 Cytometry 31:93–99 (1998)

Automated FISH Spot Counting in Interphase Nuclei: Statistical Validation and Data Correction

Carlos Ortiz de Solórzano,[1]* Andrés Santos,[1] Isabel Vallcorba,[2] José-Miguel García-Sagredo,[2] and Francisco del Pozo[1]

[1]Grupo de Bioingenieria y Telemedicina, Universidad Politécnica de Madrid, Madrid, Spain
[2]Servicio de Genética, Hospital Ramón y Cajal de Madrid, Madrid, Spain

Received 23 June 1997; Accepted 31 October 1997

The evaluation of an automated system for Fluorescence In Situ Hybridization (FISH) spot counting in interphase nuclei is presented in this paper. Different types of experiments have been performed with samples from known populations. In all of them the goal is to detect mosaicism of chromosome X in leukocytes from mixtures in known proportions of healthy male and female blood. First the initial results from the automatic FISH analysis system were obtained and evaluated. Then the analysis was modified to reduce systematic errors, so that the results are closer to what an experienced human operator would have obtained (system calibration step). Finally, an additional control probe of chromosome Y was used to detect and discard cells where incorrect hybridization or other abnormal situations had occurred. In each step the system sensitivity was determined by the use of two statistical validation tests, so that the improvement brought about by the correction methods could be assessed. The results obtained in the study showed that, using both corrections, the system is able to detect 10% monosomies with a significance level $\alpha = 0.1\%$. Cytometry 31:93–99, 1998. © 1998 Wiley-Liss, Inc.

Key terms: fluorescence in situ hybridization; interphase nucleus; optical microscopy; confusion matrix; proportion estimation; hypothesis testing

The suitability of fluorescence in situ hybridization (FISH) as a tool for the diagnosis and prognosis of different genetically aberrant related diseases has been extensively reported since the development of this cytogenetic technique (9,19). There are many research studies which show the suitability of this method for the detection and screening of numerical and/or structural genetic aberrations in both metaphase spreads and interphase nuclei (1,2,11,20,21).

The opportunity that FISH provides for analyzing specific genomic sequences in interphase nuclei is particularly advantageous in those cases where the number of cells that have to be analyzed is too high to be achieved through good quality metaphase spreads. It can be applied when the presence (or absence), number, or position of some known DNA sequences can characterize a chromosomal aberration.

The development of new image acquisition devices in parallel with the evolution and increasing computation power of the last generation computers have made feasible the automation of the analysis of FISH 2D samples, namely of blood samples or cell suspensions obtained by physical and/or chemical dissociation of solid tissues (26). This automation had been strongly recommended in some statistically oriented studies (6,8,12) as the only reasonable way to detect sporadic or low rate aberrations with small error rates.

A first attempt to detect and score FISH signals automatically in interphase nuclei has been reported by Netten et al. (16), but further insight is necessary in order to estimate the prospective application of this kind of systems.

A system based on a controllable motorized microscope and image acquisition and processing devices was developed in our laboratory. The performance of that system for automatic detection and scoring of FISH signals in interphase nuclei is presented in this paper. A brief description of the system and the image processing algorithms is included, but a more detailed description of them can be found elsewhere (17).

In this paper we mainly focus on the experiments that were carried out to evaluate the performance of our system, the controls implemented, the methodology used

Contract grant sponsor: ARCADIM Project; Contract grant number: CYCYT TIC92-0922-C02-01; Contract grant sponsor: European Concerted Action CA-AMCA; Contract grant number: BMH1-CT92-1307.

*Correspondence to: Carlos Ortiz de Solórzano, Mailstop 74/157, Resource for Molecular Cytogenetics, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720.
E-mail: carlos@rmc.lbl.gov

to calibrate the system and the final results obtained on artificially produced mosaic samples.

