Journal of Cellular Biochemistry Supplement 39:194–210 (2002)

Advances in Molecular Labeling, High Throughput Imaging and Machine Intelligence Portend Powerful Functional Cellular Biochemistry Tools

Jeffrey H. Price,[1,2]* Angela Goodacre,[2] Klaus Hahn,[3] Louis Hodgson,[1,3] Edward A. Hunter,[2] Stanislaw Krajewski,[4] Robert F. Murphy,[5] Andrew Rabinovich,[1,6] John C. Reed,[4] and Susanne Heynen[1]

[1]Department of Bioengineering and Whitaker Institute of Biomedical Engineering, University of California San Diego, La Jolla, California
[2]Q3DM, Inc., San Diego, California
[3]Department of Cell Biology, The Scripps Research Institute, La Jolla, California
[4]The Burnham Institute, La Jolla, California
[5]Departments of Biological Sciences and Biomedical Engineering and Center for Automated Learning and Discovery, Carnegie Mellon University, Pittsburgh, Pennsylvania
[6]Department of Computer Science and Engineering, University of California San Diego, La Jolla, California

Abstract Cellular behavior is complex. Successfully understanding systems at ever-increasing complexity is fundamental to advances in modern science and unraveling the functional details of cellular behavior is no exception. We present a collection of prospectives to provide a glimpse of the techniques that will aid in collecting, managing and utilizing information on complex cellular processes via molecular imaging tools. These include: 1) visualizing intracellular protein activity with fluorescent markers, 2) high throughput (and automated) imaging of multilabeled cells in statistically significant numbers, and 3) machine intelligence to analyze subcellular image localization and pattern. Although not addressed here, the importance of combining cell-image-based information with detailed molecular structure and ligand-receptor binding models cannot be overlooked. Advanced molecular imaging techniques have the potential to impact cellular diagnostics for cancer screening, clinical correlations of tissue molecular patterns for cancer biology, and cellular molecular interactions for accelerating drug discovery. The goal of finally understanding all cellular components and behaviors will be achieved by advances in both instrumentation engineering (software and hardware) and molecular biochemistry. J. Cell. Biochem. Suppl. 39: 194–210, 2002.

Key words: microscopy; protein activity; pattern analysis; cellular diagnostics; tissue microarrays; drug discovery; fluorescent probes

Imaging and molecular labeling are combining to build powerful investigative tools for functional cellular biochemistry. Knowledge gained from the genomics and proteomics revolutions are perhaps in some senses overwhelming the ability of current laboratory methods such as gel electrophoresis, mass spectroscopy, computational prediction methods, and protein chips for protein studies. Although higher throughput versions of these kinds of techniques continue to be developed [Hiroaki, 2002], the need for high throughput measurements of molecules directly in cells is apparent due to their tremendous complexity [Gibbs, 2001]. Prospects for labeling techniques combined with high throughput microscopy (HTM) and higher-level machine intelligence are presented here. HTM may at first only seem to add to the potential information overload. But the prospects for machine intelligence and smart database collation of information facilitate a

Disclosure Statement; J. Price owns equity in Q3DM Inc.

Grant sponsor: National Science Foundation Research Instrumentation; Grant number: BES-9871365. Grant sponsor: The Whitaker Foundation Biomedical Engineering Research Grant Program; National Institutes of Health; Grant numbers: R44 RR15170, R33 CA83219. Grant sponsor: WPC Research and Education Fund.

*Correspondence to: Jeffrey H. Price, MD, PhD, UCSD Department of Bioengineering, 9500 Gilman Drive, MS 0412, La Jolla, CA 92093-0412. E-mail: jhprice@ucsd.edu

Received 7 November 2002; Accepted 7 November 2002
DOI 10.1002/jcb.10448
Published online in Wiley InterScience (www.interscience.wiley.com).

vision of the future where comprehensive cellular protein-activity-structure-function is available for substantially all molecules in a cell. Realization of such a vision depends on the continued advances in computer power and storage for possibly a few decades.

Intracellular molecular function is in and of itself complex because a single enzyme may have multiple activators and/or inhibitors. The prospects for labeling intracellular protein activity are described in Section I. The complexity of obtaining information from multilabeled cells in statistically significant numbers would be simplified by intelligent automation. The prospects for imaging cells with automated HTM are described in Section II. The function of intracellular proteins can best be understood in the context of subcellular location. The prospects for automatically differentiating proteins from one another using machine intelligence to analyze subcellular localization and pattern are described in Section III. Driven by new understanding of proteins and function, molecular labels have the potential to revolutionize cellular diagnostics. The prospects for diagnosing cells using molecular labels are described in Section IV. Diagnostics based on molecular labels are best understood in the context of the patterns of expression in disease. The prospects for learning new information about molecular patterns of cancer using tissue microarrays are described in Section V. Imaging automation and the broad array of developing cellular molecular knowledge have the potential to accelerate drug discovery. The prospects for discovering drugs by imaging subcellular molecular activity are described in Section VI. These prospectives represent a snapshot of only a portion of the tremendous potential for molecular imaging.

PROSPECTING INTRACELLULAR PROTEIN ACTIVITY WITH FLUORESCENT LABELS

Using proteins tagged with fluorescent dyes, the ability to study proteins in their native environment has been a critical tool in cell biology for over two decades [Taylor and Wang, 1980]. However, this approach has been limited to proteins that could be labeled and reintroduced into correct cellular compartments via methods such as microinjection. A revolution in using fluorescence techniques to study biological systems came with the discovery of green fluorescent protein (GFP) and its analogues derived from the jellyfish *Aequoria victoria* [Chalfie et al., 1994; Heim and Tsien, 1996]. The advantage of GFP is that it can be cloned into expression constructs and used by simple transfection to attain stable expression in many cell types. Initial use of GFP was limited to tagging proteins to observe the changes in localizations in vivo [Misteli and Spector, 1997]. In these early fluorescent protein analogues, constructs were purposely designed not to disturb the fluorescence characteristics of GFP, so that the protein localizations could be quantified with ease. With this approach, bulk translocations of proteins could be observed; however, functional changes such as post-translational modification or conformational changes could not be detected.

Rapid developments in addressing the limitations of GFP-tagged fluorescent protein analogues came in the form of fluorescence resonance energy transfer (FRET) biosensors [Adams et al., 1991; Heim and Tsien, 1996; Miyawaki et al., 1997]. These sensors utilized GFP and its mutants in different "colors" and sometimes other fluorescent dyes with overlapping fluorescence excitation and emission spectra. When the two fluorophores come within approximately 80 Å of one another, FRET takes place, such that excitation of the donor fluorophore produces emission from the acceptor fluorophore. The primary advantage of this technique is the ability to monitor interactions between two components within a cell; i.e., by tagging a protein with donor CFP and monitoring acceptor emission from the YFP attached to another protein, one observes FRET emission only when the two proteins come together. Several variants of this technique have been developed to observe more specific aspects of protein activity (Fig. 1). Successful use of strategies including protein transducers, and intra- and intermolecular FRET probes have been reported [Adams et al., 1991; Hahn et al., 1992; Miyawaki et al., 1997; Llopis et al., 1999; Ting et al., 2001; Haj et al., 2002]. Of these variants of FRET biosensors, domain or antibody-based biosensors (Fig. 1d) offer the ability to monitor exposures of active sites of endogenous proteins during cellular events such as cell adhesion, spreading, and migration [Chamberlain and Hahn, 2000; Kraynov et al., 2001]. Domains used in this approach are a small peptide sequence derived from the consensus binding

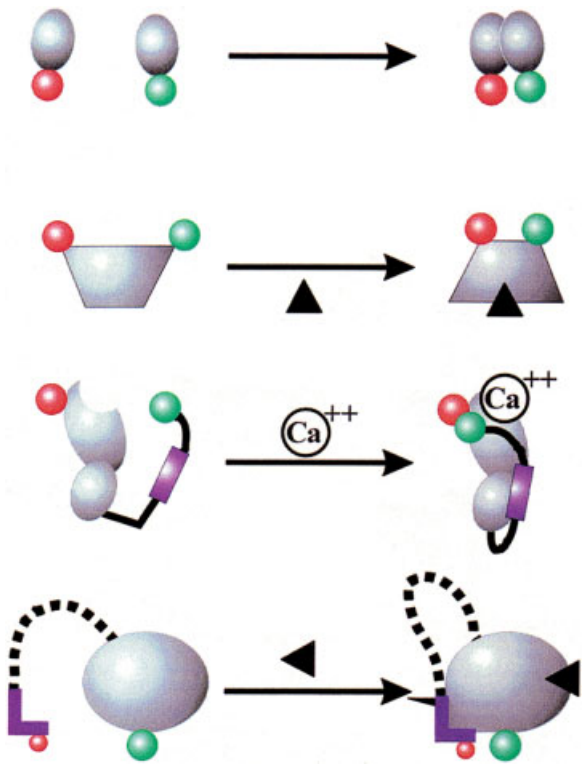


Fig. 1. Types of FRET biosensors. **a**: Intermolecular FRET. FRET between donor and acceptor fluorophore attached to two separate proteins. When proteins come together, FRET is observed. **b**: Intramolecular FRET. FRET occurs between two fluorophores attached to a single protein. Upon protein activation and a conformational change, two fluorophores are brought closer together to be detectable by FRET. **c**: Protein transducer. Proteins can be engineered in such a way as to induce large shifts in conformation upon binding to certain ligands, thus bringing the two fluorophores within FRET proximity. In this example, binding of calcium induces a conformational shift in the protein to expose a hydrophobic pocket, where the peptide (in blue) binds to cause a large shift in the linker arm attached to the second fluorophore [Miyawaki et al., 1997]. **d**: Domain/antibody-biosensors. A small peptide or antibody attached to a fluorophore selectively binds to the target protein upon activation of the protein. The binding domain/antibody can be attached or not attached to the protein of interest. [Adapted from Chamberlain and Hahn, 2000]

regions of downstream effectors for the target protein of interest. Similarly, antibodies specific to the target protein can be labeled with fluorochromes and used as biosensors [Haj et al., 2002].

