derived from S1 cells (19), continue to proliferate and form disorganized and invasive tumor-like nodules under the same culture conditions (17, 20). The distribution of nuclear proteins including retinoblastoma protein Rb, splicing factor SRm160, and nuclear mitotic apparatus (NuMA) protein is remarkably different between S1 cells in the early stage of acinar morphogenesis and S1 cells in fully formed acini (5). NuMA is diffusely distributed within the nuclei of proliferating cells but aggregates into foci of increasing size as cells arrest proliferation and complete acinar morphogenesis. Importantly, the distribution of NuMA in acinar S1 cells is similar to that observed in biopsies of normal breast tissue, indicating that the 3D model of acini formation reproduces physiologically relevant features of NuMA organization.

Earlier reports described the distribution of NuMA as a single aggregate in the center of the cell nucleus in different cell types undergoing apoptosis (21, 22) and as diffuse in cells cultured under nondifferentiating condition (i.e., on plastic surfaces producing a flat monolayer of cells) (5, 23). NuMA has also been reported to be organized in distinct foci in differentiated muscle (24) and differentiated lens cells (8). Another report has shown an association between a more punctate distribution of NuMA and a higher susceptibility to apoptosis induction in lymphocytes (25). Thus, NuMA distribution appears highly dependent on cell and tissue phenotypes, and, as such, it has been proposed to constitute a reliable indicator of cell behavior (5, 21, 25).

For decades, it has been known that cancer cells display alterations in nuclear size and chromatin organization (26, 27). Yet despite extensive use of the cell nucleus as a central diagnostic tool in cancer, there is little information available regarding specific alterations in nuclear organization in neoplasia. Understanding the relationship between nuclear organization and cell behavior has gained recent attention because it may help decipher signaling and structural events involved in differ-

Conflict of interest statement: No conflicts declared.

Abbreviations: HMEC, human mammary epithelial cell; LBF, local bright feature; NuMA, nuclear mitotic apparatus protein.

†To whom correspondence may be addressed at: Biophysics Department, Life Sciences Division, Lawrence Berkeley National Laboratory, MS: 84R0171, 1 Cyclotron Road, Berkeley, CA 94720. E-mail: dwknowles@lbl.gov.

§To whom correspondence may be addressed at: Basic Medical Sciences, Purdue University, LYNN, 625 Harrison Street, West Lafayette, IN 47907-2026. E-mail: lelievre@purdue.edu.

