

Signal to Noise Analysis of Multiple Color Fluorescence Imaging Microscopy

Y. Garini,* A. Gil, I. Bar-Am, D. Cabib, and N. Katzir

Applied Spectral Imaging, Migdal HaEmek, Israel

Received 2 March 1998; Revision Received 16 October 1998; Accepted 17 October 1998

Background: Various approaches that were recently developed demonstrate the ability to simultaneously detect all human (or other species) chromosomes by using combinatorial labeling and fluorescence in situ hybridization (FISH). With the growing interest in this field, it is important to develop tools for optimizing and estimating the accuracy of different experimental methods.

Methods: We have analyzed the principles of multiple color fluorescence imaging microscopy. First, formalism based on the physical principles of fluorescence microscopy and noise analysis is introduced. Next, a signal to noise (S/N) analysis is performed and summarized in a simple accuracy criterion. The analysis assumes shot noise to be the dominant source of noise.

Results: The accuracy criterion was used to calculate the S/N of multicolor FISH (M-FISH), spectral karyotyping,

ratio imaging, and a method based on using a set of broad band filters. Spectral karyotyping is tested on various types of samples and shows accurate classifications. We have also tested classification accuracy as a function of total measurement time.

Conclusions: The accuracy criterion that we have developed can be used for optimizing and analyzing different multiple color fluorescence microscopy methods. The assumption that shot noise is dominant in these measurements is supported by our measurements. *Cytometry* 35:214–226, 1999. © 1999 Wiley-Liss, Inc.

Key terms: spectral karyotyping; multicolor FISH; fluorescence microscopy; spectral imaging; Fourier spectroscopy; classification algorithms; spectral analysis

The desire to simultaneously detect many biological probes has been the motivation behind developments in multicolor fluorescence microscopy in the last few years. The advance was made possible by technological improvements that include the optical performance of microscopes and filters, charged coupled device (CCD) sensitivity, fluorochrome brightness, and the efficiency and specificity of biological labeling methods.

When multicolor methods are combined with fluorescence in situ hybridization (FISH), more than a single DNA target probe may be detected. This considerably broadens the spectrum of applications of FISH methods.

The potential of multicolor FISH was first demonstrated in 1989 (1) when three different fluorochromes were used to label three different centromere probes. The same group reported in 1990 that the number of different probes could be larger than the number of fluorochromes by using combinatorial labeling (2) or ratio labeling (3). This method was extended later to seven different probes (4) and even more (5).

Recently, two newly developed multicolor methods, M-FISH (6) and spectral karyotyping (7), presented the simultaneous detection of all human (or mouse) chromosomes. This means that many cytogenetic aberrations may now be detected in a relatively straightforward manner.

These new multicolor methods thus have the potential to improve both chromosome research and cytogenetic diagnosis. For example, reports have already been published on finding hidden translocations in hematological malignancies (8) and the identification of a synovial-cell sarcoma as a bone tumor (9).

The growing importance of multicolor fluorescence techniques increases the need for a substantial understanding of the advantages and limitations of the experimental methods. This paper presents a detailed study of the principles of multicolor fluorescence microscopy. A mathematical formalism is developed and summarized in a single criterion, i.e., the accuracy criterion of multicolor fluorescence. These tools may be used for optimizing and analyzing the accuracy of the results of different experimental methods.

The goal of a multicolor fluorescence microscopy measurement is to map the location of each of many probes in an image. During the measurement, a set of data points is acquired for each pixel of the image; the type of data depends on the specific experimental technique used.

*Correspondence to: Dr. Yuval Garini, Applied Spectral Imaging, POB 101, Migdal HaEmek 10551, Israel.
E-mail: yuval@spectral-imaging.com

These data are then analyzed by using a classification algorithm, and each pixel is assigned to a different probe. Assuming an optimal classification algorithm, the accuracy of the pixel classification depends on the set of measured data points. This set of data points must be unique and well-defined for all the pixels belonging to each probe. There are a large number of pixels in an image (around a million), and a very high percentage of those pixels must be classified correctly. The misclassification of even a few pixels could lead to a wrong cytogenetic conclusion. The main goal of this work was to develop quantitative tools for estimating the degree of separability of the data in a multicolor fluorescence measurement. Classification methods are most important in order to get optimized probe identifications. This deserves careful treatment and is outside the scope of this work.

FLUORESCENCE MULTIPLE COLOR MICROSCOPY Fluorescence Microscopy Measurement

Many different fluorochromes are now available with emission spectra that cover the whole visible light and near infrared spectrum. The emission spectrum of a fluorochrome is shifted toward longer wavelengths, compared to its absorption spectrum. This spectral shift (also called the Stokes shift) makes it possible to separate excitation light from emission light by using optical filters or spectroscopy (10,11).

Fluorescent microscopy measurements (12,13) are mostly based on epifluorescence. Because the emission intensity is usually a few orders of magnitude lower than the excitation intensity, it is necessary to block the excitation light from the emission path. This is done by using a set of filters (usually three) in the light path of the microscope.

In multicolor measurements, several fluorochromes are used simultaneously. In order to distinguish between the emission light from these fluorochromes, they should have the least degree of spectral overlap. This requirement is in addition to many other requirements for the fluorochromes, such as high brightness and stability.

The spectral response of a CCD is usually high enough in a limited spectral range (typically, a spectral range of about 400–500 nm). The optics of the microscope are typically also designed for high-quality images in a similar, or even smaller, spectral band. In order to get bright enough signals, the emission spectra of the chosen fluorochromes should fall inside those ranges. On the other hand, a typical full width at half maximum (FWHM) of a fluorochrome spectrum (both absorption and emission) is in the range of 50–100 nm, and the Stokes shift is also of the same order of magnitude. This fact results in a high degree of spectral overlap, as shown in Figure 1. It is this overlap that complicates the measurement of several fluorochromes simultaneously.

Labeling Methods

Combinatorial labeling was developed (14) to detect more biological probes than the number of available fluorochromes. In combinatorial labeling each probe is

labeled with a unique combination of a few fluorochromes. With N fluorochromes, $2^N - 1$ different probes can be labeled.

The methods of labeling the different probes with fluorochromes are of great importance and complexity, but fall outside the scope of this work. Detailed discussions of the labeling techniques can be found elsewhere (15).

The fluorochrome labeling method needs to meet several conditions before multicolor measurements will succeed. First, the concentration of fluorochromes on the sample must be high enough to give off a bright signal; the fluorochromes must be stable to prevent photobleaching; and the labeling must be specific. Secondly, for every probe, the concentration ratio of the fluorochromes attached to it should be well-defined across the image.

Another method described in the literature with respect to combinatorial labeling is ratio labeling (3,5). In ratio labeling, not only are different fluorochrome combinations used to label a probe, but also different concentrations of the different fluorochromes. If N fluorochromes are used, each one in K different concentration levels, and if the highest concentration for each fluorochrome is defined as 1 and the other concentrations are defined as $\frac{1}{2}^1, \frac{1}{2}^2, \dots, \frac{1}{2}^{K-2}$ and 0, then, the total number of valid combinations is

$$K^N - (K - 1)^N. \quad (1)$$

Valid combinations include combinations that have different fluorochrome concentration ratios. For example, if two fluorochromes are used, each one in three different levels, 1, 0.5, and 0, then, the valid combinations are (1,1), (1,½), (½,1), (1,0), (0,1). The combination (½,½) is not valid because it has the same concentration ratio as (1,1). If the different intensities used for each fluorochrome do not follow the above rule, a different number of valid combinations will result. Notice that the value of combinations in combinatorial labeling is just a special case ($K = 2$) of ratio labeling.

Multiple Color Measurement Methods

M-FISH. M-FISH (6,16) is performed by measuring five images, each one through a different filter set together with combinatorial labeling. Each filter set is designed to be highly specific (>90%) to only one of the fluorochromes. This specificity is achieved by using narrow band filters, which may not coincide with the peaks of the absorption and emission fluorochrome spectra (6).