Signal transduction is perhaps the area in which FRET biosensor technology has had the most impact. Work in a few laboratories has targeted activation and subsequent domain-exposure events for a number of Rho-family of GTPase proteins [Kraynov et al., 2001; Del Pozo et al., 2002; Katsumi et al., 2002]. GTPase proteins were tagged with GFP in the N-terminal region, away from areas of the protein critical for bioactivity, and a binding domain derived from the effector molecules was tagged with an acceptor fluorochrome. Upon activation of the GTPase and binding of the activated GTPase to the domain-sensor probe, GFP and the fluorochrome were brought into close proximity, effecting FRET. In the example (Fig. 2), Rac GTPase tagged with GFP was expressed as a fusion protein, and PBD (PAK Binding Domain) labeled with Alexa546 dye was microinjected into cells. As can be seen from the figure, FRET signals revealed highly localized exposures of domain binding activity correlating with the direction of cell polarization, whereas the GFP and Alexa546 images showed the very different whole-cell distribution patterns of these proteins.

There are certain limitations associated with the use of FRET biosensors. Some proteins cannot be tagged with GFP or a dye without perturbing native biology. Another potential issue may be competitive binding between the domain-biosensors and native ligands. New approaches on the horizon will likely conquer these issues by engineering sensors that can bind to endogenous, untagged proteins. Alternatively, proteins can be directly labeled with novel, solvent-sensitive dyes that report protein activation, or by using circularly permuted GFP constructs [Hahn et al., 1992; Baird et al., 1999; Nalbant et al., 2001; Hahn and Touchkine, 2002; Touchkine et al., 2003].

IMAGING CELLS WITH AUTOMATED HIGH THROUGHPUT MICROSCOPY

High throughput imaging has the potential to automatically correlate cell structure, function, and behavior with the wealth of cell molecular information generated by genomics and proteomics efforts. Instead of sitting for hours, days, months, or even years in front of the microscope to understand how molecules govern cell behavior, high throughput microscopy instrumentation automates image acquisition, cell measurement (cytometry) and will soon provide high level interpretation of cell type and behavior, all at speeds orders of magnitude faster than humans.

HTM means high speed, which cannot be achieved without complete automation. The ability to place a slide or microtiter plate on

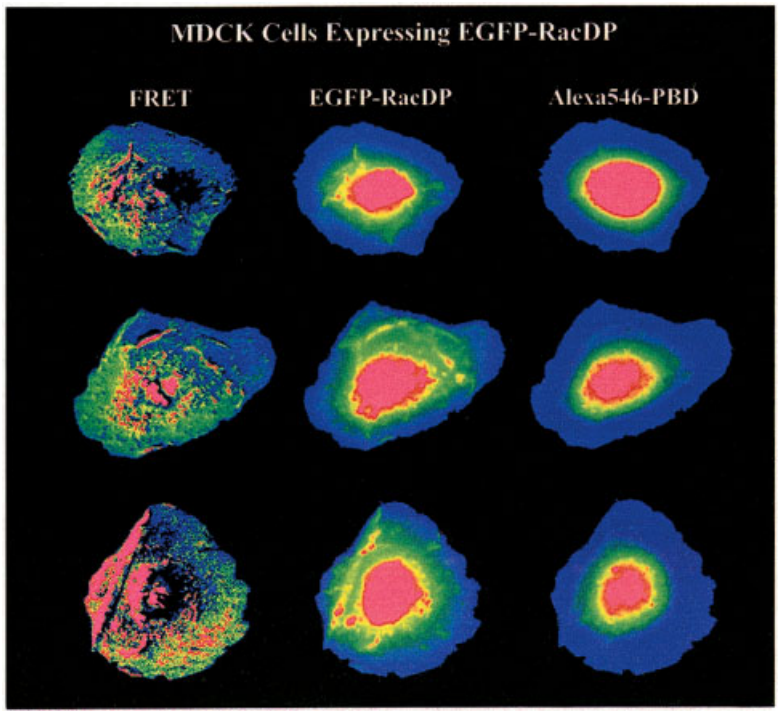


Fig. 2. Madin-Darby canine kidney (MDCK) cells expressing Rac GTPase tagged with GFP. A p21-binding domain from PAK (PBD) conjugated to Alexa 546 dye (Molecular Probes, Inc., Eugene, OR) was produced in vitro and microinjected into cells expressing the GFP-Rac GTPase fusion protein. FRET signals were detected only where the domain biosensor bound to the active Rac GTPase, showing specific localization of activity (**left column**; red, areas of high FRET signal; blue, areas of low FRET signal). Fluorescence signals from GFP-Rac GTPase (**middle column**) and the Alexa546-PBD (**right column**), show that the distribution of the protein can be seen to be quite different from that of the activated protein. Rac is activated at the leading edge of moving cells, and in the juxtanuclear region. [Figure courtesy C. Chamberlain, J. Ehrlich, J. Nelson, and K. Hahn].

the microscope, push a button, and return later with thousands to millions of cells analyzed requires automatic focus and stage motion to acquire the images, computerized image segmentation to locate the cells in the images, analysis software to collate measurements from each cell and high level tools for retrieval and visualization of the results. The process of scanning the specimen, performing autofocus and acquiring the images is shown in Figure 3. Whereas in macroscopic photography autofocus may be acceptable if it works say 9 out of 10 times with a precision of a meter or so, in high resolution microscopy, autofocus must be robust over tens of thousands of fields (depending on magnification and field of view, FOV) with submicron precision just to analyze a single microscope slide. Slides surfaces vary by as much as ± 10 μm and microtiter plate surfaces by >100 μm. The microscope itself is also unstable and temperature changes can alter focus position through ranges of 10–20 μm over 24 h in a typical lab [Bravo-Zanoguera, 2001]. This makes autofocus for microscopy a challenging task and several different criteria for finding the most in focus image have been tested. Usually for several different z-positions a focus score or focus index is calculated, which

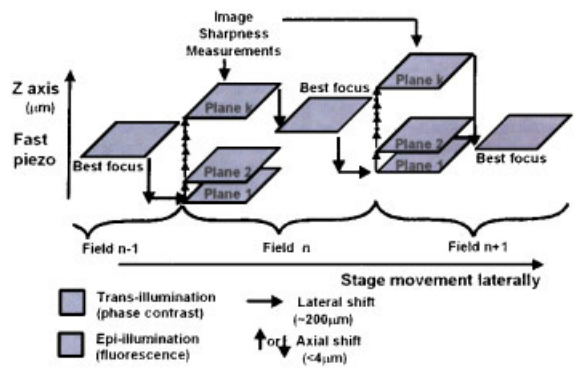


Fig. 3. Scanning and autofocus: Best focus is calculated from a series of several test focus positions (typically 5–9) and set before image acquisition and processing. The stage then moves to the adjacent field and autofocus is repeated with best focus from the previous field providing the center of the autofocus search range (~2–4 μm). Stage motion, autofocus, and imaging occur in as little as 0.3 s/field and autofocus precision is <100 nm. [Adapted from Bravo-Zanoguera, 2001].

is a measure of the resolution (high frequency content), contrast, or entropy of the image [Groen et al., 1985; Price and Gough, 1994; Bravo-Zanoguera et al., 1998; Geusebroek et al., 2000]. From these focus indices, the highest value is chosen as the optimal focus position. Research at UCSD has created HTM that operates with up to high-dry magnification objectives (0.95 numerical aperture, NA) at speeds of 3 fields/s (Hz) and recently proven technology (reports in progress) will increase this 10-fold [Bravo-Zanoguera and Price, 1998; Bravo-Zanoguera et al., 1999; Nguyen et al., 2000]. This instrumentation is based on autofocus that works in 0.25 s with a precision of < 100 nm (SD) [Price and Gough, 1994; Bravo-Zanoguera et al., 1998; Oliva et al., 1999] and lamp stabilization with 10–30 fold lower coefficients of variation (CVs) than the conventional Hg vapor lamp [Heynen et al., 1997]. This system has been demonstrated to be able to find as few as 1:20 million cells in ultra-rare cell detection (with analysis of 5 million cells/slide) [Bajaj et al., 2000].

Each image can contain from a few large cells at high magnification to thousands of small cells at low magnification. In HTM, the computer locates each cell by segmenting the image into regions. The region of pixels is then used to make measurements of, e.g., quantity, shape, pattern, and distribution of labeled intracellular molecules. There has been a wide range of image segmentation methods applied to biomedical imaging, among which are global or adaptive thresholding, texture based classification, region growing, neural networks, Markov random field approaches, and adaptive contours [Fu and Mui, 1981; Pal and Pal, 1993; Price et al., 1996; Saeed, 1998]. All these methods perform with different accuracies, different execution times and require different levels of user-interaction, which makes them suitable for certain specific applications but unsuitable for others. With images acquired at 3–30 Hz, image segmentation must be fast and because subsequent cytometry accuracy is dependent on correctly identifying pixels belonging to the cell, it must also be accurate. Because cells stained with fluorescent dyes exhibit large differences in intensity, simple thresholding does not perform well and other techniques untested on cell images [Pal and Pal, 1993] are too slow. A real-time image segmentation method for fluorescently stained cell nuclei makes cell identification largely independent of fluorescence intensity [Price et al., 1996]. The method utilizes nonlinear least-squares-designed finite impulse response (FIR) filters to create marked object-background contrast for automatic histogram-based thresholding. The contrast improvement shown on the left in Figure 4 is the key to accurate automatic image segmentation. This contrast makes result much more threshold-independent; dim cells are segmented with the same accuracy as bright ones over a much larger intensity range than was previously possible. Segmentation accuracy was 93% (percent of correct pixels as compared with computer-assisted human segmentation) [Price et al., 1996].

In addition to precise autofocus and accurate computerized image segmentation, automation must include collation of data-and-images in a relational database for rapid plotting, sorting, and access to images of cell subpopulations characterized by specified measurement ranges. Even with these tools, plotting data and iteratively experimenting with different measurement ranges to find say the G2/M dividing cells or perhaps ultimately the cancer cells mixed in with normal ones in a clinical sample, is too tedious. The machine thus needs the intelligence to locate groups of cells with shared characteristics automatically and report them. Examples of higher-level machine intelligence that may lead to this kind of automation follow in the next two sections.