We first summarize the sample preparation protocol and describe the system instrumentation and algorithms. Then the method to correct systematic errors is presented, with its application to actual scanned samples. The control strategy and statistical validation of the sensitivity results are shown next. Finally the results of the experiments are presented and discussed in the last two sections of the paper.

MATERIALS AND METHODS

Sample Preparation

Five milliliters of blood were extracted from two healthy individuals (a woman and a man). The number of lymphocytes was measured in both samples. Nine mixtures (M1–M9) of blood from the woman and the man were formed, in different proportions, $M_1 = 100/0$, $M_2 = 95/5$, $M_3 = 90/10$, $M_4 = 75/25$, $M_5 = 50/50$ and the symmetric ones M_6–M_9.

The protocol used was the following: After 72 h of phytohemaglutinin stimulated lymphocyte culture, cells were harvested, KCl hypotonic treated for 10 min, methanol:glacial acetic acid fixated and air dried after dropping on slides.

The slides were stored at −20°C until they were treated following procedures of a normal hybridization (23). Probe hybridization was achieved by denaturing slides for 2 min in 70% formamide/2 × SSC at 65°C and then quickly quenching in ice-cold 70% (v/v) ethanol and dehydrating in serial ethanol washes (80%, 90%, 100%).

Hypotonic shock was applied on the samples to break the cellular membranes and wash out the cellular proteins in order to facilitate the probes access to their complementary sites in the nuclear DNA. A Carnoy 3:1 wash was carried out to remove the cytoplasmatic debris and fix the cells to the glass.

Each preparation was hybridized to two satellite probes for chromosomes Y and X. The goal of the experiments was to detect the chromosome X probe while the chromosome Y probe was used as a control, as will be explained on below. The hybridization was performed using a combination of probes (SO CEP/SG CEO Y) provided for this experiment by Vysis Inc. (Downers Grove, Illinois). CEP X (α satellite) SpectrumOrange hybridizes the centromeric region of chromosome X (p11.1-q11, locus DXZ1). CEP Y (satellite III) SpectrumGreen hybridizes the satellite III of chromosome Y (band Yq12, locus DYZ1). After hybridization, slides were stained with 6-diamino phenylindole dihydrochloride (DAPI).

System Description

The system installed in our laboratory performs in a sequential way the set of actions which are necessary to emulate human actions in a visual inspection of the samples. Namely, the actions automatically carried out by the system are as follows: 1) focusing the microscope on every field of view, 2) acquiring the counterstained and FISH labeled images, 3) analyzing the images in order to define the position of the nuclei and the number of FISH signals inside of them; and 4) performing all the stage movements and filter changes required to scan the area under study.

The elements that make up the system, which allow the analysis to be done automatically, are briefly described in the following paragraphs.

Microscope. The microscope used is an Ergolux (Leica, Wetzlar, Germany), with a motorized scanning stage (Marzhäuser, Wetzlar, Germany), eight slides wide. The excitation/emission filter blocks, interference filters and objectives can be positioned automatically. Motor control is performed by a stepping motor controller (SMOC) (Metasystem, Sandhausen, Germany). The SMOC controller communicates with the CPU via a serial RS-232 connection. The objective used in this experiment was an oil immersion fluorescence objective, ×63 (NA 1.3), which was considered appropriate taking into account considerations about resolution, magnification, sample brightness and depth of focus.

Camera. We used a MicroImager 1400 CCD camera (XilliX Technologies Corp., Richmond, B.C., Canada). Its image sensor device is a Kodak KAF 1400 CCD chip with 1,344 × 1,038 pixels resolution, and 6.8 × 6.8 μm pixel spacing. The camera is attached to the microscope using a standard C mount adapter. The CCD can be clocked in pixel additive mode (*binning*), in which 4 adjacent pixels are combined during clock-out to increase sensitivity and to obtain a higher frame transfer rate. Hence a true 2 × 2 binning is achieved. The camera is controlled by a DC1 multi-I/O board (Access Dynamics, Alamo Gordo, California) connected through a VSB (VME subsystem Bus) to a 16 Mb memory board to store the acquired images. The memory board and the controller are attached to the VME workstation backplane.