Five images are digitized, each one through one of the filter sets. A DAPI image is digitized as well and used to define a mask for the probe images. The data are analyzed by taking each of the digitized images (that contain information about a single fluorochrome) and setting an intensity threshold that distinguishes pixels which are labeled from those which are not labeled with that fluorochrome. By combining the results from the five

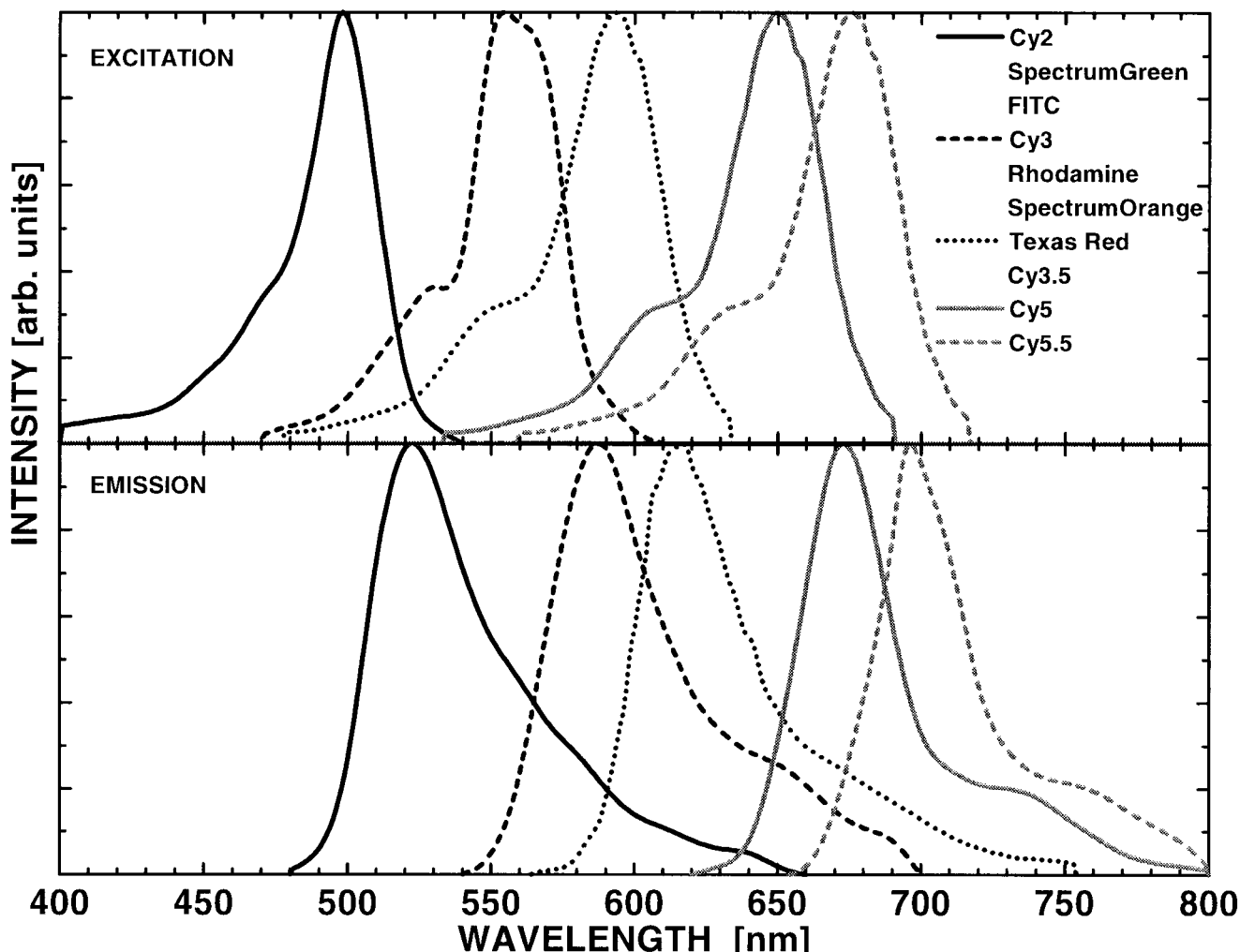


FIG. 1. The absorption and emission spectra of a family of five different fluorochromes. There is a high degree of overlap between the absorption and emission spectra of the different fluorochromes. Each spectrum shown in the graph represents few different fluorochromes (i.e., Rhodamine, SpectrumOrange, and Cy3 all have a similar spectrum).

different images and comparing them to the known labeling scheme, each pixel can be classified.

Ratio labeling. Ratio labeling (3,5) has the advantage that fewer fluorochromes can be used in more combinations. On the other hand, each fluorochrome appears in several intensity levels. The measurement is performed by using a set of filters that matches the fluorochromes. The data analysis yields the intensity ratio of the different fluorochromes for each pixel, which is then compared to the known ratio-labeling scheme.

Spectral karyotyping. Spectral karyotyping (7,17) is based on simultaneously measuring the full fluorescence spectrum of each pixel of the image. Spectral karyotyping uses Fourier spectroscopy to measure the spectrum. The fluorescence spectrum is measured through a triple dichroic filter that allows for the simultaneous excitation and detection of all the fluorochromes used. During the analysis, the spectrum at each pixel is compared to all fluorochromes' spectra, and the fluorochrome combina-

tion is determined. This combination is compared to the known labeling scheme of the probes, and the probe type of each pixel is ascertained (17).

Quantitative Analysis

Based on the Beer-Lambert law, the fluorescence intensity of a fluorochrome is linearly dependent on the following parameters:

1. The extinction coefficient $\varepsilon(\lambda)$, which describes the fluorochrome probability to absorb light quanta at wavelength λ . $\varepsilon(\lambda)$ will be described as $\varepsilon \cdot F_{\text{ABS}}(\lambda)$ where ε is the commonly described extinction coefficient and $F_{\text{ABS}}(\lambda)$ is the fluorochrome absorption spectrum, whose peak value is normalized to 1.
2. The quantum yield Φ of a fluorochrome, which defines the probability of an excited fluorescent molecule to emit a photon. The probability to emit a photon at a certain wavelength λ is obtained by multiplying the

quantum yield Φ , by the emission spectrum $F_{EM}(\lambda)$, which is normalized (its integration over wavelength equals 1).

3. The excitation source flux $L_0 \cdot L(\lambda)$ measured on the sample itself in units of photons per unit wavelength per unit time [$\text{sec}^{-1} \text{nm}^{-1}$]. L_0 is the total intensity, which when multiplied by $L(\lambda)$ gives the flux at a certain wavelength.

4. Fluorescence collection response of the microscope and the CCD detector, $R(\lambda)$. It is the ratio of the number of electrons generated in a pixel of the CCD relative to the number of photons which are emitted from the volume of the sample imaged by that pixel.

5. The transmission of the filter-cube in the excitation light path, $T_{EX}(\lambda)$ and the emission light path, $T_{EM}(\lambda)$.

6. Optical depth of the measured substance, I [cm], and the area, S , imaged by one pixel of the CCD. The area is defined by the CCD pixel size divided by the magnification of the microscope. The optical depth depends also on the numerical aperture of the microscope.

7. A_D is the digital output that results when a given number of electrons are accumulated in a pixel.

8. Fluorochrome molar concentration, ρ .

9. The exposure time, t , for one CCD frame.

An index will be used to distinguish different fluorochromes. For example, $F_{EM,j}(\lambda)$ is the emission spectrum of fluorochrome j . In the same manner, an index is used to distinguish different filters.

In the following equation, I_{ij} is the intensity that is due to fluorochrome j , in a CCD camera pixel, measured using filter i :

$$I_{ij} = t \cdot I \cdot S \cdot \rho_j \cdot \epsilon_j \cdot \Phi_j \cdot L_0 \cdot A_D \cdot \int_{\lambda} R(\lambda) \cdot F_{EM,j}(\lambda) \cdot T_{EM,i}(\lambda) d\lambda \times \int_{\lambda'} L(\lambda') \cdot F_{ABS,j}(\lambda') \cdot T_{EX,i}(\lambda') d\lambda'. \quad (2)$$

Here, all the values except the first four parameters are constants of the experimental setup. The first four parameters depend on the fluorochrome concentration, thickness, and size of every point in the image and the exposure time of the CCD.

In the case where the emission spectrum is measured at many wavelengths (e.g., spectral karyotyping), the spectral integration over λ' is taken for every wavelength due to the spectral resolution (usually 5–15 nm).

When N different fluorochromes are contributing to the fluorescence, then, I_i is the recorded intensity of a pixel due to the contribution of all the fluorochromes:

$$I_i = \sum_{j=1}^N I_{ij}. \quad (3)$$

Spectral and Vector Presentation

The set of values I_i measured for each pixel can be described as a spectrum. In the case of spectral karyotyping, it is exactly the emission spectrum. In the case of measurement through a set of filters, it can still be considered analogous to a spectrum.

A spectrum that contains M measured points can be represented by the set $I_1, I_2, \dots, I_m, \dots, I_M$. This spectrum can also be observed as a vector \mathbf{I} in an M dimensional space,

$$\mathbf{I} = (I_1, I_2, \dots, I_m, \dots, I_M). \quad (4)$$

M describes the number of wavelengths which are measured (in spectral karyotyping) or the number of filters which are being used for the measurement. An example of a spectrum and vector representation, for the case of $M = 2$, is described in Figure 2.