PROSPECTING FOR SUBCELLULAR LOCALIZATION AND PATTERN USING MACHINE INTELLIGENCE

As genome sequencing efforts reached completion over the past few years, a major effort was begun to identify the genes expressed in many cell types under various conditions (e.g., via DNA chips). More recently, as the divergence between expression levels of RNA and protein for many genes has been appreciated, the focus has shifted again from measuring RNA levels to measuring protein expression. Commonly used proteomics methods included gel electrophoresis, mass spectroscopy, computational prediction methods, and protein chips [Yarmush and Jayaraman, 2002]. Proteomics methods allow investigation of protein expression for a particular cell type under a specific condition as well as examination of protein structure, function, and interactions. However,

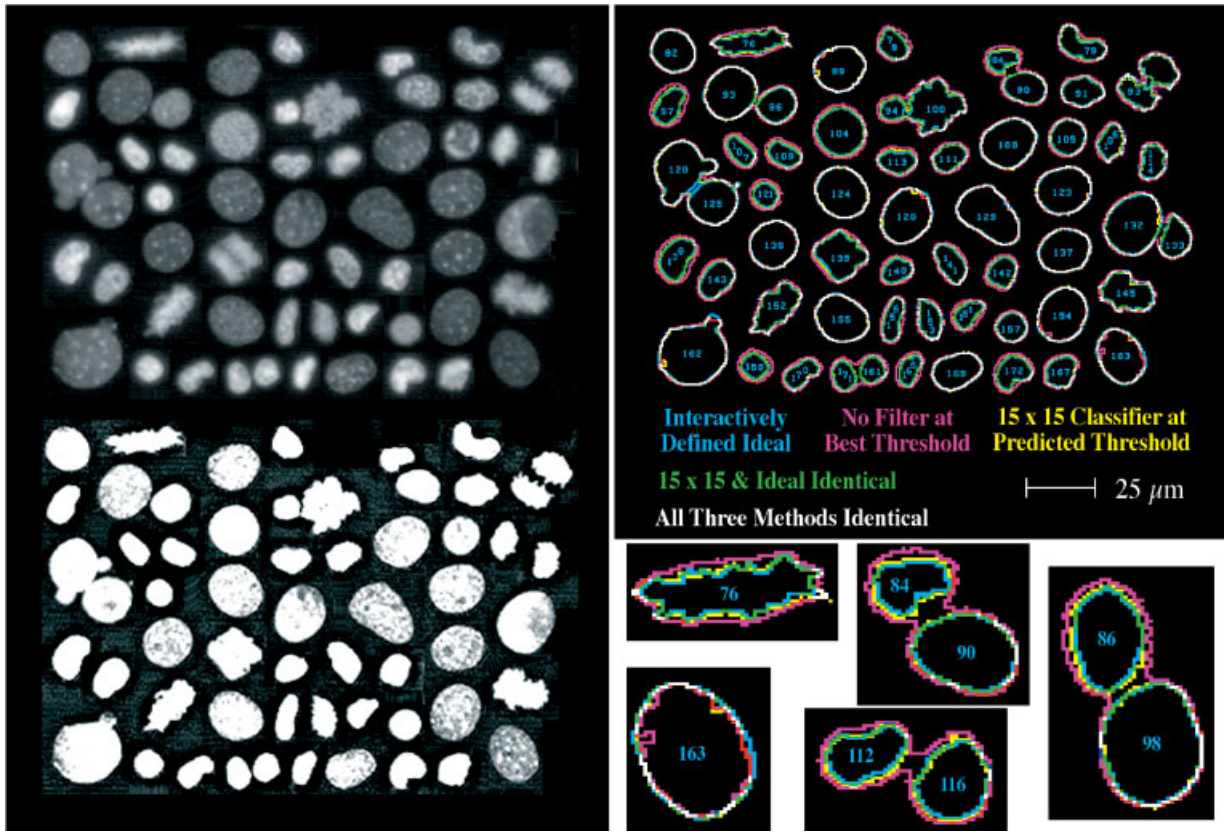


Fig. 4. Contrast enhancing least-squares-designed FIR image segmentation. Original montage image of DAPI stained cell nuclei (**upper left**), contrast enhanced filtered image (**lower left**), resulting edges (**upper right**) in yellow overlaid on the edge maps from the best threshold of the original image (magenta) and the manually segmented ideal edge maps (cyan). Magnified views of selected edge maps (**lower right**) are shown to demonstrate the improvement afforded by these filters. [Adapted from Price et al., 1996].

high throughput methods are also needed for the analysis of protein location within cells [Boland et al., 1997; Murphy et al., 2000], since knowledge of subcellular location is required for a full understanding of a protein's function. The most common method used to examine subcellular localization is to mark a specific protein with a specific fluorescent tag and then image the cell expressing it using fluorescence microscopy. The goal of this type of fluorescent tagging is distinct from the creation of fluorescent-protein biosensors that are designed to provide insights into molecular processes in living cells [Giuliano and Taylor, 1998] as described in the first section. For determination of location patterns, the goal is to specifically avoid the environmental sensitivity required for biosensors and instead to tag as many proteins as possible in as many different locations within each protein. For this purpose, an excellent method is the creation of internal fusions with fluorescent proteins (such as GFP), a method termed CD-tagging [Jarvik et al., 1996, 2002; Telmer et al., 2002].

Extensive work over the past six years has shown that proteins have characteristic cellular patterns that can be recognized computationally in 2D [Boland et al., 1997, 1998; Murphy et al., 2000, 2002; Boland and Murphy, 2001; Danckaert et al., 2002] and 3D [Velliste and Murphy, 2002] images. This work allows us to envision high throughput identification of cellular proteins and their localization from fluorescence microscopy images. Objective computerized recognition of unique subcellular patterns eliminates interobserver variations and provides a useful "fingerprint" of each protein that may also aid understanding of function. Examples of patterns (obtained by immunofluorescence microscopy) that can be distinguished with high accuracy are shown in Figure 5. The basis for this distinction is the calculation of numerical descriptions of the patterns, termed Subcellular Location Features

[Boland and Murphy, 2001]. These features describe a range of characteristics of each image, such as the number of fluorescent objects per cell, the average distance of fluorescent objects from the center of the cell, and texture measures of the correlation between gray levels of adjacent pixels. The measured values of the features can be used to train classifiers (e.g., neural networks) and the trained classifiers can be used to assign a pattern to previously unseen images. For images showing the subcellular locations depicted in Figure 5, an average classification accuracy of $83 \pm 4.6\%$ (mean $\pm$ 95% confidence interval) has been reported for classification of single images [Boland and Murphy, 2001]. If a set of ten images from the same population is analyzed and the population class is determined by assigning the class label that a plurality of the members were assigned, the classification accuracy can be increased to 98% [Boland and Murphy, 2001]. More recent studies yielded 88% accuracy for single 2D images accompanied by a parallel DNA image [Murphy et al., 2002] and 91% for single 3D images [Velliste and Murphy, 2002]. As illustrated in Table I, a key result from these studies is that the systems can discriminate between patterns that are indistinguishable by human observers (the patterns of the Golgi proteins giantin and gpp130 can be distinguished with greater than 80% accuracy while human observers cannot distinguish them with greater than the 50% accuracy that results from guessing). With such high classification accuracies, automated analysis of cellular protein localization from fluorescence images can be expected to become an independent tool for understanding the structure-function relationships of proteins.

The development of automated screening and analysis methods for microscope images will have a significant impact in several fields [Giuliano and Taylor, 1998; Boland and Murphy, 1999]. The quantitative description and automated classification of protein localization provides the opportunity for a standardized database against which a protein pattern can be compared and classified [Huang et al., 2002]. The Subcellular Location Features have also been used to provide an objective method for comparing sets of protein patterns, such as for a

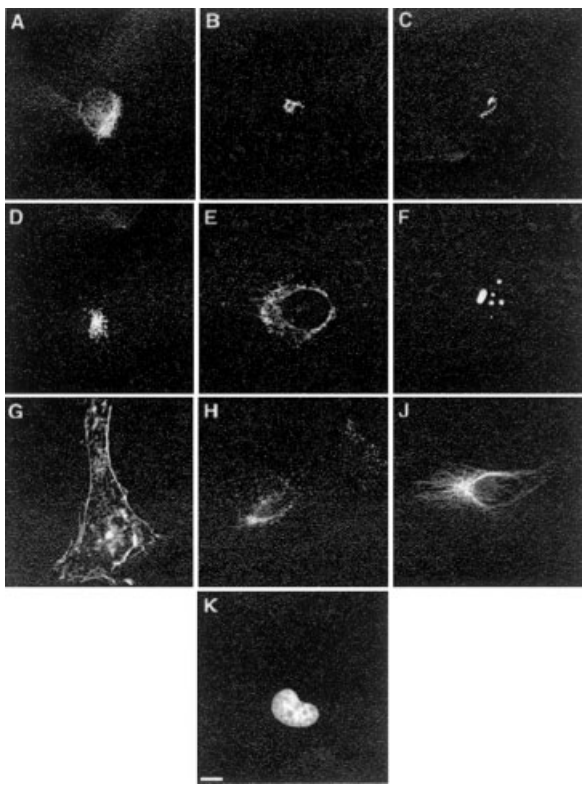


Fig. 5. Representative images from the classes used as input to the classification systems described in the text. These images have had background fluorescence subtracted and have had all pixels below a threshold set to 0. Images are shown for HeLa cells labeled with antibodies against ER protein (**A**), the Golgi protein giantin (**B**), the Golgi protein GPP130 (**C**), the lysosomal protein LAMP2 (**D**), a mitochondrial protein (**E**), the nucleolar protein nucleolin (**F**), transferrin receptor (**H**), and the cytoskeletal protein tubulin (**J**). Images are also shown for filamentous actin labeled with rhodamine-phalloidin (**G**) and DNA labeled with DAPI (**K**). Scale bar = 10 μm. [From Boland and Murphy, 2001.]