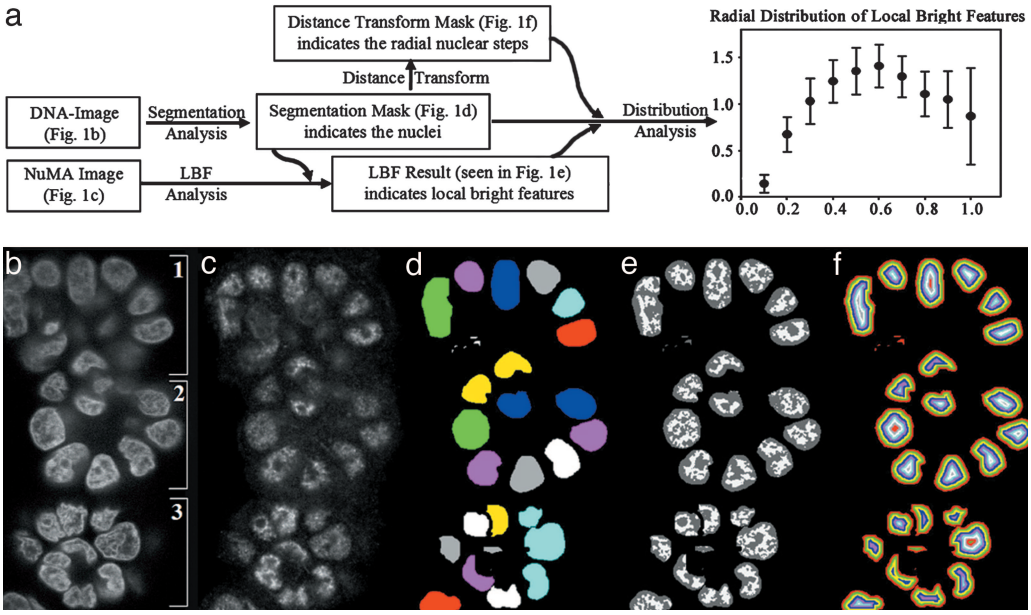


Fig. 1. LBF analysis of NuMA distribution from 3D images. (*a*) Flowchart of the imaging processing steps including a graph of the relative density of LBFs of NuMA in 77 nuclei from the three acini depicted in *b*. The radial distributions of LBFs within each nucleus in the NuMA image are calculated by first generating a segmentation mask from the image of DAPI-stained DNA. The segmentation mask not only defines the extent of each nucleus but is also used to define a set of radial steps within each nuclear volume. The graph shows the mean and standard deviation of the relative density of NuMA bright features extracted by LBF analysis (ordinate) as a function of the relative distance from the perimeter (0.0) to the center (1.0) of the nuclei (abscissa). (*b–f*) Distribution density of the bright features of NuMA in acinar cells. S1 HMECs were cultured in 3D to induce acinar morphogenesis. Each image corresponds to the application of the different steps of distribution analysis starting from the same original image. (*b*) Fluorescence micrograph of DAPI-stained nuclei from a single optical section containing three acini [1;2;3]. (*c*) Fluorescence micrograph of Texas red-immunolabeled NuMA from the optical section corresponding to the DAPI image shown in *a*. (*d*) Segmentation mask derived from the DAPI-stained image showing a single slice of individually enumerated nuclei. (*e*) Composite view of the LBFs (light gray) extracted by the LBF analysis overlaid with the segmentation mask (dark gray). (*f*) Concentric terraces resulting from the application of the distance transform on the segmentation mask.

entiation and cancer (28). With the recent refinements in confocal imaging technology, 3D high-resolution imaging has become a powerful method for recording subtle organizational features in the cell nucleus. However, although a number of recent techniques for the quantitative analysis of 3D images have been reported (7, 29–34), the availability and application of robust image analysis tools in biology remain in their infancy (26, 35, 36).

Here, we report the use of confocal imaging to record the changes in the pattern of NuMA staining in HMECs expressing different phenotypes and the development of an image analysis technique to translate the visual observations of the complexity of NuMA staining into quantitative results. In the original report (5), NuMA organization was determined by manually measuring foci sizes. However, such measurements are not possible when NuMA is diffusely organized and do not take into account the spatial reorganization of NuMA that is apparent in the differentiated cells. To circumvent this problem, we have developed the radial local bright feature (LBF) analysis. In this method, regions of local brightness in images of fluorescently immunolabeled NuMA are isolated by an adaptive LBF analysis technique. The density of local bright features is then calculated within a set of concentric, volumetric terraces that subdivide the nucleus radially from its periphery to its center. The distribution of the bright features of NuMA can be represented by a simple graph, which permits an easy quantification of the changes in the spatial organization of this protein associated with different mammary phenotypes. The method relies on the delineation of individual nuclei, and to analyze thousands of nuclei in a short period, we also have developed a previously undescribed automated, 3D segmentation technique. By using the radial-LBF analysis, we measured a striking reorganization of NuMA during acinar morphogenesis; no such reorganization occurred during the formation of tumor-like nodules. Importantly, the radial-LBF analysis of NuMA distribution permitted a clear discrimination also between proliferating nonneoplastic cells and proliferating malignant cells, which to our knowledge has not been achieved so far using other evaluation methods.

Results

Radial-LBF Distribution Analysis in Phenotypically Normal Breast Acinar Cells Reveals a Peak of Density of NuMA Bright Features Centered on a Shell Located Midway Between the Periphery and the Center of the Nucleus. The organization of NuMA in the nucleus displays intricate spatial distributions that vary with cell and tissue phenotypes (5, 8, 25). Visual analysis of NuMA immunostaining showed the formation of bright NuMA foci in a sea of diffuse NuMA staining during acinar morphogenesis (5). To quantify the distribution of bright NuMA foci, we calculated the radial distribution of the density of NuMA bright features isolated with the LBF analysis (described in *Materials and Methods*) within the 3D volume of each nucleus (Fig. 1*a*). S1-HMECs were cultured in 3D to induce phenotypically normal acinar morphogenesis where NuMA domains are most abundant. Acini were immunostained for NuMA and counterstained with DAPI (Fig. 1 *b* and *c*). A segmentation mask, which describes the position and extent of individual nuclei in 3D, was created from the DAPI image (Fig. 1*d*). Bright NuMA features were isolated from diffuse staining in the NuMA image by using the LBF analysis. To visualize the localized accumulation of NuMA foci in the nucleus, the resulting bright features were overlaid on the segmentation mask (Fig. 1*e*). This visualization indicated that the density of NuMA bright features was low at the periphery of the nucleus and varied with the depth, radially, into