The notation \mathbf{I}_j will be used to refer to the vector which is measured by the set of M filters for the j th fluorochrome.

The fluorescence intensity from different pixels along a chromosome is not uniform, mainly for two reasons. First, because of the cylinder-like shape of a chromosome, the fluorescence intensity decreases from the chromatid center outwards (Fig. 3). Assuming that the fluorochrome density ρ_j is uniform along the chromosome, I in Equation 2 varies, and therefore I_i will vary (for all i -s). Nevertheless, the ratios of the I_i components for each probe in each pixel are most likely preserved because they all proportional to the same thickness I .

The second reason is due to uneven hybridization along the chromosomes. This can be the result of either the dominance of repetitive sequences (e.g., in centromere regions), or the deficiency of some unique DNA sequences in some areas along the chromosome. This will result in a variation of fluorochrome concentration along the chromosomes, but here too, the fluorochrome concentration ratio and therefore also the ratios of the I_i components will be preserved.

Any algorithm developed for analyzing multicolor fluorescence data should take such intensity variations into account.

Fluorochrome Specificity

In order to achieve a high degree of accuracy in the classification, the spectra which are measured for the different fluorochromes should be as different as possible. This difference, or specificity, can be analyzed by calculating the angle between vectors that represent the data measured for different fluorochromes.

Let θ_{kl} be the angle between two vectors \mathbf{I}_k and \mathbf{I}_l in M -dimensional space that represent the spectra measured for two different probes. θ_{kl} can be calculated by:

$$\theta_{kl} = \arccos \left(\frac{\sum_{j=1}^M I_{jk} \cdot I_{jl}}{\sqrt{\sum_{j=1}^M I_{jk}^2} \cdot \sqrt{\sum_{j=1}^M I_{jl}^2}} \right). \quad (5)$$

Because each of the vector components is positive, θ_{kl} is always smaller than $\pi/2$.

$$I=(a, b)$$

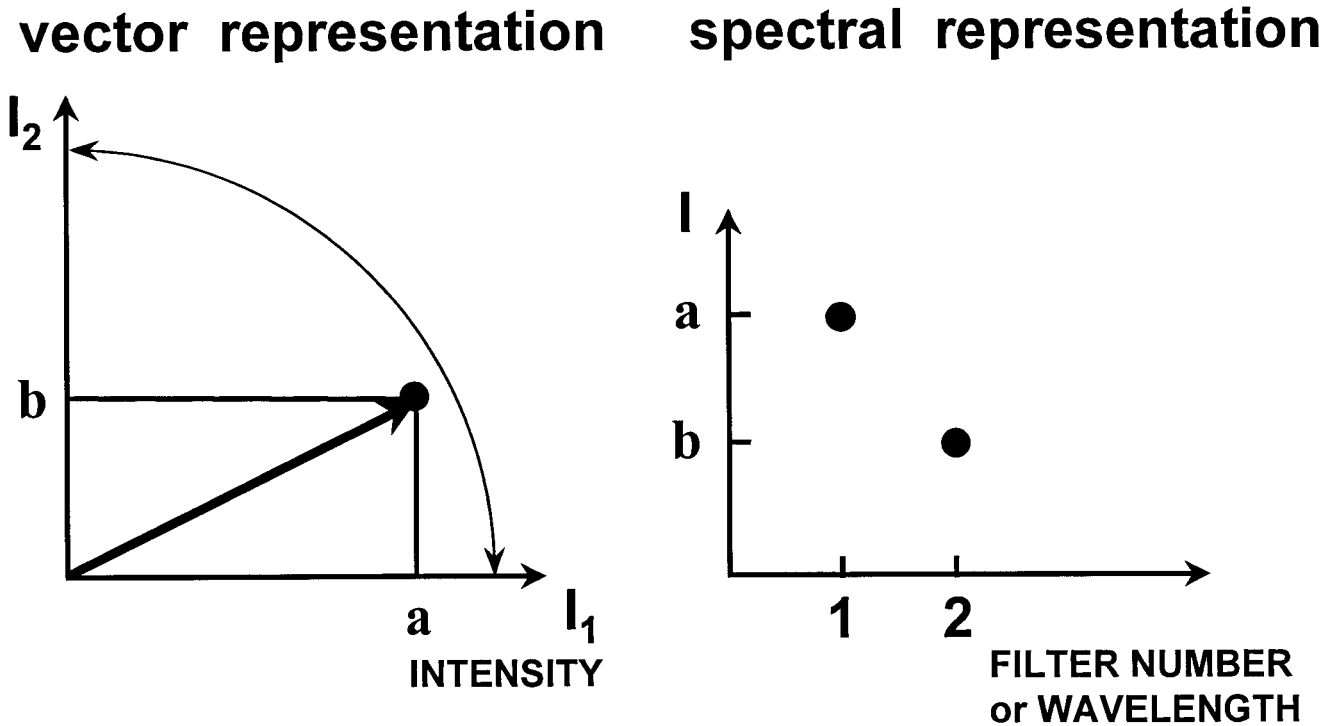


FIG. 2. Spectral and vector representation in the case of a measured data set that includes two measurement points ($M = 2$). The values measured in this case are $I_1 = a$ and $I_2 = b$.

Noise Analysis

Every measurement is accompanied by noise. Usually, there is more than one source of noise.

If I describes a single measured value, δI will be used to describe its noise. In the analysis, one or more functions that depend on the measured values are calculated, $F = F(I_1, I_2, \dots, I_M)$. F itself will have an error, δF , which results from the noise contaminating each of the parameters. δF can be calculated to first order by:

$$\delta F = \frac{dF}{dI} \cdot \delta I \approx \sqrt{\sum_{i=1}^M \left(\frac{\partial F}{\partial I_i} \cdot \delta I_i \right)^2} \quad (6)$$

This approximation of the first order can be used in those cases where δI is small enough relative to I . In general, when the noise δI is on the order of $I/10$ or less, the second term in the development of δF will be proportional to $(\delta I)^2$ which is 10 times smaller than the first order. In this noise range, Equation 6 is therefore a good approximation.

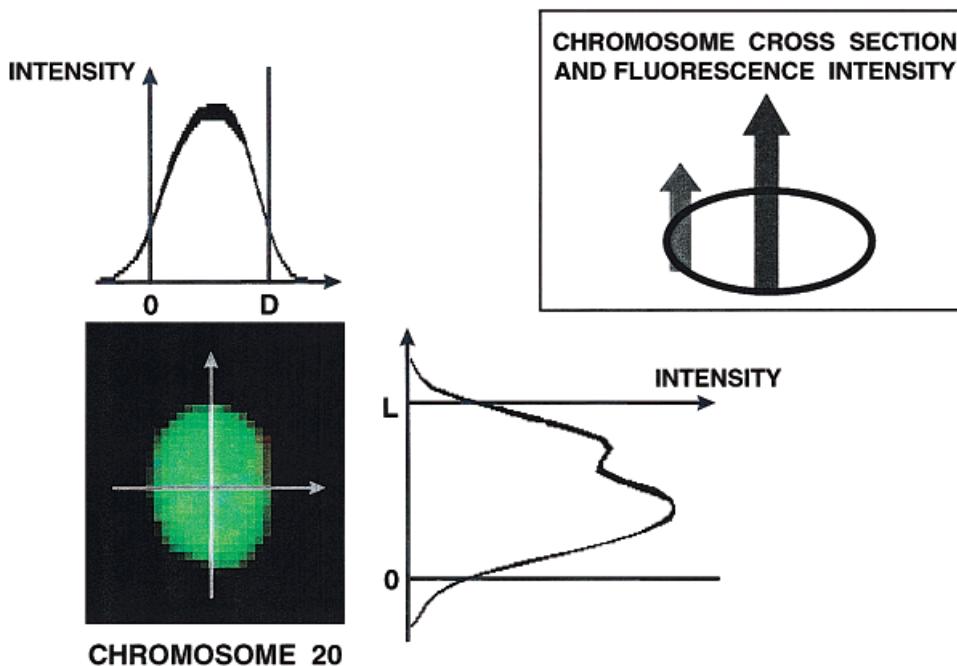
The following different types of noise sources should be taken into consideration in digital imaging fluorescence measurements:

- Bio-chemical noise:** This noise is the result of variations in the fluorochrome concentration ρ_j , variations in the fluorescence intensity due to the micro-environment on the sample, uneven hybridization etc.
- Photon noise:** Intrinsically related to the statistical nature of light. If the measured light intensity is I the root mean square (RMS) shot noise amplitude is equal to \sqrt{I} .
- Detector noise:** Limiting ourselves to CCDs, there are several noise sources. The dominant ones in a high quality CCD are the dark noise N_d and readout noise N_r . As a result of those two, the detector noise is $N_D = \sqrt{N_d^2 + N_r^2}$.