TABLE I. Classification Results for 3D Confocal Images of HeLa Cells

	Output of Classifier									
True Class	DN	ER	Gia	gpp	LA	Mit	Nuc	Act	TfR	Tub
DNA	**99**	0	0	0	0	0	0	0	0	0
ER	0	**89**	0	0	0	0	0	2	2	7
Giantin	0	0	**90**	3	2	4	1	0	0	0
gpp130	0	0	5	**81**	9	0	0	0	4	0
LAMP2	0	0	1	4	**90**	2	0	1	2	0
Mitoch.	0	1	1	0	0	**96**	0	1	1	0
Nucleolin	1	0	0	0	0	0	**98**	0	0	0
Actin	0	2	0	0	0	2	0	**92**	3	0
TfR	0	1	0	0	4	3	0	2	**85**	5
Tubulin	0	5	0	0	0	0	0	0	4	**91**

Confocal images of the same markers shown in Figure 1 were collected, 28 numerical features were calculated for each image (feature set 3D-SLF9), and a neural network classifier was trained with a portion of the images. The prediction of the classifier for the remainder of the images was compared to the true class; the values in the table show the fraction of the images that are actually from a given class (shown in the rows) that were classified as belonging to a particular class (shown in the columns). The overall accuracy is 91% across all 10 classes. [From Velliste and Murphy, 2002.]

protein in the presence and absence of a drug [Roques and Murphy, 2002]. The addition of a protein localization pattern database to the existing protein structure databases will provide exciting new insights into the relationships between function, transport, and interaction of different proteins in living cells. Combining the computerized ability to obtain protein pattern signatures with the high throughput automated acquisition of images described in the previous section, provides the prospect for rapid development of comprehensive databases that store this information for virtually all cellular proteins.

PROSPECTING FOR NEW CYTODIAGNOSTICS BY IMAGING MOLECULAR LABELS

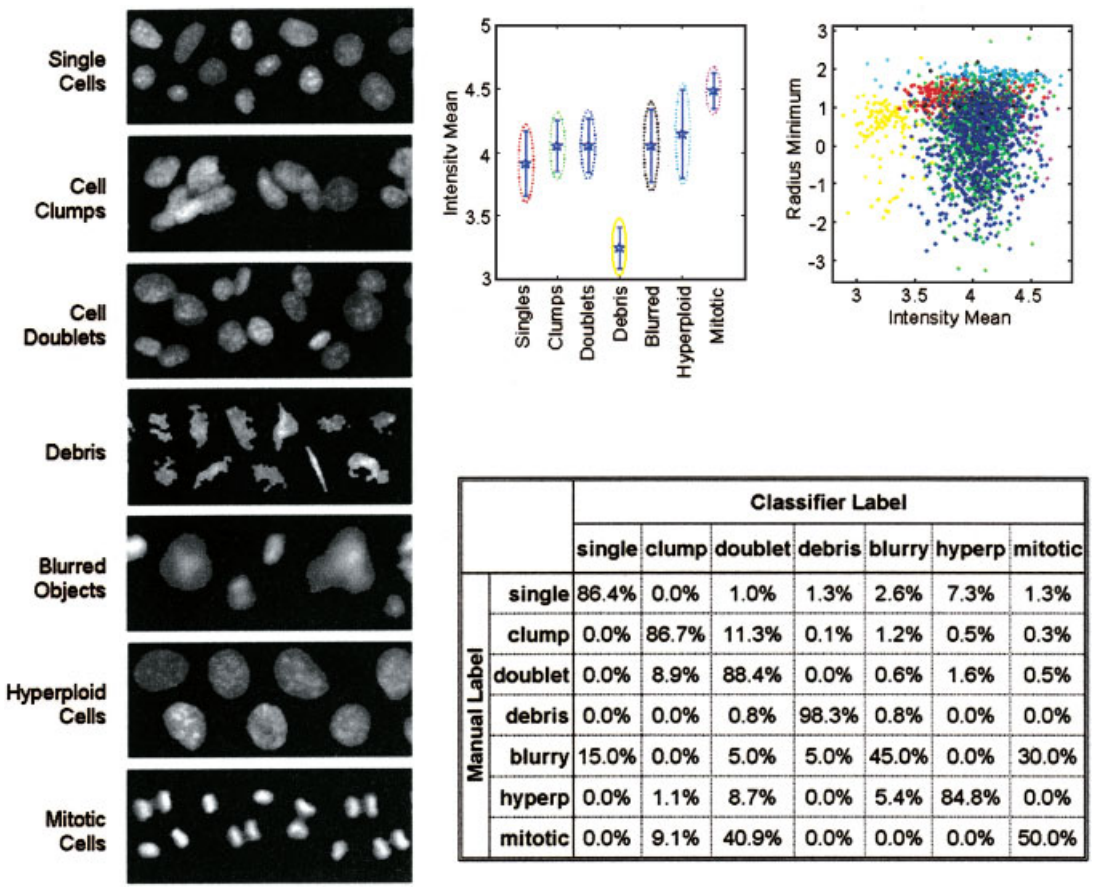
Over the last few decades, the advances in computation speed, imaging, and high throughput analysis methods have opened up new diagnostic opportunities in the biological and health sciences. For example, the search for molecular markers for cancer detection, a subset of molecular biomarkers, has continued to expand [Srivastava and Gopal-Srivastava, 2002; Wu et al., 2002]. Biomarkers that have become useful in cancer diagnostic include prostate specific antigen (PSA), cancer antibody or tumor marker 125 (CA125), cancer antigen 15-3 (CA15-3), several cytokeratins, some cell surface antigens (MUC1 & MUC2), and several growth factors (TGF-α, TGF-β, erb-2, erb-3) [Brotherick et al., 1998; Braun et al., 1999; Zimmerman et al., 2000; Srivastava and Gopal-Srivastava, 2002]. The combination of new biomarkers and advanced high-throughput imaging technologies for analyzing subcellular molecular characteristics using, e.g., fluorescence microscopy, offers tremendous potential for new cytodiagnostics.

If one or more biomarkers for a specific disease are found, high throughput imaging methods may be applied to screen body fluids containing exfoliated cells for early cancer detection. Automated microscopy image acquisition at high throughput rates for slides, cell culture chambers, or well plates has made substantial progress [Price, 1990; Bravo-Zanoguera et al., 1999; Bajaj et al., 2000] and higher-level image interpretation may provide powerful automated diagnoses. The addition of sophisticated image analysis methods will provide new, completely automated imaging avenues for studying, classifying, and interpreting cellular molecules. These methods may also aid in understanding the underlying molecular disease mechanisms and thereby facilitate development of new anticancer therapeutics [Giuliano and Taylor, 1998; Srivastava and Gopal-Srivastava, 2002; Wu et al., 2002].

High-level image interpretation has been used in a wide range of applications [Hudson and Cohen, 2000; Duda et al., 2001] including microscope images of: cervical smears [Kemp et al., 1997; Mackin et al., 1998; Van der Laak et al., 2002]; premalignant prostate, colon and esophageal tissue [Weyn et al., 2000]; and cultured cells [Boland et al., 1998; Boland and Murphy, 2001]. However, automated image interpretation using classical clinical stains (e.g., hematoxylin and eosin, and variations of the Pap stain) has only shown limited success [Bartels and Vooijs, 1999]. The combination of using fluorescent biomarkers and automated high-level image interpretation offers an exciting new alternative for classification (diagnoses) of cells and tissue. One cellular diagnostic category enabled by high throughput cellular classification is detection of rare and ultra-rare cells in blood and other large cell populations. Classification of cellular objects in images of tissue stained with specific fluorescent labels is performed as shown in Figure 6. In this simple example, all the objects (including debris) encountered in HTM images of large populations of fixed cells are first grouped into single cells, cell clumps, cell doublets, debris, blurred objects, hyperploid cells, and mitotic cells. This highlights the fact that in fully automated imaging, the system must first identify general classes of cells in order to then carry out, e.g., cancer versus normal classification. This is because a mitotic cancer cell may differ so much from a normal G0/G1 cell that it could not be classified. That is, it may be necessary to find mitotic and G0/G1 cells and then classify normal and cancer mitotic cells and normal and cancer G0/G1 cells. From the images, the objects are extracted using image segmentation techniques [Duda and Hart, 1973; Price et al., 1996; Jaehne, 1997] and different measurements, referred to as features, are performed on these objects. The features with the highest discriminating power are selected and combined in the feature set (too large of a number of features makes the classifier too complex and requires a large amount of training data, a too

small number of features might not separate the classes with the desired accuracy). An example of one pair of features (mean intensity and radius minimum) in the set used for the seven classes is shown on the upper right panel of Figure 6. A classifier (e.g., a linear discriminant function, a neural network, a probability-based Bayesian classifier, a nearest neighbor classifier, or clustering techniques [Hudson and Cohen, 2000; Duda et al., 2001]) processes the features and finds classes of grouped features that characterize types of cells. In most cases the classifier requires a training set to determine its parameters. The performance of the developed classifier on naïve test data is reported as confusion matrix, shown at the lower right of Figure 6, where the diagonal represents the correctly classified number of cells. Such a confusion matrix also allows analysis of similarities of classes, e.g., for this case the 'mitotic cells' had a very large confusion ratio with 'doublets,' about 40%. This suggests similarity in odd shapes occurring both in 'doublets' and in 'mitotic cells.'