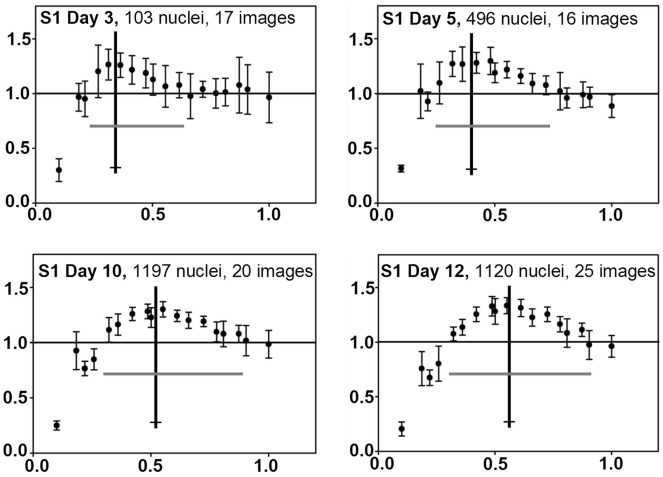


Fig. 2. Average relative density of LBFs of NuMA during acinar morphogenesis. S1 cells were cultured in 3D for 3, 5, 10, and 12 days. Plots represent the relative density of NuMA bright features extracted by LBF analysis (ordinate) of a population of nuclei as a function of the relative distance from the perimeter (0.0) to the center (1.0) of the nuclei (abscissa) for each time point. Vertical lines (black) represent the location of the peak of bright feature density in the nucleus. Horizontal lines (gray) represent the extent of nuclear volume with densities of bright features above the average. The number of days the cells were in culture and the number of nuclei analyzed are indicated above each corresponding graph. Bars represent the standard deviations of the relative density of NuMA bright features calculated from multicellular units of the same phenotype, on a per-image basis.

the nucleus. To quantify NuMA bright features, a distance transform was applied to the segmentation mask to subdivide each nuclear volume into a set of concentric terraces of equal thickness, starting at the nuclear perimeter (Fig. 1*f*). The distance transformed-segmentation mask was combined with the LBF image to calculate the variation of the relative density of NuMA bright features as a function of the relative radial distance measured from the perimeter of the nucleus to its center. To demonstrate the consistency of this radial-LBF analysis, the radial distribution was plotted for 77 nuclei analyzed within a single image (Fig. 1*a*, graph). As the visual representation indicated (Fig. 1*e*), the density of NuMA bright features was below average at the perimeter of the nuclei. As the radial depth into the nucleus increased, the density of bright features increased and reached a peak, above the average density, at radial distance 0.55 ± 0.05. Then the density decreased to a value close to the average density as the center of the nucleus was reached.

Distribution of NuMA Bright Features Changes as a Function of Acinar Morphogenesis. Previous analysis of the distribution of NuMA in S1 HMECs during the cell proliferation stage (day 3 of 3D culture) and upon acinar differentiation (day 10 of 3D culture) suggested that there was an increase in the foci-like aggregation of NuMA upon completion of acinar morphogenesis (5). This analysis was based on visual estimation and the manual measurement of the size of NuMA foci on images of NuMA staining. To assess the efficiency of the radial-LBF analysis in measuring the changes in NuMA distribution reproducibly along the morphogenesis process, we analyzed images of S1 HMECs cultured in 3D as a function of time over a period of 12 days. At day 3, the density of NuMA bright features was the highest in a region toward the perimeter of the nucleus (Fig. 2). After 10 days of culture, there was a clear reorganization of NuMA away from the perimeter and toward the center of the nucleus (Fig. 2). This reorganization was accompanied by a significant decrease of the density of NuMA bright features at the perimeter of the nucleus.

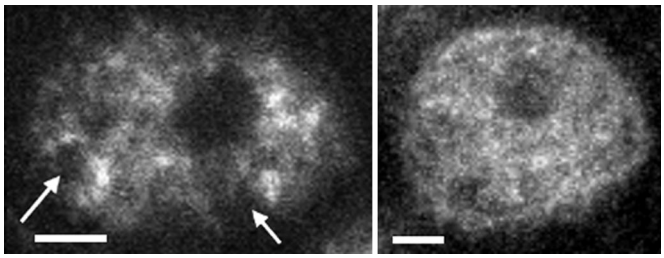


Fig. 3. Immunostaining of NuMA in acinar S1 and malignant T4-2 HMECs. S1 and T4-2 cells were cultured in 3D for 10 days. Each image shows an optical section of NuMA immunostaining through the center of the nucleus in an acinar S1 cell (*Left*) and a malignant T4-2 cell (*Right*). One nucleus is shown per image. Arrows indicate typical empty areas in NuMA staining at the periphery of the nucleus of the S1 cell. (Scale bar: 2 μm.)

Visual inspection of images of NuMA distribution in acinar cells showed that this protein was usually absent from regions located at the periphery of the nucleus (Fig. 3). Similar density distributions were obtained with two different antibodies directed against NuMA (data not shown). To establish the statistical significance of the differences measured for NuMA distribution between proliferating S1 cells, cultured for 5 days or fewer, and differentiated S1 cells, cultured for 10 days or more, we compared the average radial position of the distribution maxima, shown by the vertical black line in Fig. 2. The *P* value between days 5 and 10 was 0.023, indicating significant difference between the distributions at those days. Thus, these results quantitatively confirmed our initial visual observation that acinar morphogenesis is accompanied by the reorganization of NuMA foci (5).