THE ACCURACY CRITERION OF MULTIPLE COLOR FLUORESCENCE

As explained above, in a multicolor measurement, each probe is labeled with a different combination of fluorochromes. During the measurement, a certain number of values I_i , $i = 1 \dots M$, is measured for each pixel of the image. Then, a mathematical analysis, or classification, is performed in order to find the values of ρ_j for each fluorochrome j and to deduce what fluorochrome combi-

FIG. 3. The intensity variation as measured along the two main axes of a chromosome is demonstrated. This intensity variation is a result of the varying thickness of the cylinder-like shape of the chromosome (chromosome 20 in this case) and the uneven hybridization along the chromosome (especially in the centromere region). Other optical effects such as diffraction and depth of focus might contribute as well. The curves show the intensity as a function of the longitudinal and transversal cross section as measured for chromosome 20.



nation is present. The accuracy of the classification is dependent on both the specificity and noise level of the measured vectors.

We have chosen the angle θ_{kl} between the vectors measured for two fluorochromes k and l (or two combinations of fluorochromes) as a parameter for analyzing the signal to noise ratio in a multicolor measurement. The angle θ_{kl} is chosen because it is directly related to the relative fluorescence intensity measured for all the fluorochromes, and at the same time neutralizes the brightness variation of pixels which is expected to be large, as explained before.

The angle θ_{kl} can be calculated from Equation 5. It depends only on the I_{ij} values given in Equations 2 and 3. Those can be either calculated from the equations or directly measured in an experiment. The error $\delta\theta_{kl}$ can be calculated from Equation 5 for θ_{kl} , based on the errors in I_{ij} . If the errors δI are in the order of $1/10$ of I , we can use Equation 6 for the calculation:

$$\delta\theta_{kl} \approx \sqrt{\left(\frac{\partial\theta_{kl}}{\partial I_{1k}} \cdot \delta I_{1k}\right)^2 + \dots + \left(\frac{\partial\theta_{kl}}{\partial I_{Mk}} \cdot \delta I_{Mk}\right)^2 + \left(\frac{\partial\theta_{kl}}{\partial I_{1l}} \cdot \delta I_{1l}\right)^2 + \dots + \left(\frac{\partial\theta_{kl}}{\partial I_{Ml}} \cdot \delta I_{Ml}\right)^2}. \quad (7)$$

In other cases where δI is not so small, $\delta\theta_{kl}$ can be calculated directly from Equation 5 by setting $\delta\theta_{kl} = \theta_{kl}(I + \delta I) - \theta_{kl}(I)$.

In the general case, δI will be different for each I_{ij} and Equation 7 can then be calculated numerically. It is, however, sufficient here to refer to those fluorochrome

combinations that have the highest potential to be misclassified, e.g., the combinations that involve the highest number of similar fluorochromes. For these combinations we can assume that the values of I_{ij} are of the same order,¹ and therefore also the δI_{ij} values. Under these assumptions, Equation 7 can be solved. The details of the calculation are not fully described here.² The rather simple solution that results is:

$$\delta\theta_{kl} \approx \sqrt{\frac{2}{M}} \frac{\delta I}{I}. \quad (8)$$

Here I describes an average value of the intensities measured by all the filters for probe l and k (where each one of them is labeled with several fluorochromes). In the same way, δI stands for an average of the error values measured.

In order to get high accuracy in the classification, the value of $\delta\theta$ must be small relative to θ . The ratio between them can lead to a definition of a criterion for classification

¹Such an assumption is in general justified because we are mainly interested in evaluating the errors in the more complex cases, which are for those probes that are labeled with many fluorochromes. As an example, in both spectral karyotyping and M-FISH, more than half of the chromosomes are labeled with three fluorochromes, which means that the measurement through most of the filters will give a relatively large value. We also assume here that the intensities measured for different fluorochromes are not much different from one another, a condition that can be met by tuning the different fluorochromes combinations as explained in the text, see Equation 10.

²The derivative of θ is performed by using: $\partial/\partial x \arccos U(x) = -1/\sqrt{1 - U^2(x)} \cdot \partial/\partial x U(x)$. The internal differentiation of the I values is straight forward, although rather long.

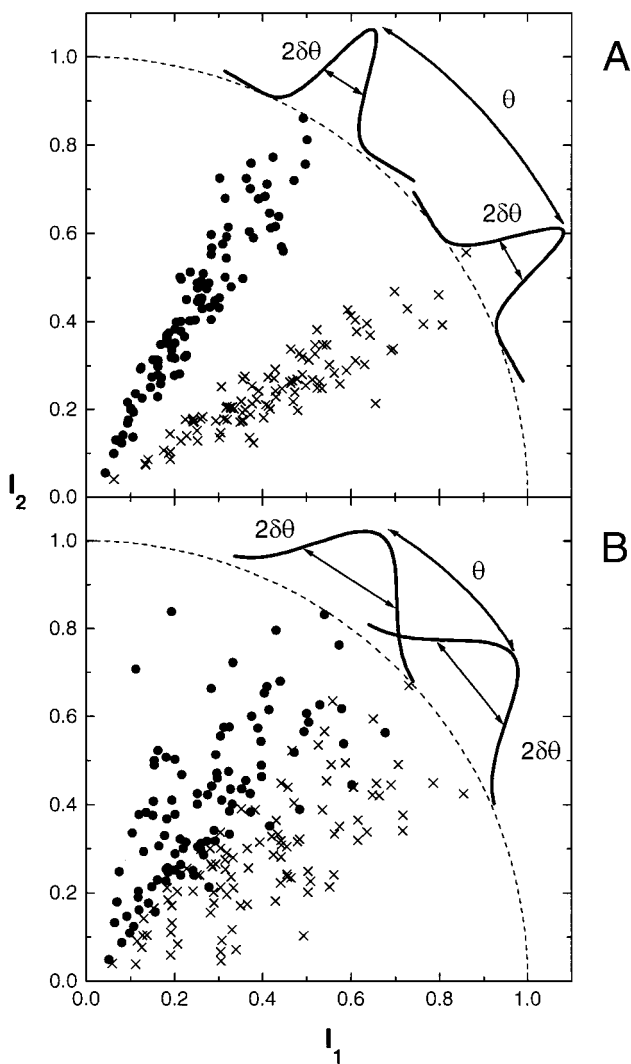


FIG. 4. The relation between θ and $\delta\theta$ in the case of a two-dimensional space. Shown here are the two sets of data points, assumed to follow a Gaussian distribution around two different vectors. A: A case with a high degree of accuracy. Here $\theta \gg \delta\theta$ and therefore, each vector can be classified unambiguously. B: A case where $\theta \approx 4 \times \delta\theta$, and here more than 10% of the points will be misclassified.

accuracy. Figure 4 shows the relation between θ and $\delta\theta$ in the case of $M = 2$. Two sets of vectors that were generated with Gaussian distributions are shown. Figure 4A shows a case with a high degree of accuracy. Here $\theta \gg \delta\theta$ and therefore each vector can be classified unambiguously. Figure 4B shows a case where $\theta \approx 4 \times \delta\theta$, and here, more than 10% of the points will be misclassified.

In the general case of $M > 2$ we should look at all the normalized vectors, i.e., vectors whose length is set to unity (as explained, the intensity or length of the vector can vary, and it is the vector direction which is important). This normalization projects each one of the M dimensional vectors onto a point on the first quadrant of the M dimensional hypersphere surface. Due to noise, all the vectors measured for a single probe form a cluster on this hypersphere surface. With the definitions given above, θ is

a measure of the distance between the average vectors of two different clusters, and $\delta\theta$ is a measure of the variation of θ as a result of the distribution of the points in the clusters.

If the measurement technique takes advantage of K different fluorochrome concentrations, then, the minimal angle that will be measured between certain combinations of two pure fluorochrome vectors is equal to $\theta_{MIN} = \theta_{PURE}/K$, where θ_{PURE} is the angle between the vectors measured for the pure fluorochromes. Combining this with the definition of θ , the following expression can be defined for the S/N ratio of a multicolor measurement:

$$S/N = \frac{\theta_{PURE}}{\delta\theta \cdot K}. \quad (9)$$

We refer to this expression as the accuracy criterion of multiple color fluorescence. It can be calculated by using Equations 5 and 8. This value should be understood only as a criterion; the higher the S/N is, the better the accuracy of the results will be. For a given experimental method (where M , N , and K are fixed), any improvement to the S/N ratio improves the accuracy of the classification. Such improvements can be made, for example, by optimizing the filter parameters or by changing the fluorochromes used in the labeling scheme.