The resulting performances of the automated high-level image interpreters can vary widely (anywhere between 60 and 100%) depending on how similar the object classes are, how discriminating the molecular biomarker(s) used for imaging is (are), and whether the differences in the imaged objects are simple on-off decisions, like for rare event detection of fetal nucleated red blood cells in maternal peripheral blood [Bianchi, 1999; Bajaj et al., 2000; Bianchi et al., 2002; Yamanishi et al., 2002], images of which are shown in Figure 7 (left), or fine differences, e.g., in texture of the molecular marker that can classify cells with malignancy associated changes [Kemp et al., 1997]. An example of a breast cancer specific marker, anti-cytokeratin-



		Classifier Label						
		single	clump	doublet	debris	blurry	hyperp	mitotic
Manual Label	single	86.4%	0.0%	1.0%	1.3%	2.6%	7.3%	1.3%
	clump	0.0%	86.7%	11.3%	0.1%	1.2%	0.5%	0.3%
	doublet	0.0%	8.9%	88.4%	0.0%	0.6%	1.6%	0.5%
	debris	0.0%	0.0%	0.8%	98.3%	0.8%	0.0%	0.0%
	blurry	15.0%	0.0%	5.0%	5.0%	45.0%	0.0%	30.0%
	hyperp	0.0%	1.1%	8.7%	0.0%	5.4%	84.8%	0.0%
	mitotic	0.0%	9.1%	40.9%	0.0%	0.0%	0.0%	50.0%

Fig. 6. Examples of automated cell classification using the molecular DNA probe DAPI. In this case classification was performed to distinguish the different cell objects encountered during cell-by-cell analysis of slides using HTM. The **left panel** shows the seven different object classes, the **right top panel** shows a scatter plot of two of the features selected for the feature set, and the **right bottom panel** gives the results of a linear discriminant function classifier for previously unseen cell objects. [From Heynen, 2002.]

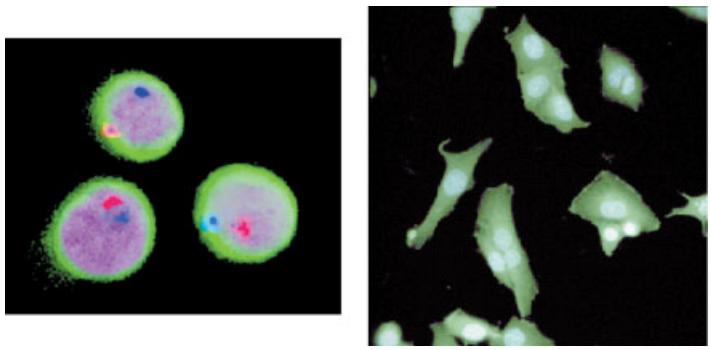


Fig. 7. Left: Fetal nucleated RBCs stained with centromeric FISH probes for the X- and Y-chromosomes, DAPI for the nuclei and anti-γ-globin-FITC for fetal hemoglobin. **Right**: FITC-DAPI cell image obtained showing cytokeratin-19-positive cells (green) with DAPI-stained cell nuclei (blue).

19 is shown in Figure 7 (right). In these and other applications, molecular imaging has the potential to alter cytodiagnostics, especially when combined with powerful new machine intelligence algorithms for interpreting images of millions of cells acquired with high-throughput microscopy.

PROSPECTING FOR MARKERS BY IMAGING TISSUE MICROARRAYS

Molecular microarrays, including DNA, RNA and protein arrays, have become important tools for finding disease characteristics and mapping normal expression by testing patterns from thousands of molecules on one or a few homogenized tissues. Tissue microarray (TMA) analysis is a complementary method that can test the resulting hundred or so molecular array hits on many intact tissue sections on a single slide. Hundreds to thousands of cylindrical cores (of 0.6–1.5 mm diameter and up to 2.5 mm long) are biopsied from tissue blocks, and arrayed in a single new paraffin block [Kononen et al., 1998; Skacel et al., 2002]. The new block is sectioned and placed on slides, and the molecular patterns are analyzed directly in the tissues to confirm that they are unique to a particular type of tissue [Kononen et al., 1998]. Validation of normal versus cancer expression or one type of normal tissue against another is a key advantage of TMAs over molecular arrays. The presence of many tissue samples on one slide dramatically increases the processing and analysis rates. Many patients with, e.g., breast cancer, can be studied for protein expression patterns simultaneously after identical preparation. A single tissue microarray block can be analyzed for genotype or expression by DNA and RNA in situ hybridization (ISH, or fluorescence, FISH), and by immunohistochemistry (IHC) or immunofluorescence (IF) [Skacel et al., 2002].

In the program on Apoptosis and Cell Death Regulation at the Burnham Institute, researchers can immunostain up to 240 TMA slides/week, but it takes one pathologist 2–3 h to read each slide, making the analysis about 10-fold more time consuming than the preparation (and preparation automation efforts continue, see e.g., Beecher Instruments, www.beecherinstruments.com). Although several semi-automated TMA readers are commercially available [examples include: Chromavision, Inc., (www.chromavision.com), Biogenix (www.biogenex.com), and Aperio (www.aperio.com)], fully automating the process of scoring the TMAs is challenging because it means both machine identification of normal versus cancer tissue (or different types of normal tissue) and densitometry of expression markers in clinically stained samples. Using HTM developed at UCSD, TMA slides can be scanned and the images collated and montaged for scoring and storing in databases as shown in Figure 8. Although some researchers are exploring the use of fluorescence markers [Camp et al., 2000], pathologists often prefer to diagnose the cancer tissue using clinical stains [Krajewski et al., 1994, 1999; Krajewska et al., 2002]. Thus, whereas measuring the intensities of a few fluorescence signals with wavelengths is relatively trivial, separating overlapping broad-spectrum wavelengths of, e.g., hematoxylin and the immunohistochemical stain diaminobenzidine (DAB), is more challenging. Multispectral techniques are being explored and have shown great promise for determining the contribution of each dye to the color and density of each pixel [Gat, 2000; Levenson and Hoyt, 2000; Macville et al., 2001; Ruifrok and Johnston, 2001]. We have also begun exploring texture

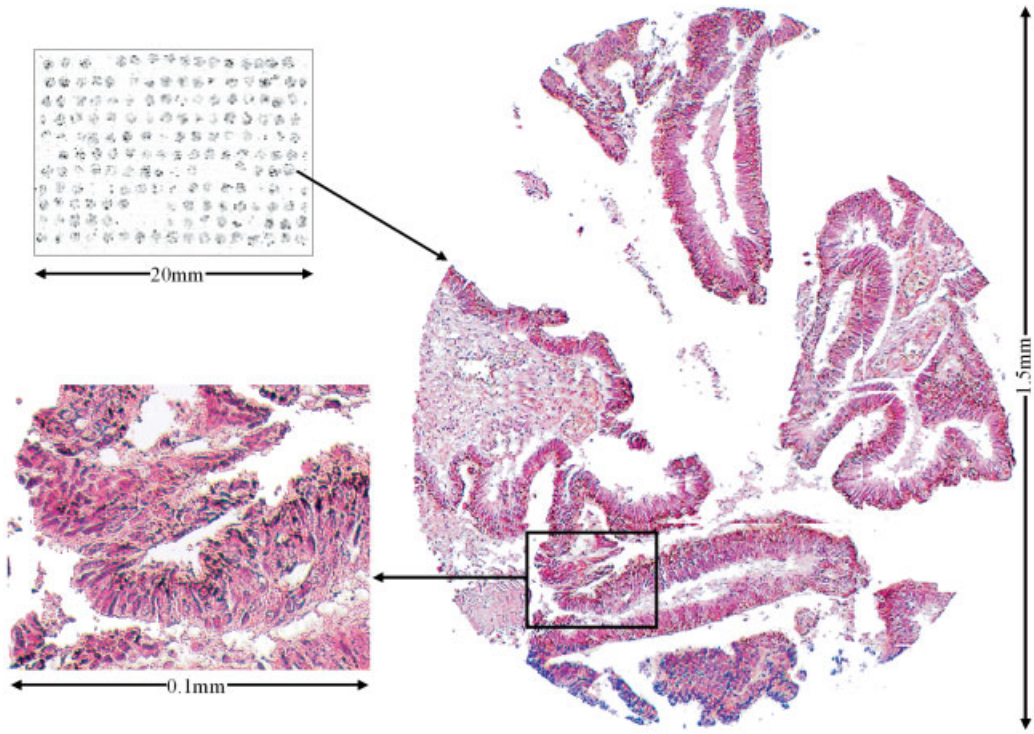


Fig. 8. A rectangular area was automatically scanned at 20× by HTM (with several images/core) and the images were processed to produce a downsampled montage image (**upper left**). The individual cores were automatically located in the downsampled montage and used to guide computerized reassembly of high-resolution montages of each core (**right**). The native high-resolution views were then available by selecting regions for zooming (**lower left**).

segmentation techniques that may successfully separate normal and cancer tissue. As an example, Figure 9 shows a texture-based technique largely successfully delineating normal glandular and stromal tissue.

Combining automated TMA preparation with fully automated analysis holds promise for very rapid prospecting of molecular patterns of cancer. Database software for storing and retrieving TMA images and data will provide efficient access to the large sets of data generated by automated processes [Manley et al., 2001]. For rapid understanding of the relationship of this molecular information to disease and treatment, the patient treatment and outcome data must also be available to clinical researchers. A multi-hospital tissue and clinical records database that can serve as a model for completing the picture is under construction in Australia (Personal communication, Nikolajs Zeps, Western Australia Institute for Medical Research, and www.nbcf.org.au/prioritie-sperth.shtml#d). Ultimately, worldwide clinical databases that facilitate tracking of treatment and outcome data as a function of molecular markers would create an unparalleled opportunity for cancer diagnostic and therapeutic discoveries.