CPU. The core of the system is a SparcStation 4/370 (SUN Microsystems, Mountain View, California), 32 Mb RAM, 1 Gb HD, VME bus, with a UNIX SUN OS 4.1.1 Operating System. It controls the SMOC via RS232, and the Xillix camera through the DC1 board. Images are retrieved and stored in the RAM from the 16 Mb memory board through the VME bus.

Image Acquisition

Since a monochrome CCD camera was used, a set of excitation/emission filters was used to acquire images containing the emission of the different fluorochromes present in the samples:

Counterstain: A Leica A filter block (exc. BP 340-380 nm, ems. LP 470 nm) was used. The acquisition time used to get bright, high contrast images was 0.1 s.

Chromosome X Probe (Spectrum Orange): The filter used was a Leica L4 (exc. BP 515-560 nm, ems. LP 580 nm). Acquisition time: 2.0 s.

Chromosome Y Probe (Spectrum Green): The filter used was a Leica N 2.1 (exc. BP 450-490 nm, ems. BP 515-560 nm) Acquisition time: 5.0 s.

The acquisition was performed using the binning feature of the Xillix camera. The effective pixel size is then 13.6 × 13.6 μm, which, considering the magnification used, corresponds to a microscopic square 0.21 μm wide. The source of light was a mercury arc lamp HBO 50/AC-L (OSRAM , Germany), 50w power.

Sequence of Work and Image Analysis

The purpose of this paragraph is to summarize the sequence of actions that the system carries out in order to perform the analysis. A complete description of the algorithms can be found in (17). Initially, some user interaction is required to determine the area of work in the sample and also to manually focus the microscope in one field of view of the microscope. This in-focus position is used as a reference position when automatically focusing the whole area under analysis.

The automatic focusing in every field is carried out using a three-step range focusing algorithm (4). An autocorrelation-based measure (25) was used to determine the degree of focusing of images taken at different positions on the focus axis for every field of view that was focused. The autocorrelation-based focus function provides an outstanding measurement of the degree of focusing when compared with other functions described in the literature. In order to substantially reduce the amount of time necessary to focus the area under study, only 1 out of N fields were focused using the range focusing algorithm, and the in-focus position for the intermediate fields was calculated by bicubic interpolation from the surrounding fields. The value of N was calculated from the maximum error allowed in the interpolation.

Once the images are focused, three images, the DNA counterstained image and both FISH signal emission images, are acquired and preprocessed before being analyzed. The preprocessing steps involve shading correction, background subtraction, autofluorescence correction, and color-shift compensation (7). After being corrected, images are segmented in order to extract the required information. The strategy used is different for the counterstained and FISH labeled images. Nuclei are segmented by automatically thresholding the histograms of the DNA counterstained images using an ISODATA algorithm (18). Clusters of nuclei are divided into their individual components using a morphological Watershed algorithm (3) over the Distance Transform of the ISODATA binarization of the original image (15). FISH signals are extracted by means of a Top Hat Transform (22) followed by a recursive Reconstruction algorithm (24) whose aim is removing secondary peaks and refining the contours of the real FISH signals.

Systematic Error Correction

The standard FISH analysis performance, to which other systems are compared, is the one achieved by an experienced human operator. This assumption made, the system has to be ''calibrated'' by training, detecting systematic deviations from the results provided by a ''trained'' operator.

Systematic deviations can be due to errors in the discrimination of the signals (touching probes detected as single probes, signals of irregular size and shape due to a different condensation state of the chromatin or due to the sample preparation, slightly out-of-focus images or overlapping FISH signals, etc.). The method to correct estimated populations described by Castleman in (8) was used to reduce the influence of some of these errors.