The actual degree of misclassification can be determined from Equation 9 by calculating the percentage of overlap of two clusters of vectors (each one measured for a different probe), when approximated by two M -dimensional spheres with a Gaussian distribution around the center of each (the standard deviation equals $\delta\theta$) and separation between the centers of the two clusters that equal θ . This calculation is outside the scope of this article.

EVALUATING THE S/N OF DIFFERENT METHODS

Using Equation 9, we have calculated the S/N ratios that our model predicts for a few different experimental methods. It is assumed that the highest-quality components will be used, including a CCD camera, optimized microscope optics, illumination light, and filters.

For the calculation, a computer program was developed that takes as input the absorption and emission spectra of all the fluorochromes that are used, the filters' spectral windows (for both the excitation and emission), and all the other parameters that appear in Equation 2. The set of fluorochromes shown in Figure 1 was chosen for the calculation. The calculation was simplified by assuming that:

$$\rho_i \cdot \epsilon_j \cdot \Phi_i = \rho_j \cdot \epsilon_j \cdot \Phi_j \quad \forall i, j. \quad (10)$$

³One of the results that is found by performing the calculation is that the degree of misclassification depends on the value calculated above for the S/N ratio and on a function $f(M)$. This function shows that the dependence of the degree of misclassification on M is even stronger than $1/\sqrt{M}$ that appears in Equation 9.

This can be achieved by refining the fluorochromes' concentrations ρ_j during the labeling, so that Equation 10 holds. This condition simply requires emission intensity to be similar for all the fluorochromes. It is our experience that such a condition is not difficult to achieve, and it does not depend on a specific measurement method.

By defining C as $C = I \cdot S \cdot \rho_j \cdot \epsilon_j \cdot \Phi_j \cdot I_0 \cdot A_D$ (these values are not dependent on the experimental method that is used), Equation 2 can be written in a simpler form:

$$I_{ij} = t \cdot I'_{ij} = t \cdot C \cdot \int_{\lambda} R(\lambda) \cdot F_{EM,j}(\lambda) \cdot T_{EM,i}(\lambda) d\lambda \times \\ \times \int_{\lambda'} L(\lambda') \cdot F_{ABS,j}(\lambda') \cdot T_{EX,j}(\lambda') d\lambda' \quad (11)$$

where I' is defined as the emission intensity per unit time, $I = tI'$.

For the excitation light source intensity, a white spectrum is assumed, which is very close to the spectral characteristics of a xenon lamp. If a different light source is used (e.g., an Hg lamp), the fluorescence intensity of the different fluorochromes might be different, but not the spectral shape of the emission spectra. Therefore, by appropriately adjusting the fluorochrome concentrations, Equation 10 can still hold. This assumption therefore simplifies the calculations without limiting the generality of the results.

A typical response curve of a CCD camera that peaks at 600 nm is used. The shot noise of light is the dominant noise source as long as the intensity measured by the CCD fills a significant part of the CCD's dynamic range (18). The treatment can be extended to low light levels where the readout noise is dominant, but in this case, very poor S/N ratios will result. The biochemical noise described earlier is not treated here, because it is not dependent on the measurement method being used.

In principle, the S/N ratio should be calculated for each pair of fluorochromes, but it is sufficient to make the calculation for the worst case of two fluorochromes with the most similar spectra. For the following calculations in M-FISH and spectral karyotyping, Cy3 and Texas Red were used, since their spectra are the closest (a similar pair of fluorochromes was used in both methods that have been published so far: Cy3 and Cy3.5 in M-FISH (6) and Cy3 and Texas Red in spectral karyotyping (7)).

M-FISH

In M-FISH, five narrow-band-pass filters are used. These filters are highly specific in that each is designed for one of the five fluorochromes. Here, $M = N = 5$ and $K = 2$. For a total measurement time T , each filter will have on average $T/5$ s for measuring an image. The intensity measured will therefore be $I_{ii} = T/5 \times I'_{ii}$. The index in this case is $i = j$ because perfectly specific filters are assumed (if $i \neq j$, then $I_{ij} = 0$). The shot noise is $\delta I_{ii} = \sqrt{I_{ii}} = \sqrt{T/5 \cdot I'_{ii}}$ and therefore, $\delta\theta_{kl} = \sqrt{2/M} \delta I/I = \sqrt{2/5} \sqrt{I'_{ii}/I_{ii}} \cdot T$.

For fully specific filters, $\theta_{PURE} = \pi/2$. Because rather narrow filters are used on the tails of the fluorochrome

spectra (6), the value of I'_{ii} , as calculated from Equation 11, is relatively small, and therefore δI_{ii} is relatively large. The result for $\theta_{PURE}/\delta\theta \cdot K$ is 0.2 in units of $1/\sqrt{C \cdot T}$. It is shown in Figure 5.

Spectral Karyotyping

In spectral karyotyping, a triple dichroic filter is used, and the measurement is performed with a Fourier imaging spectrometer (17). The Fourier spectroscopy method has the advantage (19,20) that throughout the whole measurement, data are collected for all the wavelengths simultaneously. Using Equation 5, it is found that $\theta_{PURE} \approx \pi/4.5$. The throughput of the filter, with respect to an ideal filter, is on the order of 0.5.

With a typical spectral resolution of 10 nm at 600 nm, $M = 30$. When five fluorochromes are used, $N = 5$ and $K = 2$. The calculation of I and δI is performed by starting with the intensity measured in the interferogram and then Fourier-transforming it to the wavelength domain (20). The result for the intensity is $I_{\lambda,j} = 0.5 \cdot I'_{\lambda,j} \cdot T$ where one can see that the whole measurement time is contributing to the intensity measured at each wavelength. The error in θ can be calculated in a similar way (i.e., by calculating the shot noise in the interferogram domain, and then performing the Fourier transformation to find the noise at each wavelength (19)), and it is found that $\delta\theta_{kl} = \sqrt{2/M} \delta I/I \approx 2/30 \times 1/\sqrt{I'_{\lambda,j} \cdot T}$.

By using Equation 11 for $I'_{\lambda,j}$, $\theta_{PURE}/\delta\theta \cdot K$ is calculated as 1.8 in units of $1/\sqrt{C \cdot T}$. This is shown in Figure 5.

The result is relatively high due to two important factors:

1. There are a high number of measured points for each pixel ($M = 30$), which reduces the noise as shown in Equation 8.
2. The Fourier spectroscopy method allows for the simultaneous measurement of all the wavelengths during the whole measurement time. When compared to filter-based methods, this is advantageous because when measuring with five filters, one at a time, only 20% of the total acquisition time is being used for collecting data for each fluorochrome.

Ratio Imaging

The advantage of ratio imaging is that fewer fluorochromes are required to label a given number of chromosomes. There are few options for labeling 24 different probes by ratio labeling (see Equation 1). One option is to use four fluorochromes, each one in three different concentrations ($K = 3$ in Equation 1). Another option is to use three fluorochromes with at least four different intensity levels ($K = 4$). The last option is to use two fluorochromes, each one with 13 different intensity levels.

Because fewer fluorochromes are used, filters which are relatively wide can be used and give both high specificity and throughput. On the other hand, K is larger and therefore the S/N is reduced (see Equation 9).