Distribution of NuMA Bright Features in Malignant T4-2 Cells Differs from both Proliferating and Differentiated Nonneoplastic S1 Cells. In proliferating S1 cells (day 3 of 3D culture), NuMA distribution is more diffuse than at day 5 of 3D culture and in fully differentiated (day 10 of 3D culture) acinar cells. We asked whether the diffuse distribution of NuMA was a characteristic of a cell population that was actively proliferating, regardless of whether or not it was malignant. As expected, immunostaining of malignant T4-2 cells for NuMA after 10 days of 3D culture showed that this protein was mostly diffusely distributed (Fig. 3) and that overall this distribution did not appear to be visually different from that observed in proliferating nonneoplastic S1 cells. To quantitatively assess this visual observation, we applied the radial-LBF analysis to 3D cultures of T4-2 cells as a function of time up to 11 days. During this culture period, T4-2 cells formed disorganized tumor-like nodules of increasing sizes. In contrast to nonneoplastic S1 cells, the radial-LBF analysis showed a fairly flat distribution of NuMA bright features in malignant cells, regardless of the number of days in 3D culture (Fig. 4*a*). Thus, despite the increase in mass, there were no significant alterations in the phenotype of tumor nodules during 11 days of 3D culture, and there was no apparent change in the density of NuMA bright features in the nuclei of malignant cells during the entire culture period.

The distribution curves of the density of bright features of NuMA in T4-2 cell nuclei did not show a clear peak at any of the time points, suggesting that there was a difference in NuMA distribution not only between malignant T4-2 cells and acinar S1 cells, but also between malignant T4-2 cells and proliferating S1 cells. To better visualize the differences in the distribution of the bright features of NuMA for the different phenotypes at the time points described above, we plotted the cumulative density of NuMA bright features that exceeded unity on Figs. 2 and 4*a* as a function of the distance from the nuclear boundary. The cumulative plots unambiguously show that the distribution of the

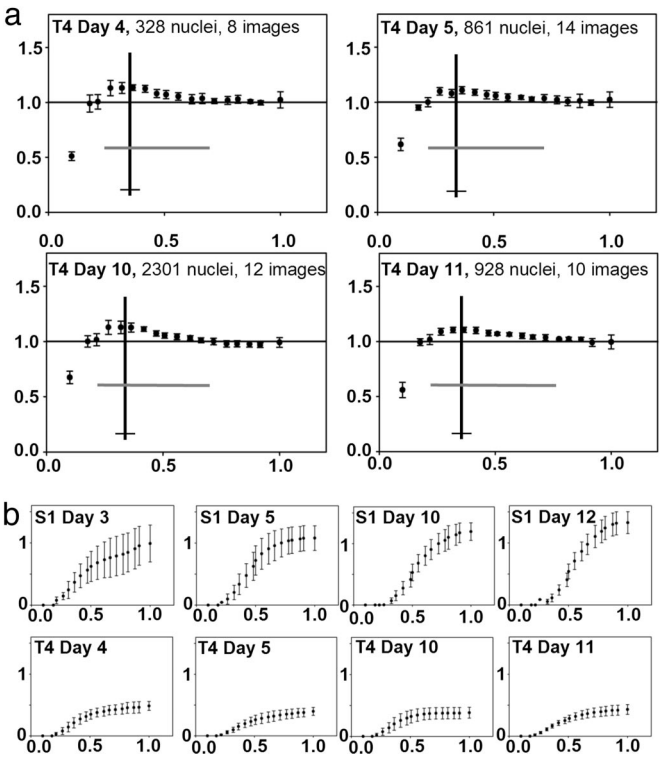


Fig. 4. Differences in the relative density of NuMA bright features between nonneoplastic and malignant cells. (*a*) T4-2 cells were cultured in 3D for 4, 5, 10, and 11 days. Plots represent the relative density of NuMA bright features extracted by LBF analysis (ordinate) of a population of nuclei as a function of the relative distance from the perimeter (0.0) to the center (1.0) of the nuclei (abscissa) for each time point. Vertical lines (black) represent the location of the peak of bright feature density in the nucleus. Horizontal lines (gray) represent the extent of nuclear volume with densities of bright features above the average. The number of days the cells were in culture and the number of nuclei analyzed are indicated above each graph. (*b*) Cumulative plots of the relative density of NuMA bright features above unity (ordinate) at different time points of 3D culture of S1 and T4-2 cells as a function of the relative distance from the perimeter (0.0) to the center (1.0) of the nucleus (abscissa). Cumulative plots for S1 and T4-2 cells were prepared from the relative density data shown in Figs. 2 and 4*a*, respectively. Bars represent the standard deviations of the relative density of NuMA bright features calculated from multicellular units of the same phenotype, on a per image basis.

bright foci of NuMA is consistently similar for the different culture time points of the malignant T4-2 cells and that such a distribution is remarkably different from any of the stages of acinar morphogenesis of S1 cells, including the proliferation phase (Fig. 4*b*). To establish the statistical significance of the differences measured for NuMA distribution between S1 and T4-2 cells, we calculated the *P* values of the average maxima accumulation. Fig. 4*b* shows that the average maximum accumulation in S1 cells lies >7 SD away from the average maximum accumulation for T4-2 cells. This finding results in a *P* value of <0.001, indicating that the distributions for S1 and T4-2 cells are significantly different.