In his paper, Castleman proposes an unbiased estimator p', of the vector p ($p_1, p_{2,\ldots}, p_k$) − proportions of cells with 0, 1, 2, . . . , k FISH signals-based on the observations q' obtained by the automatic system. The estimator p' is the value which minimizes the quadratic error (MQE).

$$MQE = \frac{1}{N}\sum_{i=1}^{k} \lambda_i E[(\, p_i - p_i')^2] \tag{1}$$

where $E[\,]$ is the expectation operator, N is the total number of cells analyzed and λ is a cost vector which allows us to penalize the most serious errors. In case that all errors are considered equally harmful, all the values in vector λ should be equal.

Castleman's unbiased estimator is,

$$p' = [C^T]^{-1}\, q' \tag{2}$$

where C is a so-called *Confusion Matrix* whose elements $\epsilon_{i,j}$, are the probability (experimentally evaluated) of a nucleus actually having i FISH signals, being detected by the system as having j FISH signals.

The minimum quadratic error for this estimator is

$$MQE = \frac{1}{N}\sum_{i=1}^{k} \lambda_i\, p_i\, (1 - p_i) + \frac{1}{N}\, p^T B \lambda \tag{3}$$

where the elements of B (*Befuddlement Matrix*) can be calculated from the elements in C as

$$B_{ml} = \sum_{i=1}^{k} C_{ml}\, [C_{li}^{-1}]^2 - \delta_{ml} \tag{4}$$

$$\delta_{ml} = \begin{cases} 1 & m = l \\ 0 & m \neq l \end{cases}$$

To calculate the matrix C, 100 nuclei of each of the M1-M9 samples were analyzed by the automatic system and classified as having 0, 1, 2, or more than 2 FISH signals. An experienced operator visually determined on the microscope the accuracy of the results provided by the system. In this way, nine confidence matrices were obtained, and finally a ''global'' matrix combining the previous matrices was calculated. The confusion matrix for sample M_1 and the global confusion matrix are presented in Table 1.

The correction matrix $[C^T]^{-1}$ can be obtained on a sample by sample basis (nine different C matrices for the

Table 1
Confusion Matrices for Sample M1 (left) and Global Confusion Matrix (right)

	System				System			
Operator	0	1	2	≥3	0	1	2	≥3
0	0.75	0.25	0	0	0.87650	0.0312	0.0625	0.0312
1	0.1	0.9	0	0	0.0637	0.9023	0.0297	0.0042
2	0	0.1463	0.8049	0.0488	0.0310	0.0960	0.8483	0.0248
≥3	0	0.25	0.25	0.50	0	0.1429	0.1429	0.7143

nine samples) or globally from the whole set of samples analyzed. Both types of correction were tested in our experiments and the results compared.

Internal Control of the Experiments

The previously described strategy for systematic error correction uses an unbiased estimator of the ''detectable'' population on the samples to get the same results as the experienced human observer. However these detectable populations do not necessarily correspond to the underlying ''actual'' populations; there are many reasons why the fluorescence signals sometimes are not correctly detected (both by the system and by the human operator) or are detected in a proportion different from what was expected, namely: the non-ideal hybridization efficiency; local defects in parts of the slides; high levels of uncorrected autofluorescence; damage in the cells; overlapping of signals; misclassification of artifactual organic residues as FISH signals; etc.

In order to avoid these differences between the ''detectable'' and the ''actual'' populations, a control strategy was used. The strategy allowed us to detect and discard some of the cases where hybridization defects may have occurred.

The possibility of hybridizing two or more FISH probes labeled with fluorochromes that can be spectrally resolved by using different sets of excitation/emission filters allowed us to use a control strategy internal to the experiment. In that way, one of the probes was used as a control probe for the experiment. This assured a more accurate experiment control than using an external control.

In general, the internal control can be implemented by using a probe whose copy number is ''a priori'' known, and the control is achieved by looking at the number of FISH control signals: those nuclei whose control copy numbers are different than the expected number are discarded.