The calculation itself is very similar to the one performed for M-FISH. The above three options (with $N = 4$,

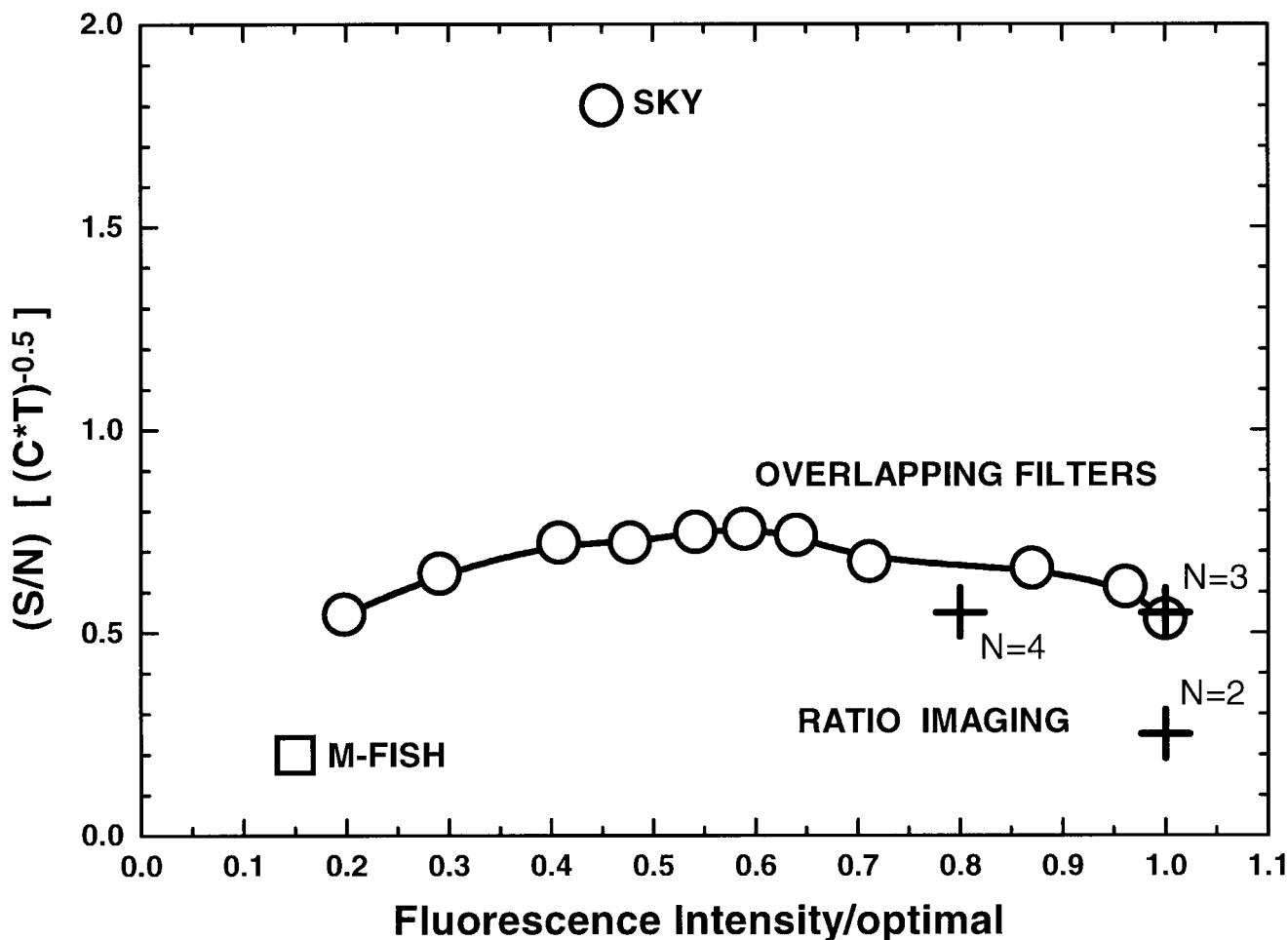


FIG. 5. The calculation for the S/N ratio for different experimental methods, based on the accuracy criterion of multiple color fluorescence. The x-axis represents the excitation efficiency of the fluorochrome relative to the excitation with an optimal filter cube. The results (in the y-axis) are shown in units of $1/\sqrt{C \cdot T}$. The values are normalized so that 1.0 describes the case of $S/N = 12$, as calculated for a total acquisition time of 2 min (see Equation 9). The results for ratio imaging are shown for the case of using four, three, or two fluorochromes. The results for multicolor fluorescence with a set of broad band filters are shown for different sets of five filters each, which are designed to have (when moving from right to left) higher specificity and therefore less throughput.

$K = 3$; $N = 3$, $K = 4$; and $N = 2$, $K = 13$) are calculated and shown in Figure 5.

Multiple Color Fluorescence With Broad Band Filters

In M-FISH, narrow-band filters are used (5–10-nm width (6)), which result in a relatively low throughput and therefore low S/N. In ratio imaging, the number of fluorochromes is smaller but the larger number of intensity levels lowers the S/N ratio. We suggest and analyze here another approach, based on using a set of filters that are not fully fluorochrome-specific. In this case, broad band filters are used and therefore the throughput is higher than in M-FISH. As a result, more than one fluorochrome contributes to the intensity measured by each filter and therefore, the measured vectors (or spectra) must be analyzed by spectral methods that are similar to those developed for spectral karyotyping (17).

Broad band filters are commonly used in FISH, and methods for compensating the spectral overlap of several fluorochromes have been described (21,22). These methods can be used to improve fluorochrome separability for the classification of combinatorial labeled probes.

It is possible to design a filter set whose spectral windows are wide enough to allow high throughput, while keeping each filter relatively specific to one of the fluorochromes. If five fluorochromes are used, at least five different filters are needed.

We have written a computer program to design an optimized set of filters with a given set of constraints. For example, the constraints can be defined by setting limits on the specificity of each filter in the set. Using this program, 11 different filter sets were designed. Each set contains five different filters, and each filter is defined by four wavelengths (two for the excitation and two for the emission). Each set of five filters has different fluoro-

chrome specificity. The first set is designed without any requirement on the specificity at all and each next set is designed with a somewhat higher specificity requirement. Equation 9 was used to compute the S/N of each set of filters, and the results are also shown in Figure 5.

As seen in Figure 5, this method yields a S/N higher than M-FISH. At the same time, even the set of filters that gives the best result has a S/N ratio which is about 2.5 times lower than the S/N calculated for spectral karyotyping.

SPECTRAL KARYOTYPING EXPERIMENTAL RESULTS AND S/N

To test the S/N that can be obtained with the spectral karyotyping method, a quantitative analysis of a normal human male metaphase was made. The sample preparation method is described elsewhere (7). The measurement was performed on a Leica DM-RXA microscope with a 63× immersion oil objective and a specially designed triple dichroic filter cube (Chroma Technology). For illumination, a 75-W Xe lamp was used. The measurement time with the SpectraCube (SD-200) system was 120 s. The analysis was done with the SkyView 1.2 program, which is part of the SpectraCube system. The complete analysis, including the pixel classification as well as the arrangement of chromosomes in the karyotyping table, took less than 5 min. About 3 min were spent to edit the contours of touching and overlapping chromosomes. The contours around the chromosomes are defined automatically, using a threshold-based scheme. This method leaves a generous area around the chromosomes as seen in the DAPI image, which means that many pixels at the edges of the chromosomes are included in the analysis although their intensity is relatively low. To demonstrate this, one of the contours around a chromosome is shown in Figure 6B (the others are not shown in order to leave the bands information clear). A DAPI image was captured as well (5-s measurement time). Although not used for the analysis, it adds important banding information which is complementary to the classification information.

The principles of classification are described elsewhere (7,17), with one major improvement: the analysis is automatic, with no need to define reference spectra to the system. The analysis is done in two steps: 1) Each pixel of each chromosome is classified based on its spectrum. 2) Each chromosome is placed in the karyotyping table according to the class of which the majority of pixels were found to be members.

The measured image in display colors and the DAPI image are shown in Figure 6A,B. The classification results are shown in Figure 6C. The chromosome classification in this case was 100%: the system placed all the chromosomes in their correct position, as was verified by using the DAPI image.

The classification accuracy of each chromosome was checked by counting the misclassified pixels and the total number of pixels. The results are shown in Figure 7. Regions of the chromosomes lying underneath other chromosomes in an overlapping region (chromosomes 2, 9, and 12) and the centromeric regions of the acrocentric

chromosomes were not taken into account (the chromosome probe that was used is not specific for these regions and therefore the labeling of these regions is not well-defined). We also did not take into account a small spot on chromosome 5 that turned out to be a piece of dirt, as found by checking its spectrum (the spectrum of that spot was different from the spectrum of chromosome 5, and the difference was not due to noise). The percentage of misclassified pixels for each chromosome is shown in Figure 7.

The accuracy is found to be very high, with an average misclassification error of 1.3%. This accuracy, of more than 98%, agrees with the calculated S/N for the measurement conditions as found by using Equation 9.

We have also measured a few metaphases with several different total measurement times T , and calculated the classification accuracy for each one. The results are shown in the inset to Figure 7.

Each point describes the average classification accuracy that was measured. The error bars represent the classification accuracy variation of different chromosomes and metaphases. A square root dependence function of the S/N ratio on the total measurement time, as predicted in our calculations (see Equation 9 and Spectral Karyotyping above) was fitted to the measured results. The fitted curve is shown in the insert to Figure 7.

Note that when high accuracy is reached, although the S/N still grows as a function of \sqrt{T} , the classification accuracy asymptotically reaches the maximal value of 100%.