Discussion

We have developed an automated image analysis method that quantifies the radial distribution of nuclear proteins, on a per-nucleus basis. Although the focus here has been the analysis of NuMA distribution, the tools developed are expected to be fully applicable to other nuclear proteins. The analysis identified individual nuclei within an image, revealed bright features of NuMA staining within each nucleus, and calculated the relative density of the bright features of NuMA staining as a function of the distance from the perimeter of the nucleus to its center. The results demonstrate quantitatively that the organization of NuMA is dynamic and is linked directly to the phenotype of the HMECs. During the process of acinar morphogenesis, there is a marked decrease in the relative density of NuMA bright features at the perimeter of the nucleus and a marked increase in this same parameter toward the center of the nucleus. In contrast, the relative density of NuMA is more uniformly distributed in malignant cells and there is no measurable variation in its distribution during the growth of tumor-like nodules. Furthermore, the distribution of NuMA in malignant cells is clearly different from that in nonneoplastic cells regardless of the stage of acinar morphogenesis.

One of the key steps of the image analysis is the delineation of individual nuclei from 3D fluorescence images. To permit the analysis of large numbers of nuclei, we have developed a segmentation method that is automated. Our method builds on approaches described in refs. 29 and 37–39. Irinopoulou and colleagues (30) used a global threshold, a distance transform, and a watershed method to segment nuclei on a per-image-slice basis. Their final 3D reconstruction then was produced by implementing a rule-set to correctly join nuclei in adjacent slices. Our technique uses an adaptive threshold (29) that enables us to correct for inherent image anisotropy and work directly in three dimensions. Then, much like Irinopoulou and colleagues (30), we apply a distance transform, but instead of a watershed method, we use template-matching and region-growing techniques, which are directed by the results of the distance transform. These techniques allow us to use the known geometry of the nuclei and produce a more accurate segmentation than an unconstrained watershed technique (40–43). Our segmentation method is optimized with the help of tools that present a user with the raw DAPI-stained image overlaid with the corresponding segmentation mask and permit the visual scoring of the segmentation accuracy. Application of these tools shows that although some segmentation errors occur, the number of errors is insignificant compared with the number of correctly segmented nuclei. Our ongoing efforts are focused on improving the nuclear segmentation technique to maintain accuracy and efficiency in cases where the morphology of the nuclei is even more complex.

A major concern of using 3D confocal images for quantitative analysis is the inherent image anisotropy, which is linked to the nature of image acquisition. Confocal images are more highly resolved in directions perpendicular than parallel to the optical axis. This characteristic is due to the spatial asymmetry of both the point-spread function of the excitation illumination and the microscope's "pinhole" spatial filter. Also, the collection efficiency and hence the brightness of confocal images decreases with depth into the object. The severity of this penetrative loss depends on the physical properties of the object, the mounting medium, and the objective lens. Standard image analysis techniques are often based on ideal imaging assumptions, which neglect inherent properties of confocal images. In such cases, images must be preprocessed using restoration techniques to remove artifacts created by confocal imaging before the quantitative analysis may proceed. In contrast, our analysis techniques take inherent properties of confocal images into account and allow results from different images to be quantitatively compared, independently of variations in fluorescence staining efficiency and acquisition parameters. The LBF analysis isolates local bright and local dark features within an image by using an adaptive approach where a kernel of neighboring imaging pixels is defined around each point in the image. The LBF analysis then uses the relative brightness of the neighboring pixels in the kernel to classify each pixel. These types of nonlinear techniques are powerful because they mimic human visual perception, espe-

cially the ability to isolate rare events, such as a small number of foci in a diffuse background. Furthermore, the size of the kernel sets a spatial sensitivity limit to the LBF technique, and its relative dimensions can be easily adjusted to match the spatial sampling asymmetry of the microscope. Consequently, the LBF technique is not affected by the absolute brightness of an image or long-scale brightness variations like penetrative loss, and restoration techniques such as background subtraction, attenuation correction, and image interpolation are not necessary.

The ability to quantify the spatial distribution of fluorescent bright cellular features has many biological applications ranging from the study of gene expression and protein movement in live cells and the exploration of the structural aspects of cell division to the investigation of the role of nuclear alterations in pathologies (30, 31, 34, 44–48). We believe that the LBF analysis, which isolates LBFs, and the radial-LBF analysis, which quantifies the distribution of the bright features, are examples of powerful tools capable of measuring differences in the complex distribution of endogenously expressed nuclear proteins from 3D images acquired following simple immunostaining procedures. Radial-LBF analysis has led to findings that strongly support the concept that specific cell and tissue phenotypes are reflected by the organization of nuclear components. These findings underline the importance of reorganization within the nucleus during the differentiation process and the alterations in nuclear organization that may be associated with tumor behavior. It was not the purpose of this study to measure independently the effect of specific cellular events that may account for tumor phenotypes, like the cell cycle phase or changes in the number of chromosomes, on the distribution of NuMA. However, the investigation of such effects will be of great value to refine the phenotypic classification, especially when working with small numbers of cells. A possible future goal is to create a quantitative 3D "view" of cells and tissues, based on the redistribution of nuclear proteins that helps understand the organization of the nucleus and aids in the classification of pathological samples.