In our specific experiment, in which we are using the Y chromosome probe as control, and the percentage of male and female cells in the samples is the variable to be detected, we can apply an even more specific control. The control was used as follows: for every nucleus we counted the number of X and Y FISH signals; we discarded nuclei having a single test probe (X chr.) and no control probe (Y chr.), and nuclei with two test probes and one or more control probes. These cases represent erroneous situations in which a problem in the hybridization or some artifactual error can be presumed.

Evaluation of the System Sensitivity

In order to evaluate the sensitivity of the automatic system to detect monosomic mosaicisms, we applied statistical tests to analyze the differences between each monosomic distribution (M_2–M_9) and the ''pure'' disomic (sample M_1). The tests were applied to the distributions before and after the systemic correction, after the internal control alone, and after both correction and control. The statistics were compared to determine the improvement in sensitivity provided by each one of the correction methods.

Following (12), two tests were used. A discrete version of the Kolmogorov–Smirnov (KS) maximum deviation test (5) was used to detect non-specific differences between each pair of distributions (M_1 versus M_i, for i = 2, . . , 9). The null hypothesis H0 (both populations are not significantly different) was tested against the hypothesis H1 (the populations are significantly different). No assumptions were made for the underlying populations, besides random sampling, since the KS is a non-parametric, distribution-free statistical test.

Then we used a second, more powerful, test in order to detect significant differences associated to one particular element of the distribution. In this case the null hypothesis H0 (both populations are not significantly different) was tested as opposed to the alternate hypothesis H1 (some element of the distributions is significantly different). The Z'max test was used to detect significant differences between the number of FISH signals per nucleus ($i = 1, \ldots, k$) in the disomic (M_1) sample and in the rest of the samples (population M_1 is considered as a control population for the rest of the samples).

Therefore the null hypothesis for this test was

$$H_0 : p_t(i) = p_c(i) \quad \forall i$$

being $i = 1 \ldots k$, and p_t and p_c the distributions of the test and control samples, and the alternate hypothesis was

$$H_1 : p_t(j) > p_c(j)$$

for at least one value of j. The statistic applied for each value i should be ,

$$z(i) = \frac{p_t(i) - p_c(i)}{\sqrt{\frac{p_t(i)(1 - p_t(i))}{n_t} + \frac{p_c(i)(1 - p_c(i))}{n_c}}} \qquad (5)$$

where n_t and n_c are the number of test and control nuclei analyzed. The relevant statistic for the test is z'max =

Table 2
Proportion of Cells Obtained with Each Number of FISH Signals by the System (s) and by the Operator (o) on Samples M_1–M_9

Sample	Number of FISH signal detected (*n*) 0	1	2	≥3	K–S	Euclidian distance
M_1 (s)	0.036	0.263	0.645	0.055	0.1091	80.92
M_1 (o)	0.032	0.159	0.768	0.040		
M_2 (s)	0.095	0.204	0.680	0.018	0.0984	64.24
M_2 (o)	0.058	0.143	0.787	0.010		
M_3 (s)	0.104	0.286	0.595	0.013	0.0673	4.28
M_3 (o)	0.087	0.235	0.666	0.010		
M_4 (s)	0.203	0.453	0.343	0	0.0785	48.15
M_4 (o)	0.158	0.421	0.421	0		
M_5 (s)	0.062	0.562	0.362	0.012	0.0101	9.04
M_5 (o)	0.056	0.569	0.367	0.005		
M_6 (s)	0.056	0.690	0.242	0.011	0.0619	52.56
M_6 (o)	0.032	0.775	0.185	0.005		
M_7 (s)	0.030	0.778	0.180	0.011	0.0519	42.45
M_7 (o)	0.011	0.849	0.139	0		
M_8 (s)	0.288	0.688	0.042	0	0.1311	94.14
M_8 (o)	0.146	0.814	0.039	0		
M_9 (s)	0.142	0.819	0.038	0	0.0613	44.52
M_9 (o)	0.080	0.884	0.035	0		

The K–S statistic for the comparison between the system and operator counts (critical value for $\alpha = 0.1$ is $c = 0.49$), and the eulidian distance between manual and automatic counts.