Table 1 summarizes the results for a few more typical cases that were selected. For each sample, 3–10 metaphases were measured with total exposure times which were similar to that of the case described above. The spectral karyotyping results are shown for each case. The misclassification percentages were calculated as explained above. In some cases, the metaphase spreads were more difficult to analyze because of their poor quality (cases 4–6). This is also reflected in the classification accuracy for these cases.

DISCUSSION

The principles of a multicolor fluorescence microscopy measurement have been described in this work. The accuracy of pixel classification depends on the accuracy of the spectrum obtained for every pixel. Tools to analyze the accuracy of the measured data were developed and summarized in the expression for the S/N ratio. This expression that we call the accuracy criterion for multiple color fluorescence can be used to analyze and optimize the parameters of a given experimental method.

By applying this criterion to different experimental methods, one can predict the S/N of each and compare among them. Spectral karyotyping was predicted to yield the highest S/N ratio and pixel classification accuracy. This is due to two basic points: the first is the relatively high number of measured points for each pixel; the second is the relatively high signal to noise ratio achieved by the Fourier spectroscopy method. M-FISH was predicted to have a relatively low S/N, mainly because of the very

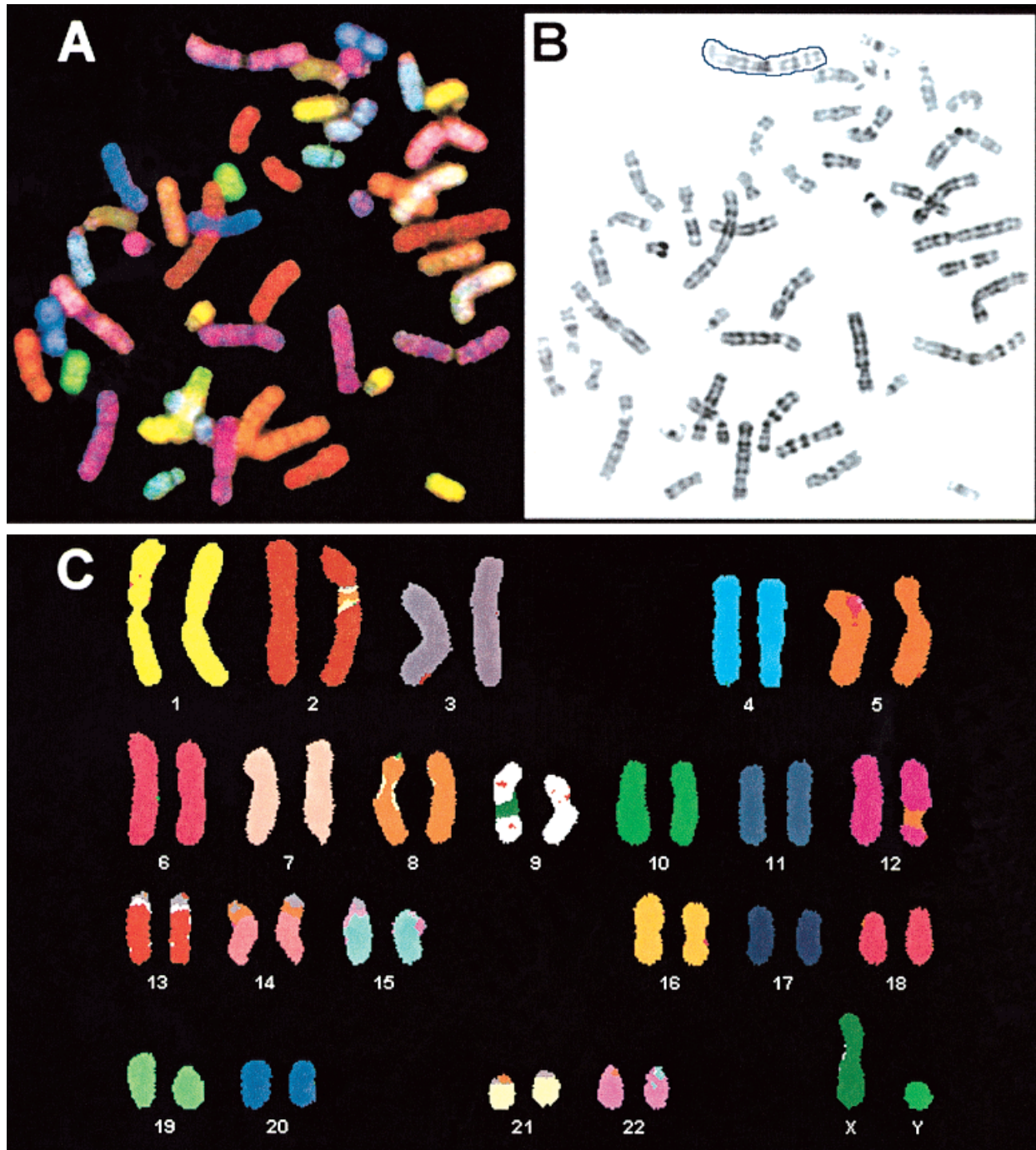


FIG. 6. Example of a normal male metaphase measured with the SpectraCube (SD-200) system. **A:** The metaphase in display colors. Those colors are reproduced from the full spectrum measured at each pixel by using a color assignment algorithm. **B:** The DAPI image as measured for the same cell, after performing a band-enhancement algorithm. One of the chromosomes is shown with the contour defined to it. As shown, the contour is defined in a way that leaves all the pixels that belong to the chromosome inside. **C:** The classification results, as calculated. The results show a high degree of accuracy (more than 98%). The overlapping areas, as well as the centromere regions of the acrocentric chromosomes, are not taken into account in this calculation.

narrow spectral range of the filters that are used. An additional experimental method was suggested in this paper. This multicolor method uses a set of five broad band filters that are less specific to the fluorochromes, but have a higher throughput. This method was shown to have

an intermediate S/N ratio. Ratio labeling was also analyzed, and its S/N was found to be intermediate as well.

The accuracy of the classification of a normal metaphase was tested in a spectral karyotyping measurement, and a high pixel classification accuracy was found.