Materials and Methods

Cell Culture. HMT-3522 nonneoplastic (S1) cells (18) and HMT-3522 malignant T4-2 cells (19) were cultured in serum-free H14 medium as described in refs. 17 and 20. To induce acinar morphogenesis, S1 cells were cultured in 3D for up to 12 days on 40 μl/cm^2 Matrigel (BD Biosciences)-coated surfaces in the presence of culture medium containing 5% Matrigel (10). Tumor-like nodule formation was achieved by culturing T4-2 cells under similar conditions for a maximum of 11 days to avoid overgrowth.

Immunostaining and Image Acquisition. 3D cultures of S1 and T4-2 cells in four-well chamber slides were permeabilized with 0.5% peroxide and carbonyl-free Triton X-100 (Sigma Biosciences) in cytoskeleton buffer (100 mM NaCl/300 mM sucrose/10 mM Pipes, pH 6.8/5 mM $MgCl_2$) containing protease and phosphatase inhibitors (1 mM Pefabloc/10 μg/ml aprotinin/250 μM NaF), before fixation in 4% paraformaldehyde and immunostaining (5). Primary monoclonal antibodies against NuMA were from clone 204.4 (Oncogene Research Products, San Diego) and B1C11 (a gift from Jeffrey Nickerson, University of Massachusetts, Amherst). Secondary antibody was Texas red-conjugated (Jackson ImmunoResearch). Nuclei were counterstained with DAPI. After immunostaining, 3D cultures were mounted in antifade medium (ProLong; Molecular Probes) under #1 coverglass. Optically sectioned images of DAPI-stained DNA and Texas red-labeled NuMA were acquired sequentially throughout the volume of the acini and assembled into 3D images. DAPI and Texas red signals were acquired simultaneously into separate channels by using a Zeiss 410 confocal laser-scanning microscope with a planapochromatic 63×, 1.4 numerical aperture objective. The resulting voxel dimensions of the 3D images were 0.08×0.08 μm in the plane of the slide and 0.5 μm along the optical direction.

Segmentation of Individual Nuclei. To isolate individual nuclei in the 3D DAPI-stained image, a model-based automatic nuclear segmentation method was developed on the assumption that nuclei of epithelial cells are of simple geometry in that they encompass a single spherical core. An adaptive threshold was first applied to the DAPI-stained image to produce the binary segmentation mask of the nuclei. The technique, which normalizes for penetrative loss along the optical direction, uses a difference-of-Gaussians filter (49) followed by a morphological closing filter and a flood-fill algorithm (40–43). Although this technique accurately delineates nuclei from their background, it does not completely separate neighboring nuclei when they are tightly clustered. To separate nuclei that are connected by the binary mask, the central core of each nucleus was located, by using standard template-matching techniques, and dilated into the rest of the nucleus by using standard region-growing techniques (40–43). Briefly, a template was constructed with dimensions that approximated those of the average spherical core of nuclei and was convolved with the binary nuclear mask. This convolution produced a map that indicated the percentage of the template that fits within the binary mask at each point in the image. Then the template was stamped into the binary mask at locations where there were corresponding local maxima in the map that exceeded 70%. The templates were stamped at the center-of-mass of the local maximum, in an order ranked by their percentage, starting from the highest. A template was not stamped if the local maximum was $<70\%$, if it overlapped a previously stamped template by $>70\%$, or if the local maximum was at the boundary of the binary mask. Once all of the nuclear cores were located, each template was dilated in a semiintelligent fashion. The template dilation was done independently in the positive and negative X, Y, and Z directions. Dilation along any direction was halted when 60% of the dilating template boundary reached the boundary of the binary mask. Halting the dilation prevented the template from squeezing through narrow regions in the binary mask that connected two adjacent nuclei. Also, dilation was stopped in all directions if the volume of the dilated template exceeded nine times its original volume. This phenomenon occurred if nuclei were clustered so closely that the initial segmentation mask failed to separate them adequately. The resulting object was reported as an undersegmentation error.

Distribution Analysis of NuMA. After immunostaining, image acquisition, and nuclear segmentation, NuMA bright features were first isolated by the LBF analysis technique, and then their radial distribution was calculated by our radial-LBF technique. In the LBF analysis, pixel brightness in the raw NuMA images was normalized by the local average brightness using an extension of the difference-of-Gaussians technique (49). The raw NuMA image was masked by the binarized segmentation result derived from the DAPI image as described above. Then the image brightness within each nucleus was rescaled by dividing the brightness at each point by the average brightness within a local region surrounding that point. The dimension of the local region was chosen to be half that of the dimension of the nuclear core. This choice resulted in a local region that was significantly larger than the bright NuMA foci of interest but smaller than the nuclear dimension. This feature was important because the LBF technique sensitively resolves light or dark features that lie within the local region while ignoring features that are larger. Using this approach allowed the bright foci and dark regions of interest within the nucleus to be resolved and the low-frequency brightness variations, due to nuclear geometry and finite axial