PROSPECTING FOR DRUGS BY IMAGING SUBCELLULAR MOLECULAR ACTIVITY

Biological drug discovery is struggling to integrate the flood of new information emerging from the genomics and proteomics revolutions. While many high throughput tools provide lists of genes and proteins in cells, pharmaceutical researchers are increasingly asking for a comprehensive tool set for rapidly acquiring and integrating cell-molecular, cell-structural, and cell-functional data linked directly to compound responses [Giuliano et al., 1997]. Although the simplest living cell is probably too complex for exact modeling by the most powerful computers, even relatively simple models may shake the foundations of biology [Gibbs, 2001]. A typical high throughput screening (HTS) instrument processes thousands to tens of thousands of compounds a day by rapidly measuring the bulk response of all of the cells in each well in a

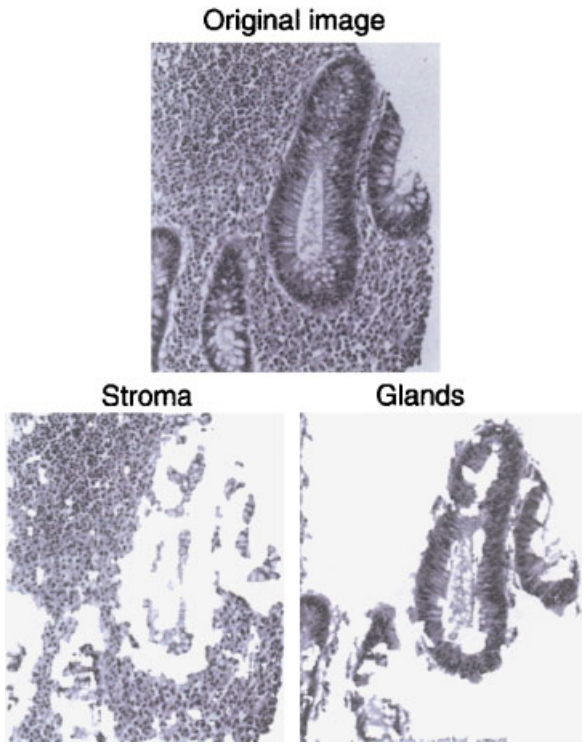


Fig. 9. The original image containing both stroma and glandular regions was broken in two independent regions based on texture elements.

microtiter plate. The amount of a cellular product, such as cytokine interleukin 2 (IL2), can be assayed in this manner as an end-point to inhibition or excitation by a library of compounds. A given compound may result in, e.g., the secretion of IL-2, but the researcher does not know which signaling pathway was utilized, or if there were other cellular responses. If the signaling pathway used also led to cell toxicity or death, the efficacy of the candidate drug would be compromised and the information might not be learned until failure in more costly and controversial animal testing. Multiplexed cellular responses need to be investigated to decrease these kinds of false positive leads in drug discovery. Dissecting the steps in cellular pathways is important because multiple pathways converge and diverge to provide redundancy and for coordinated cellular behavioral responses.

A number of commercially available instruments (including those from Cellomics, Universal Imaging, Axon, Q3DM, and Amersham) are addressing the need for automated microtiter plate instruments that automatically provide cell measurements from images. Nuclear-cytoplasmic translocation is a commonly addressed cellular dynamic whereby a labeled protein moves between the cytoplasm and the nucleus upon stimulation [Auphan et al., 1995; Verma et al., 1995; Ding et al., 1998; Mercié et al., 2000]. In Figure 10, some results of a nuclear-cytoplasmic translocation experiment involving stimulation of the GR receptor by dexamethosone are shown. The software in Figure 10 is also an example of the integrated image acquisition, processing, and analysis that provides for walk-away automation of these kinds of measurements over large numbers of cells, wells, and microtiter plates. Automation of cell-by-cell measurements for routine use in every assay has been very challenging. Advances in image segmentation (see Fig. 4) have led to instrumentation where cell-by-cell measurements are the basis of all image cytometry assays. Another type of translocation where image segmentation enabling cell-by-cell measurements is even more challenging is the formation of pits (or vesicles) as shown in Figure 11. In this example, the pits are very small and the measurement precision is much better using higher resolution (0.95 NA) optics. The corresponding dose response curve is for stimulation of pit formation by isoproterenol.

Automation and speed are important but must be accompanied by measurement fidelity for efficient analyses. Figure 12 shows that the number of cells required for a given statistical significance increases exponentially for a given SD as statistical significance requirements become more stringent. This family of curves demonstrates that it may be much less costly if high precision measurements can be made because lower SDs achieve the same minimum significant response with much fewer cells. This dependence on the fidelity of image measurements may also be true in the examples of high-level cell classification in the previous sections. The impact of the measurement precision is easy to overlook but may have substantial contributions to the usefulness of the conclusions.

The examples of translocation in Figures 10 and 11 included stains for both the protein of interest and the DNA. In preliminary NFκB translocation experiments performed under similar conditions as previously published [Ding et al., 1998], the SD of FLIN was found to be smaller for G0/G1 cells. In transient translocation experiments, only a few cells may be

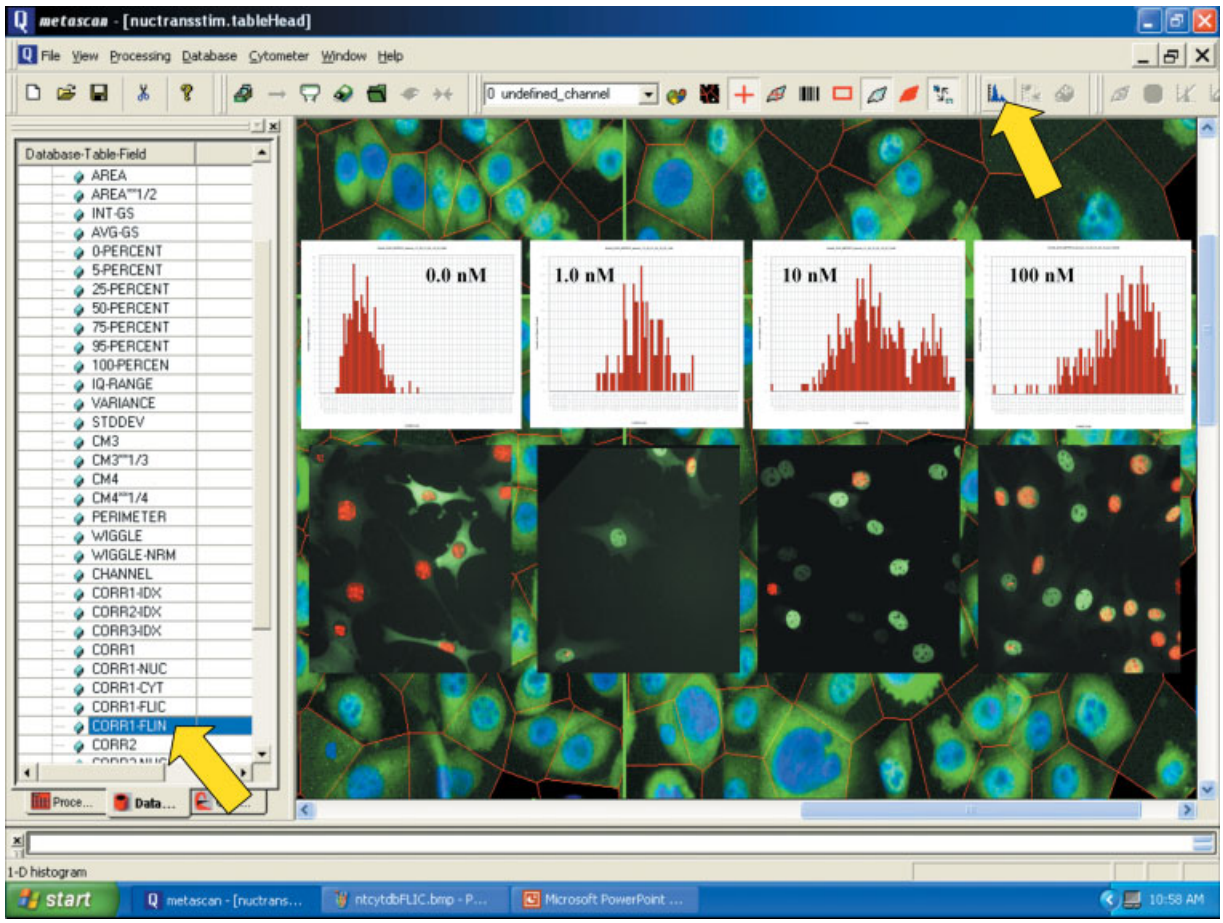


Fig. 10. Cytoshop™ software by Q3DM, Inc., shows the cytometrics list (**left**), an image table with a pan-and-scroll view of the entire scan area (background, large window) and a series of images with corresponding histograms of the fluorescent localized intensity of the nucleus (FLIN = nuclear intensity/cellular + nuclear intensity). Cells were labeled with GFP for the glucocorticoid receptor (GR) and stimulated with increasing concentrations of dexamethosone, which stimulates GR causing it to translocate into the nucleus. [Labeled cells courtesy of Gordon Hager, National Institutes of Health.]

expressing the desired tag as verified by a co-transfected label. Analyzing only the transfected cells, which can represent as little as a few percent of the total cells, is very important for measuring the dose responses. Cell heterogeneity increases dose response complexity. Even in cloned cell lines, cell cycle position and phenotypic variations create differences in cell

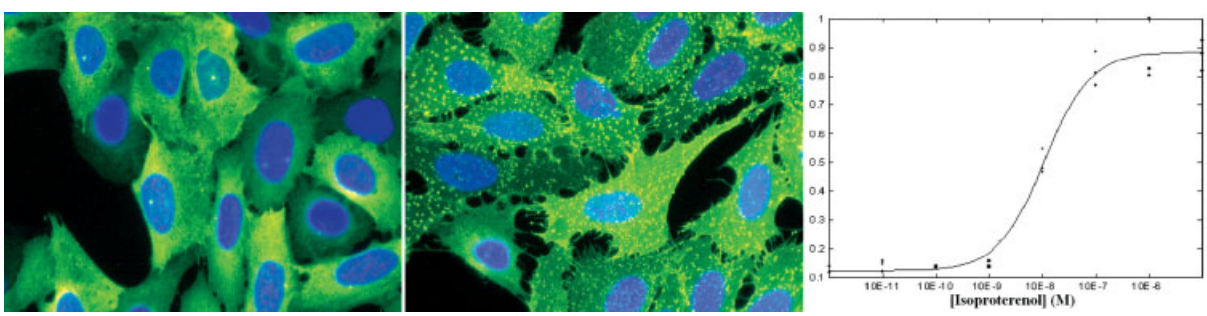


Fig. 11. Pit formation in human osteosarcoma cells. Example 40x 0.95 NA fluorescent micrographs of the Norak Transfluor™ GPCR pit formation assay using the Q3DM EIDAQ™ 100 HTM instrument. GFP expression (green) and Hoechst staining (blue) are visualized before (**left**) and after (**middle**) pit formation. The dose response curve (**right**) shows measured normalized pit formation in U20S cells expressing the beta-2-adrenergic receptor and beta arrestin 2 fused to GFP as a function of isoproterenol concentration. The $EC_{50} = 1.48$ nM (±0.11). (Cells courtesy of Norak Biosciences, Morrisville, NC).