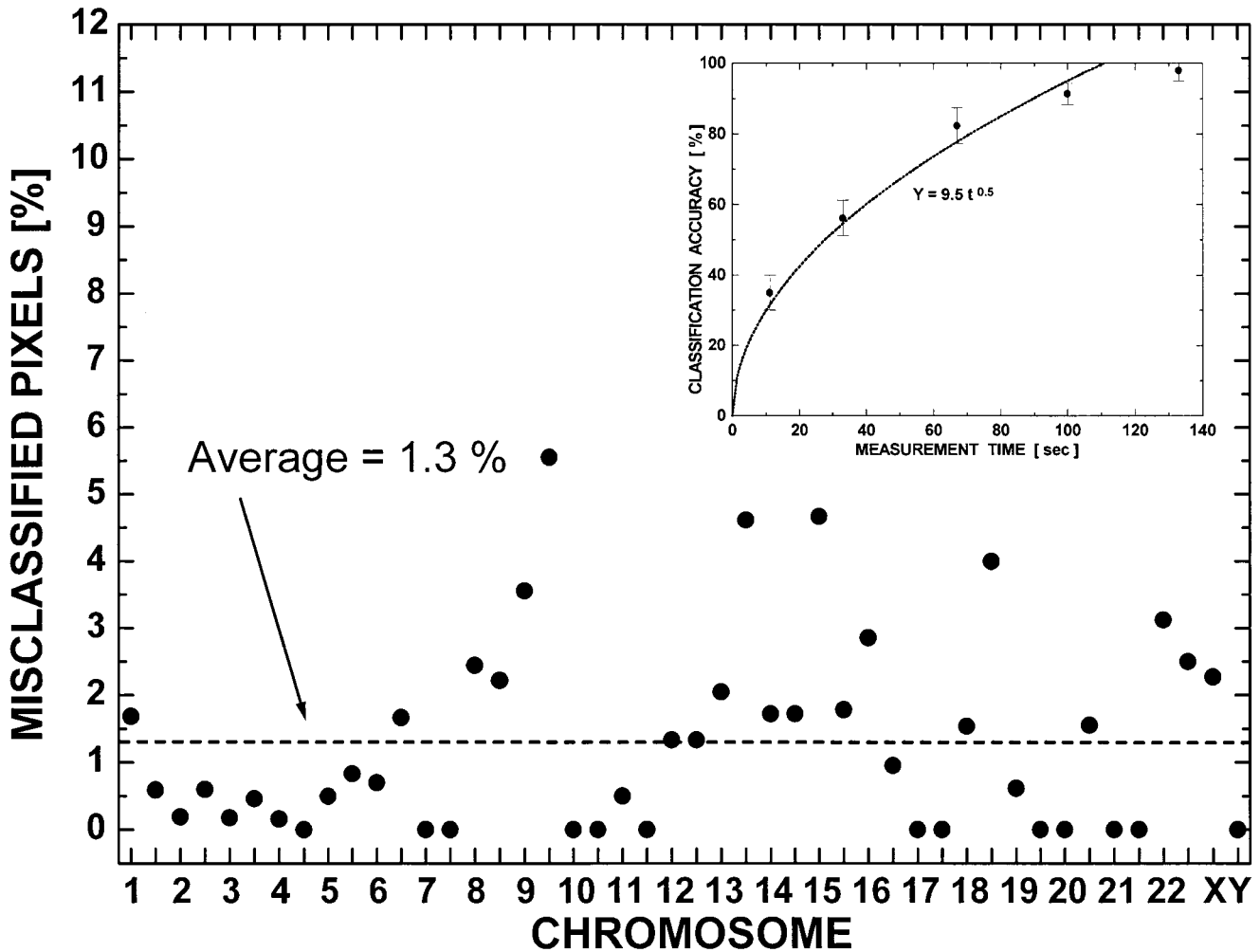


FIG. 7. Accuracy of the spectral karyotyping classification results for each chromosome in the normal human metaphase shown in Figure 6. The results are given in percentage of misclassified pixels (y-axis). The average misclassification percentage is 1.3%. **Inset:** Classification accuracy as a function of total measurement time, as measured for few metaphases. Each point describes the average classification accuracy that was measured. Error bars represent the variation of different chromosomes and metaphases. Dashed line shows a fitted function that is predicted from the calculations to the S/N ratio. Note that when high accuracy is reached, although the S/N still grows as a function of \sqrt{t} , the classification accuracy asymptotically reaches the maximal value of 100%.

Table 1
Summary of Spectral Karyotyping Analysis and Misclassification Results of Six Different Samples*

| Sample | Type | Sex | Spectral karyotyping | Misclassification (% of pixels) |
|--------|-------------------------|-----|---|---------------------------------|
| 1 | Normal (blood) | M | 46,XY | 1.5% |
| 2 | Normal (blood) | M | 46,XY | 2% |
| 3 | Lymphoma (lymph node) | F | 46,XX,t(1;16),del(6),der(9)t(9;18),t(10;15),der(14)t(6;14) | 2% |
| 4 | Normal (amniotic fluid) | F | 46,XX | 1.8% |
| 5 | Lymphoma (lymph node) | M | 56,XXY,+der(2),+der(3),+5X2,+8X2,t(10;22),t(11;16),+der(12)t(12;14),iso(15),+21,+der(X) | 3% |
| 6 | Solid tumor | M | 44,XY,-11,-17,der(9)t(9;11),der(10)t(10;X) | 5% |

*For each sample, 3-10 different metaphases were measured. The misclassification result shows the average calculated for all of them.

Many important aspects of multiple color fluorescence microscopy are still to be further understood. They include a study of the smallest detectable translocation, insertion, deletion, and further progress of labeling methods.

Classification methods are also an important aspect in multicolor methods. The classification methods used here

(17,23) led to very accurate results, but this subject should be treated separately.

ACKNOWLEDGMENTS

We thank Dr. Thomas Ried and Dr. Evelin Schröck from the National Human Genome Research Institute at the NIH

(Bethesda, MD) and Dr. Stan Du Manoir, now at INSERM (Strasbourg, France) for many fruitful discussions. We also thank Prof. Zvi Malik from the Bar Ilan University (Ramat Gan, Israel), Prof. Steve Lipson from the Technion (Haifa, Israel), Prof. Zvi Kam from the Weizmann Institute (Rehovot, Israel), and Dr. Dan Farkas from Carnegie Mellon University (Pittsburgh, PA) for their help during the different development stages of the spectral imaging system. We thank Prof. Elio Cabib of the University of Udine (Udine, Italy), for his suggestions in multidimensional considerations. Finally, we thank all the team at Applied Spectral Imaging for their contributions to the development of all aspects of spectral karyotyping.

LITERATURE CITED

1. Nederlof PM, Robinson D, Abuknesha R, Wiegant J, Hopman AH, Tanke HJ, Raap AK. Three-color fluorescence in situ hybridization for the simultaneous detection of multiple nucleic acid sequences. *Cytometry* 1989;10:20-27.
2. Nederlof PM, van der Flier S, Wiegant J, Raap AK, Tanke HJ, Ploem JS, van der Ploeg M. Multiple fluorescence in situ hybridization. *Cytometry* 1990;11:126-131.
3. Nederlof PM, van der Flier S, Vrolijk J, Tanke HJ, Raap AK. Fluorescence ratio measurements of double-labeled probes for multiple in situ hybridization by digital imaging microscopy. *Cytometry* 1992;13:839-845.
4. Ried T, Baldini A, Rand TC, Ward DC. Simultaneous visualization of seven different DNA probes by in situ hybridization using combinatorial fluorescence and digital imaging microscopy. *Proc Natl Acad Sci USA* 1992;89:1388-1392.
5. Dauwerse JG, Wiegant J, Raap AK, Breuning MH, van Ommen GJ. Multiple colors by fluorescence in situ hybridization using ratiolabelled DNA probes create a molecular karyotype. *Hum Mol Genet* 1992;1:593-598.
6. Speicher MR, Ballard SG, Ward DC. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet* 1996;12:368-375.
7. Schröck E, du Manoir S, Veldman T, Schoell B, Wienberg J, Ferguson-Smith MA, Ning Y, Ledbetter DH, Bar-Am I, Soenksen D, Garini Y, Ried T. Multicolor spectral karyotyping of human chromosomes. *Science* 1996;273:494-497.
8. Veldman T, Vignon C, Schröck E, Rowley JD, Ried T. Hidden chromosome abnormalities in haematological malignancies detected by multicolour spectral karyotyping. *Nat Genet* 1997;15:406-410.
9. Cohen IJ, Issakov J, Avigad S, Stark B, Meller I, Zaizov R, Bar-Am I. Synovial sarcoma of bone delineated by spectral karyotyping. *Lancet* 1997;350:1679-1680.
10. Kasten FH. Introduction to fluorescent probes: properties, history and applications. In: Mason WT, editor. *Fluorescent and luminescent probes for biological activity*. London: Academic Press Ltd.; 1993. p 12-33.
11. Mason WT, editor. *Fluorescent and luminescent probes for biological activity*. London: Academic Press Ltd.; 1993.
12. Ploem JS, Tanke HJ. *Introduction to fluorescence microscopy*. Oxford: Oxford University Press; 1987.
13. Lakowicz JR. *Topics in fluorescence spectroscopy: techniques (volume 1); principles (volume 2); biochemical applications (volume 3)*. New York: Plenum Press; 1991.
14. Nederlof PM, van der Flier S, Verwoerd NP, Vrolijk J, Raap AK, Tanke HJ. Quantification of fluorescence in situ hybridization signals by image cytometry. *Cytometry* 1992;13:846-852.
15. Bouffler SD. Whole-chromosome hybridization. *Int Rev Cytol* 1994;153:197-227.
16. Speicher MR, Ballard SG, Ward DC. Computer image analysis of combinatorial multi-fluor FISH. *Bioimaging* 1996;4:52-64.
17. Garini Y, Macville M, Du Manoir S, Buckwald RA, Lavi M, Katzir N, Wine D, Bar-Am I, Schröck E, Cabib D, Ried T. Spectral karyotyping. *Bioimaging* 1996;4:65-72.
18. Aikens R. Properties of low-light-level slow-scan detectors. In: Mason WT, editor. *Fluorescent and luminescent probes for biological activity*. London: Academic Press Ltd.; 1993. p 277-286.
19. Bell RJ. *Introductory Fourier transform spectroscopy*. London: Academic Press; 1972.
20. Chamberlain JG. *The principles of interferometric spectroscopy*. New York: Wiley; 1979. p 262-310.
21. Castleman KR. Color compensation for digitized FISH images. *Bioimaging* 1993;1:159-165.
22. Castleman KR. Digital image color compensation with unequal integration periods. *Bioimaging* 1994;2:160-162.
23. Garini Y, Katzir N, Cabib D, Buckwald RA, Soenksen D, Malik Z. Spectral bio-imaging. In: Wang XF, Herman B, editors. *Fluorescence imaging spectroscopy and microscopy*. New York: John Wiley & Sons; 1996. p 65-72.