resolution, to be correctly normalized. In the resulting LBF images, bright image features have values above unity, whereas dark image features have values below unity. For the radial-LBF analysis, a distance transform (40–43) was applied to the nuclear segmentation mask. The transform calculates the shortest distance of each point within a nucleus to the nuclear boundary and, in doing so, divides each nucleus into a set of concentric terraces of equal thickness. The LBF image then was used in conjunction with the nuclear segmentation mask and the distance transform to compute the density of LBFs in each terrace of each nucleus (see Fig. 1*a*). In each terrace, the density was calculated as the number of pixels in LBFs divided by the total number of pixels. The relative distribution of the density of bright features within each nucleus was revealed by normalizing the density per terrace, so that the average density of bright features was unity for each nucleus. The distances defined by the distance transform also were normalized so that the distance at the nuclear perimeter was 0 and the distance at the center of the nucleus was 1.0. This normalization was done to account for variation in the number of terraces per nucleus due to variations in nucleus size and shape. Finally, normalized density of bright features was plotted against normalized distance from the perimeter of the nucleus to its center.

Statistical Analysis. P values were derived from the standard normal distribution using the Z score [$Z = (X - \mu)/\sigma$]. Two averages with a P value of <0.05 were considered significantly different.

We thank John Turek and Fuhui Long for critical reading of the manuscript. This work was supported by Department of Defense–Breast Cancer Research Program Grant DAMD-170210440 (to D.W.K.), Department of Energy, Office of Health and Environmental Research Grant DE-AC03-76SF0098 (to M.J.B., D.W.K., and D.S.), Walther Cancer Institute Grant WCI-110-114 (to S.A.L.), "Friends You Can Count On" and Purdue Research Foundations (S.A.L.), and a Department of Defense Breast Cancer Research Program Innovator Award (to M.J.B.).