Table 3
Proportion of Cells Obtained by the System with Each Number of Fish Signals, Modified with the Correction of the Systematic Error

Sample	Number of FISH signals (*n*) 0	1	2	≥3	K–S	Euclidian distance
M_1 (pc)	0.0288	0.1473	0.7920	0.0317	0.0148	14.40
M_1 (gc)	0	0.2086	0.7494	0.0421	0.0320	30.94
M_2 (pc)	0.0837	0.1592	0.7817	−0.0237	0.0398	22.56
M_2 (gc)	0.0709	0.1406	0.7930	−0.0046	0.0152	10.33
M_3 (pc)	0.0670	0.2285	0.7087	−0.0043	0.0277	24.74
M_3 (gc)	0.0768	0.2427	0.6890	−0.0105	0.0185	15.84
M_4 (pc)	0.1556	0.4225	0.4219	0	0.0006	0.44
M_4 (gc)	0.1852	0.4592	0.3787	−0.0239	0.0645	33.43
M_5 (pc)	0.0371	0.5939	0.3662	0.0028	0.0150	13.08
M_5 (gc)	0.0143	0.5794	0.4056	−0.0013	0.0377	26.52
M_6 (pc)	0.0365	0.7703	0.1814	0.0118	0.0059	5.03
M_6 (gc)	0.0012	0.7368	0.2591	0.0020	0.0698	44.19
M_7 (pc)	0.0016	0.8262	0.1748	−0.0025	0.0329	21.82
M_7 (gc)	−0.0336	0.8429	0.1842	0.0055	0.0501	32.09
M_8 (pc)	0.1641	0.8026	0.0332	0	0.0180	11.28
M_8 (gc)	0.2740	0.7550	0.0057	−0.0166	0.1155	85.55
M_9 (pc)	0.0728	0.8983	0.0288	0	0.0078	8.86
M_9 (gc)	0.0961	0.9051	0.0077	−0.0098	0.0361	19.43

(pc) Data corrected with the sample matrix; (gc) data corrected with the global matrix. The table includes the euclidian distance between the corrected counts and the manual counts displayed in Table 2.

max(z(i)). This test provides *k* related statistics that follow a multivariate normal distribution.

The comparison of the statistic and a critical value $c(\alpha)$ provides the probability α of accepting the null hypothesis (both distributions are similar) when the alternate hypothesis is true (both distributions are different). Ewens (10) suggested that $c(\alpha)$ can be approximated by a multinomial distribution and gives the significance levels for different numbers of elements in the distribution. The systemic error correction affects the independence of the distributions, invalidating the Z'max test. Therefore we applied this test only to the original distributions and to the distributions after the internal control.

RESULTS

Data Gathering

Five hundred nuclei from each one of the nine samples M_1–M_9 were analyzed by the automatic system and by an experienced operator. The results are summarized in Table 2. The distributions obtained by the system and by the operator on the same sample were compared with a K-S test. The column KS in Table 2 shows the value of the statistic for each sample. Finally, the last column in Table 2 shows a second measurement of the ''proximity'' between the manual and automatic count, which is the Euclidean distance. None of the automatic counts was found significantly different (significance $\alpha = 0.1$, critical value = 0.49) from the corresponding automatic counts. Even though the differences detected between the system and operator performance were further reduced, as will be seen, by using the systematic error correction and internal control methods described, we can point out here that this divergence is smaller than the one found in other studies for brightfield spot counting (13,15).

Results after Systematic Error Correction and Internal Control of the Experiments

The distributions obtained for the M_1–M_9 samples were corrected, following the procedure described in ''Systematic Error Correction'', with the Confusion Matrices calculated from each sample and with the global matrix calculated as a combination of the whole set of samples.