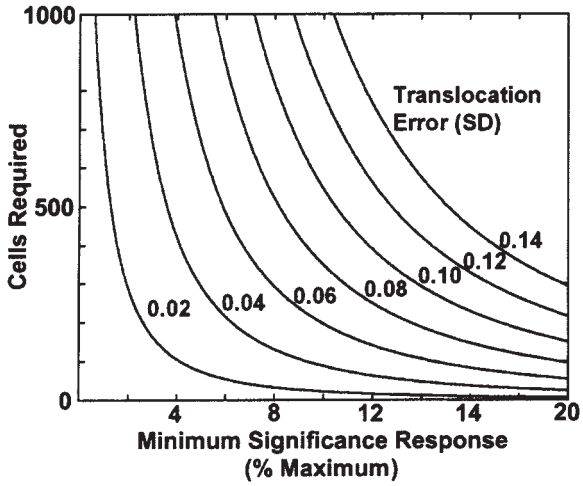


Fig. 12. The number of cells (and hence the time) required as a function of minimum significant response. The family of curves are for a range of underlying cell-by-cell standard deviations. $\alpha = 0.05$ Pr(identical samples show response, type I error); $\beta = 0.20$ Pr(different samples show no response, type II error).

behavior. A lead that acts on a cellular process correlated to DNA synthesis will elicit a maximal change in response in S-phase cells. Averaging the response from all the cells in a well may require laborious cell cycle synchronization or result in overlooking an effective drug. Image analyses of multiplexed DNA and protein stimulus-response labels correlates the compound response to the cell cycle phase thereby providing direct elucidation of more complex compound responses. These are relatively simple examples of the use of multiple markers. As experiments using high throughput cell-by-cell imaging become more common, the demand for automated analyses of multiplexed intracellular probes for direct measurements of more complex cellular responses is only likely to increase. The prospects for improved living cell labels of protein activity (summarized in the first section) are likely to continue to drive the need for multiplexed probes and measurements.

CONCLUSIONS AND PROSPECTS FOR RICH INFORMATION FROM CELLULAR MOLECULAR IMAGING

The tremendous prospects for molecular imaging are best recognized in the context of advances in both computing and biology. Engineers and computer scientists (and perhaps societies as a whole) have adapted to the expectation of speed and storage continuing to advance at least at the Moore's law rate of two fold every 18 months. Scientists have experienced some of the potential benefits for biological research through the relatively rapid advances in gene sequencing over the last 10–15 years that were enabled by high throughput technologies. The quest to understand cellular molecular biochemistry from molecular activity to cellular behaviors will be similarly enabled by high throughput technologies. But the information explosion represented in the detailed molecular structure and function of hundreds of thousands of different proteins in each cell along with detailing signaling pathways will require more than just a list. Computer models (see, e.g., Physiome Sciences, www.physiome.-com) and high-level machine intelligence will also rapidly collate this information. And although these computer models may never lead to a perfect understanding of the cell [Gibbs, 2001], the prospects are intriguing.

ACKNOWLEDGMENTS

Funding for the work reported here included: National Science Foundation Major Research Instrumentation grant BES-9871365, the Whitaker Foundation Biomedical Engineering Research Grant Program, National Institutes of Health grants R44 RR15170 and R33 CA83219, and the WPC Research and Education Fund.

REFERENCES

Adams SR, Harootunian AT, Buechler YJ, Taylor SS, Tsien RY. 1991. Fluorescence ratio imaging of cyclic AMP in single cells. Nature 349:694–697.

Auphan N, DiDonato JA, Rosette C, Helmberg A, Karin M. 1995. Immunosuppression by gluco-corticoids: Inhibition of NF-κB activity through induction of IκB synthesis. Science 270:286–290.

Baird GS, Zacharias DA, Tsien RY. 1999. Circular permutation and receptor insertion within green fluorescent proteins. Proc Natl Acad Sci USA 96:11241–11246.

Bajaj S, Welsh JB, Leif RC, Price JH. 2000. Ultra-rare-event detection performance of a custom scanning cytometer on a model preparation of fetal nRBCs. Cytometry 39:285–294.

Bartels PH, Vooijs PG. 1999. Automation of primary screening for cervical cancer—Sooner or later? Acta Cytologica 43:7–12.

Bianchi DW. 1999. Fetal cells in the maternal circulation: Feasibility for prenatal diagnosis. Br J Haematol 105: 574–583.

Bianchi DW, Simpson JL, Jackson LG, Elias S, Holzgreve W, Evans MI, Dukes KA, Sullivan LM, Klinger KW, Bischoff FZ, Hahn S, Johnson KL, Lewis D, Wapner RJ, de la Cruz F. 2002. Fetal gender and aneuploidy

detection using fetal cells in maternal blood: Analysis of NIFTY I data. Prenat Diagn 22:609–615.
Boland MV, Murphy RF. 1999. Automated analysis of patterns in fluorescence-microscope images. Trends Cell Biol 9:201–202.
Boland MV, Murphy RF. 2001. A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. Bioinformatics 17:1213–1223.
Boland MV, Markey MK, Murphy RF. 1997. Classification of protien localization patterns obtained via fluorescence light microscopy. Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. p 594–597.
Boland MV, Markey MK, Murphy RF. 1998. Automated recognition of patterns characteristic of subcellular structures in fluorescence microscopy images. Cytometry 33:366–375.
Braun S, Hepp F, Sommer HL, Pantel K. 1999. Tumor-antigen heterogeneity of disseminated breast cancer cells: Implications for immunotherapy of minimal residual disease. Int J Cancer 84:1–5.
Bravo-Zanoguera ME. 2001. Autofocus for high speed scanning in image cytometry. Dissertation, University of California San Diego.
Bravo-Zanoguera ME, Price JH. 1998. Simultaneous multiplanar image acquisition in light microscopy. SPIE Proc Optical Diagnostics of Biological Fluids and Advanced Techniques Analytical Cytology 3260:194–200.
Bravo-Zanoguera ME, von Massenbach B, Kellner AL, Price JH. 1998. High performance autofocus circuit for biological microscopy. Review of Scientific Instruments 69(11):3966–3977.
Bravo-Zanoguera ME, von Massenbach B, Price JH. 1999. Automatic on-the-fly focusing for continuous image acquisition in high-resolution microscopy. SPIE Proc Optical Diagnostics of Biological Fluids and Advanced Techniques Analytical Cytology 3604:243–252.
Brotherick I, Robson CN, Browell DA, Shenfine J, White MD, Cunliffe WJ, Shenton BK, Egan M, Webb LA, Lunt LG, Young JR, Higgs MJ. 1998. Cytokeratin expression in breast cancer: Phenotypic changes associated with disease progression. Cytometry 32:301–308.
Camp RL, Lori CA, Rimm DL. 2000. Validation of tissue microarray technology in breast carcinoma. Lab Invest 80:1943–1952.
Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. 1994. Green fluorescent protein as a marker for gene expression. Science 263:802–805.
Chamberlain C, Hahn KM. 2000. Watching proteins in the wild: Fluorescence methods to study protein dynamics in living cells. Traffic 1:755–762.
Danckaert A, Gonzalez-Couto E, Bollondi L, Thompson N, Hayes B. 2002. Automated recognition of intracellular organelles in confocal microscope images. Traffic 3: 66–73.
Del Pozo MA, Kiosses WB, Alderson NB, Meller N, Hahn KM, Schwartz MA. 2002. Integrins regulate GTP-Rac localized effector interactions through dissociation of Rho-GDI. Nature Cell Biol 4:232–239.
Ding G, Fischer P, Boltz R, Schmidt J, Colaianne J, Gough A, Rubin R, Uiller D. 1998. Characterization and quantitation of NFκB nuclear translocation induced by interleukin-1 and tumor necrosis factor-α. J Biol Chem 273:28897–28905.
Duda RO, Hart PE. 1973. Pattern classification and scene analysis. New York: John Wiley & Sons.
Duda RO, Hart PE, Stork DG. 2001. Pattern classification. 2nd Edition. New York: John Wiley & Sons.
Fu KS, Mui JK. 1981. A survey on image segmentation. Pattern Recognit 13:3–16.
Gat N. 2000. Imaging spectroscopy using tunable filters: A review. SPIE Proc 4056:50–64.
Geusebroek J-M, Cornelissen F, Smeulders AWM, Geerts H. 2000. Robust autofocusing in microscopy. Cytometry 39:1–9.
Gibbs WW. 2001. Cybernetic cells. Sci Am 265:53.
Giuliano KA, Taylor DL. 1998. Fluorescent-protein biosensors: New tools for drug discovery. Trends Biotechnol 16:135–140.
Giuliano KA, DeBiasio RL, Dunlay RT, Gough A, Volosky JM, Zock J, Pavlakis GN, Taylor DL. 1997. High-content screening: A new approach to easing key bottlenecks in the drug discovery process. J Biomol Screen 2:249–258.
Groen FCA, Young IT, Ligthart G. 1985. A comparison of different focus functions for use in autofocus algorithms. Cytometry 6:81–91.
Hahn KM, Toutchkine A. 2002. Live-cell fluorescent biosensors for activated signaling proteins. Curr Opinion Cell Biol 14:167–172.
Hahn KM, DeBasio R, Taylor DL. 1992. Patterns of elevated free calcium and calmodulin activation in living cells. Nature 359:736–738.
Haj FG, Verveer PJ, Squire A, Neel BG, Bastiaens PI. 2002. Imaging sites of receptor dephosphorylation by PTP1B on the surface of the endoplasmic reticulum. Science 295: 1708–1711.
Heim R, Tsien RY. 1996. Engineering green fluorescent protein for improved brightness, longer wavelengths, and fluorescence resonance energy transfer. Curr Biol 6:178–182.
Heynen S. 2002. Fluorescent feature fidelity and exploration of automated cell-by-cell classification in model populations for high-throughput image cytometry. Dissertation, University of California, San Diego.
Heynen S, Gough DA, Price JH. 1997. Optically stabilized mercury vapor short arc lamp as UV-light source for microscopy. SPIE Proc Optical Diagnostics of Biological Fluids and Advanced Techniques Analytical Cytology 2982:430–434.
Hiroaki K. 2002. Systems Biology: A brief overview. Science 295(5560):1662–1664.
Honda A, Adams SR, Sawyer CL, Lev-Ram V, Tsien RY, Dostmann WR. 2000. Spatiotemporal dynamics of guanosine 3′, 5′-cyclic monophosphate revealed by a genetically encoded, fluorescent indicator. PNAS 98: 2437–2442.
Huang K, Lin J, Gajnak JA, Murphy RF. 2002. Image content-based retrieval and automated interpretation of fluorescence microscope images via the protein subcellular location image database. Proceedings of the 2002 IEEE International Symposium on Biomedical Imaging (ISBI-2002). p 325–328.
Hudson DL, Cohen ME. 2000. Neural networks and artificial intelligence for biomedical engineering. New York: Institute of Electrical and Electronics Engineers.