1. Lelièvre, S. A., Bissell, M. J. & Pujuguet, P. (2000) *Crit. Rev. Eukaryotic Gene Expression* **10,** 13–20.
2. Dillon, N. & Festenstein, R. (2002) *Trends Genet.* **18,** 252–258.
3. Mancini, M. A., He, D., Ouspenski, I. I. & Brinkley, B. R. (1996) *J. Cell Biochem.* **62,** 158–164.
4. Gerlich, D., Beaudouin, J., Gebhard, M., Ellenberg, J. & Eils, R. (2001) *Nat. Cell Biol.* **3,** 852–855.
5. Lelièvre, S. A., Weaver, V. M. Nickerson, J. A. Larabell, C. A. Bhaumik, A. Petersen, O. W. & Bissell, M. J. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14711–14716.
6. Antoniou, M., Carmo-Fonseca, M., Ferreira, J. & Lamond, A. I. (1993) *J. Cell Biol.* **123,** 1055–1068.
7. Beil, M., Durschmied, D., Paschke, S., Schreiner, B., Nolte, U., Bruel, A. & Irinopoulou, T. (2002) *Cytometry* **47,** 217–225.
8. Gribbon, C., Dahm, R., Prescott, A. R. & Quinlan R. A. (2002) *Eur. J. Cell Biol.* **81,** 557–566.
9. Cammas, F., Oulad-Abdelghani, M., Vonesch, J. L., Huss-Garcia, Y., Chambon, P. & Losson, R. (2002) *J. Cell Sci.* **115,** 3439–3448.
10. Plachot, C. & Lelièvre, S. A. (2004) *Exp. Cell Res.* **298,** 122–132.
11. Powers, M. J., Janigian, D. M., Wack, K. E., Baker, C. S., Beer Stolz, D. & Griffith, L. G. (2002) *Tissue Eng.* **8,** 499–513.
12. Gudjonsson, T., Ronnov-Jessen, L., Villadsen, R., Bissell, M. J. & Petersen, O. W. (2003) *Methods* **30,** 247–255.
13. Jacks, T. & Weinberg, R. A. (2002) *Cell* **111,** 923–925.
14. Bissell, M. J., Weaver, V. M., Lelievre, S. A., Wang, F., Petersen, O. W. & Schmeichel, K. L. (1999) *Cancer Res.* **59,** 1757–1763.
15. Bissell, M. J., Rizki, A. & Mian, S. (2003) *Curr. Opin. Cell Biol.* **15,** 753–762.
16. Weaver, V. M., Lelievre, S., Lakins, J. N., Chrenek, M. A., Jones, J. C., Giancotti, F., Werb, Z. & Bissell, M. J. (2002) *Cancer Cell* **2,** 205–216.
17. Petersen, O. W., Ronnov-Jessen, L., Howlett, A. R. & Bissell, M. J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9064–9068.
18. Briand, P., Petersen, O. W. & Van Deurs, B. (1987) *In Vitro Cell Dev. Biol.* **23**, 181–188.
19. Briand, P., Nielsen, K. V., Madsen, M. W. & Petersen, O. W. (1996) *Cancer Res.* **56,** 2039–2044.
20. Weaver, V. M., Petersen, O. W., Wang, F., Larabell, C. A., Briand, P., Damsky, C. & Bissell, M. J. (1997) *J. Cell Biol.* **137**, 231–245.
21. Weaver, V. M., Carson, C. E. Walker, P. R., Chaly, N. Lach, B., Raymond, Y., Brown, D. L. & Sikorska, M. (1996) *J. Cell Sci.* **109,** 45–56.
22. Zweyer, M., Riederer, B. M., Ochs, R. L., Fackelmayer, F. O., Kohwi-Shigematsu, T., Bareggi, R., Narducci, P. & Martelli, A. M. (1997) *Exp. Cell Res.* **230,** 325–336.
23. Mattagajasingh, S. N., Huang, S. C., Hartenstein, J. S., Snyder, M., Marchesi, V. T. & Benz, E. J. (1999) *J. Cell Biol.* **145,** 29–43.
24. Merdes, A. & Cleveland, D. W. (1998) *J. Cell Sci.* **111,** 71–79.
25. Sodja, C., Walker, P. R., Brown, D. L. & Chaly, N. (1997) *Biochem. Cell Biol.* **75,** 399–414.
26. Gil, J. & Wu, H. S. (2003) *Cancer Invest.* **21,** 950–959.
27. Zink, D., Fische, A. H. & Nickerson, J. A. (2004) *Nat. Rev. Cancer* **4,** 677–687.
28. Shumaker, D. K., Kuczmarski, E. R. & Goldman, R. D. (2003) *Curr. Opin. Cell Biol.* **15,** 358–366.
29. Ancin, H., Roysam, B., Dufresne, T. E., Chestnut, M. M., Ridder, G. M., Szarowski, D. H. & Turner, J. N. (1996) *Cytometry* **25,** 221–234.
30. Irinopoulou, T., Vassy, J., Beil, M., Nicolopoulou, P., Encaoua, D. & Rigaut, J. P. (1997) *Cytometry* **27,** 99–105.
31. Lieb, J. D., Ortiz de Solorzano, C., Rodriguez, E. G., Jones, A., Angelo, M., Lockett, S. & Meyer, B. J. (2000) *Genetics* **156,** 1603–1621.
32. Schupp, S., Elmoataz, A., Herlin, P. & Bloyet, D. (2001) *Anal. Quant. Cytol. Histol.* **23,** 257–267.
33. Eils, R. & Athale, C. (2003) *J. Cell Biol.* **161,** 477–481.
34. Voss, T. C., Demarco, I. A., Booker, C. F. & Day, R. N. (2004) *BioTechniques* **36,** 240–247.
35. Swedlow, J. S., Goldberg, I., Brauner, E. & Sorger, P. (2003) *Science* **300,** 100–102.
36. Ecker, R. C. & Steiner, G. E. (2004) *Cytometry* **59A**, 182–190.
37. Rigaut, J. P., Vassy, J., Herlin, P., Duigou, F., Masson, E., Briane, D., Foucrier, J., Carvajal-Gonzalez, S., Downs, A. M. & Mandard, A. M. (1991) *Cytometry* **12,** 511–524.
38. Belien, J. A., van Ginkel, H. A., Tekola, P., Ploeger, L. S., Poulin, N. M., Baak, J. P. & van Diest, P. J. (2002) *Cytometry* **49,** 12–21.
39. Lin, G., Adiga, U., Olson, K., Guzowski, J. F., Barnes, C. A. & Roysam, B. (2003) *Cytometry* **56A,** 23–36.
40. Castleman, K. R. (1996) *Digital Image Processing* (Prentice–Hall, Englewood Cliffs, NJ).
41. Russ, J. C. (1992) *The Image Processing Handbook* (CRC, Boca Raton, FL).
42. Jahne, B. (2002) *Digital Imaging Processing* (Springer, Berlin), 5th Ed.
43. Soille, P. (2003) *Morphological Image Analysis* (Springer, Berlin), 2nd Ed.
44. Tvarusko, W., Bentele, M., Misteli, T., Rudolf, R., Kaether, C., Spector, D. L., Gerdes, H. H. & Eils, R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 7950–7955.
45. Eils, R., Gerlich, D., Tvarusko, W., Spector, D. L. & Misteli, T. (2000) *Mol. Biol. Cell* **11,** 413–418.
46. Platani, M., Goldberg, I., Swedlow, J. R. & Lamond, A. I. (2000) *J. Cell Biol.* **151,** 1561–1574.
47. Kyan, M. J., Guan, L., Arnison, M. R. & Cogswell, C. J. (2001) *IEEE Trans. Biomed. Eng.* **48,** 1306–1318.
48. Chin, K., Ortiz de Solorzano, C., Knowles, D., Jones, A., Chou, W., Rodriguez, E. G., Kuo, W. L., Ljung, B. M., Chew, K., Myambo, K., *et al.* (2004) *Nat. Genet.* **36**, 984–988.
49. Marr, D. (1982) *Vision* (Freeman, New York).