Jaehne B. 1997. Digital image processing: Concepts, algorithms, and scientific applications. Berlin, New York: Springer.

Jarvik JW, Adler SA, Telmer CA, Subramaniam V, Lopez AJ. 1996. CD-Tagging: A new approach to gene and protein discovery and analysis. Biotechniques 20:896–904.

Jarvik JW, Fisher GW, Shi C, Hennen L, Hauser C, Adler S, Berget PB. 2002. In vivo functional proteomics: Mammalian genome annotation using CD-tagging. BioTechniques 33:852–867.

Katsumi A, Milanini J, Kiosses WB, del Pozo MA, Kaunas R, Chien S, Hahn KM, Schwartz MA. 2002. Effects of cell tension on the small GTPase Rac. J Cell Biol 158:153–164.

Kemp RA, MacAulay C, Garner D, Palcic B. 1997. Detection of malignancy associated changes in cervical cell nuclei using feed-forward neural networks. Anal Cell Pathol 14:31–40.

Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP. 1998. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 4:844–847.

Krajewska M, Zapata JM, Meinhold-Heerlein I, Hedayat H, Monks A, Bettendorf H, Shabaik A, Bubendorf L, Kallioniemi O-P, Kim H, Reifenberger G, Reed JC, Krajewski S. 2002. Expression of Bcl-2 family member Bid in normal and malignant tissues. Neoplasia 4:129–140.

Krajewski S, Krajewska M, Shabaik A, Miyashita T, Wang HG, Reed JC. 1994. Immunohistochemical determination of in vivo distribution of Bax, a dominant inhibitor of Bcl-2. Am J Pathol 145:1323–1336.

Krajewski S, Krajewska M, Ellerby LM, Welsh K, Xie Z, Deveraux QL, Salvesen GS, Bredesen DE, Rosenthal RE, Fiskum G, Reed JC. 1999. Release of caspase-9 from mitochondria during neuronal apoptosis and cerebral ischemia. Proc Natl Acad Sci USA 96:5752–5757.

Kraynov VS, Chamberlain C, Bokoch GM, Schwartz MA, Slabaugh S, Hahn KM. 2001. Localized Rac activation dynamics visualized in living cells. Science 290:333–337.

Levenson RM, Hoyt CC. 2000. Spectral imaging and microscopy. Am Laboratory 32:26–33.

Llopis J, Westin S, Ricote M, Wang Z, Cho CY, Kurokawa R, Mullen TM, Rose DW, Rosenfeld MG, Tsien RY, Glass CK, Wang J. 1999. Ligand-dependent interactions of coactivators steroid receptor coactivator-1 and peroxisome proliferator-activated receptor binding protein with nuclear hormone receptors can be imaged in live cells and are required for transcription. PNAS 97:4363–4368.

Mackin RW, Jr. Newton LM, Turner JN, Holmes TJ, Roysam B. 1998. Accuracy of nuclear classification in cervical smear images. Anal Quant Cytol Histol 20: 77–91.

Macville MVE, Speel EJM, van der Laak JAWM, de Wilde PCM, Hanselaar AGJM, Soenksen D, McNamara G, Hopman AHN, Ried T. 2001. Spectral imaging of multicolor chromogenic dyes in pathological specimens. Anal Cell Pathol 22:133–142.

Manley S, Mucci NR, De Marzo AM, Rubin MA. 2001. Relational database structure to manage high-density tissue microarray data and images for pathology studies focusing on clinical outcome. Am J Path 159:837–843.

Mercié P, Belloc F, Bihlou-Nabera C, Barthe C, Pruvost A, Renard M, Seigneur M, Bernard P, Marit G, Boisseau MR. 2000. Comparative methodologic study of NFκB activation in cultured endothelial cells. J Lab Clin Med 136:402–411.

Misteli T, Spector DL. 1997. Application of the green fluorescent protein in cell biology and biotechnology. Nat Biotechnol 15:961–964.

Miyawaki A, Llopis J, Heim R, McCaffery JM, Adams JA, Ikura M, Tsien RY. 1997. Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. Nature 388:882–887.

Murphy RF, Boland MV, Velliste M. 2000. Towards a systematics for protein subcellular location: Quantitative description of protein localization patterns and automated analysis of fluorescence microscope images. Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology 8:251–259.

Murphy RF, Velliste M, Porreca G. 2002. Robust classification of subcellular location patterns in fluorescence microscope images. Proceedings of the 2002 IEEE International Workshop on Neural Networks for Signal Processing (NNSP 12). p 67–76.

Nalbant P, Kraynov V, Touchkine A, Hahn KM. 2001. Fluorescent biosensors report activation of endogenous unlabelled Cdc42 in living cells. ASCB conference (abstract).

Nguyen LK, Bravo-Zanoguera ME, Kellner AL, Price JH. 2000. Magnification corrected optical image splitting module for simultaneous multiplanar acquisition. SPIE Proc Optical Diagnostics of Living Cells III 3921:31–40.

Oliva M, Bravo-Zanoguera ME, Price JH. 1999. Filtering out contrast reversals for microscopy autofocus. Applied Optics, Optical Tech & Biomed Optics 38(4):638–646.

Pal N, Pal S. 1993. A review of image segmentation techniques. Pattern Recognit 26(9):1277–1294.

Price JH. 1990. Scanning cytometry for cell monolayers. Dissertation, University of California, San Diego.

Price JH, Gough DA. 1994. Comparison of digital autofocus functions for phase-contrast and fluorescence scanning microscopy. Cytometry 16(4):283–297.

Price JH, Hunter EA, Gough DA. 1996. Accuracy of least squares designed spatial FIR filters for segmentation of images of fluorescence stained cell nuclei. Cytometry 25(4):303–316.

Roques EJS, Murphy RF. 2002. Objective evaluation of differences in protein subcellular distribution. Traffic 3: 61–65.

Ruifrok AC, Johnston DA. 2001. Quantification of histochemical staining by color deconvolution. Anal Quant Cytol Histol 23:291–299.

Saeed N. 1998. Magnetic resonance image segmentation using pattern recognition, and applied to image registration and quantitation. NMR Biomed 11:157–167.

Skacel MS, Skilton B, Pettay JD, Tubbs RR. 2002. Tissue microarrays: A powerful tool for high-throughput analysis of clinical specimens. Appl Immunohistochem Molecul Morphol 10:1–6.

Srivastava S, Gopal-Srivastava R. 2002. Biomarkers in cancer screening: A public health perspective. J Nutr 132 (Suppl):2471S–2475S.

Taylor DL, Wang YL. 1980. Fluorescently labeled molecules as probes of the structure and function of living cells. Nature 284:405–410.

Telmer CA, Berget PB, Ballou B, Murphy RF, Jarvik JW. 2002. Epitope tagging genomic DNA using a CD-tagging Tn10 minitransposon. Biotechniques 32:422–430.

Ting AY, Kain KH, Klemke RL, Tsien RY. 2001. Genetically encoded fluorescent reporters of protein tyrosine kinase activities in living cells. PNAS 98:15003–15008.

Touchkine A, Kraynov V, Hahn KM. 2003. Solvatochromic dye to report protein conformational changes in living cells. J Am Chem Soc (in press).

Van der Laak JAWM, Siebers AG, Cuijpers VMJI, Pahlplatz MMM, de Wilde PCM, Hanselaar AGJM. 2002. Automated identification of diploid reference cells in cervical smears using image analysis. Cytometry 47: 256–264.

Velliste M, Murphy RF. 2002. Automated determination of protein subcellular locations from 3D fluorescence microscope images. Proceedings of the 2002 IEEE International Symposium on Biomedical Imaging (ISBI-2002) p 867–870.

Verma IM, Stevenson JK, Schwarz EM, Van Antwerp D, Miyamoto S. 1995. Rel/NF-κB/IκB family: Intimate tales of association and dissociation. Genes Dev 9:2723–2735.

Weyn B, Jacob W, da Silva VD, Montironi R, Hamilton W, Thompson D, Bartels HG, van Daele A, Dillon K, Bartels PH. 2000. Data representation and reduction for chromatin texture in nuclei from premalignant prostatic, esophageal, and colonic lesions. Cytometry 41:133–138.

Wu W, Hu W, Kavanagh JJ. 2002. Proteomics in cancer research. Int J Gynecol Cancer 12:409–423.

Yamanishi DY, Junquan X, Hujsak PG, Yang Z, Wang X, Wu L. 2002. Enrichment of rare fetal cells from maternal peripheral blood. Expert Rev Mol Diagn 2:303–311.

Yarmush ML, Jayaraman A. 2002. Advances in proteomic technologies. Annu Rev Biomed Eng 4:349–373.

Zimmerman RL, Fogt F, Goonewardene S. 2000. Diagnostic value of a second generation CA15-3 antibody to detect adenocarcinoma in body cavity effusions. Cancer Cytopathology 90:230–234.