The results obtained after this correction are shown in Table 3, and the estimator's Minimum Quadratic Errors are displayed in Table 4. From these MQE values, we can conclude that there is not a substantial difference, regarding to the estimation error, between using the particular and the global matrices when performing the correction. This fact suggests the possibility of using ''historical'' global matrices associated with each type of cells, probes hybridized, protocols, fluorochromes and/or filter settings.

The improvement associated to the use of the correction method is clear. The KS statistics of the corrected distributions (Table 3) can be compared with the statistics of the uncorrected distribution shown in Table 2. The superiority of the correction based on the particular confidence matrices is neat when considering that the average KS value of the nine corrected populations is eighteen times the value of the average KS of the uncorrected distributions while the average KS value of the distributions corrected using the general matrix is only twice the value of the uncorrected. A similar result conclusion can be drawn by comparing the Euclidean Distances between the distributions before and after correction. After systematic error correction, we have applied the internal control method described in ''Internal Control of the Experiments''. The results after both

Table 4
Minimum Quadratic Error of the Estimation of the Underlying Population, when the Data is Corrected Using the Particular and Global Confidence Matrices

	MQE	
Sample	Particular matrix	Global matrix
M_1	0.0023 (6.98 10^{-4} + 0.0016)	0.0015 (7.86 10^{-4} + 7.62 10^{-4})
M_2	0.0017 (7.12 10^{-4} + 0.48 10−4)	0.0015 (6.92 $10^{\dagger-}$4 + 7.67 10^{-4})
M_3	0.0023 (8.18 10^{-4} + 0.0014)	0.0017 (9.20 10^{-4} + 7.35 10^{-4})
M_4	0.0017 (0.0012 + 4.41 10^{-4})	0.0019 (0.0012 + 6.52 10^{-4})
M_5	0.0015 (0.0010 + 4.52 10^{-4})	0.0016 (9.98 10^{-4} + 6.45 10^{-4})
M_6	0.0012 (7.4 10^{y4} + 4.11 10^{-4})	0.0014 (7.78 10^{-4} + 6.01 10^{-4})
M_7	0.0008 (5.73 10^{-4} + 2.64 10^{-4})	0.0011 (5.03 10^{-4} + 5.74 10^{-4})
M_8	0.0015 (6.55 10^{-4} + 8.52 10^{-4})	0.0013 (7.51 10^{-4} + 5.54 10^{-4})
M_9	0.0007 (3.43 10^{-4} + 7.16 10^{-4})	0.0008 (3.43 10^{-4} + 5.34 10^{-4})

The error componets corresponding to the distribution and to the system error appear between brackets (see eq. 3).

Table 5
Proportion of Cells Obtained by the System with Each Number of FISH Signals

	Number of FISH signals (n)			
Sample	0	1	2	≥3
M_1	0.040	0	0.916	0.044
M_2	0.074	0.050	0.873	0.003
M_3	0.055	0.112	0.832	0.001
M_4	0.149	0.440	0.411	0
M_5	0.046	0.524	0.421	0.009
M_6	0.057	0.704	0.223	0.016
M_7	−0.002	0.996	0.100	0.006
M_8	0.232	0.730	0.038	0
M_9	0.080	0.920	0	0

Modified with the correction of the systematic errors (based on each sample) and the internal control, as described in the text.

Table 6
Results of the K–S Statistic

	Step			
Sample	Initial data	After error correction	After control correction	After both corrections
M_2	* (0.06)	* (0.07)	* (0.08)	* (0.08)
M_3	* (0.09)	* (0.11)	* (0.18)	* (0.12)
M_4	* (0.35)	* (0.43)	S (0.56)	S (0.55)
M_5	* (0.32)	* (0.38)	S (0.53)	S (0.53)
M_6	* (0.44)	S (0.53)	S (0.71)	S (0.72)
M_7	S (0.52)	S (0.60)	S (0.84)	S (0.85)
M_8	S (0.66)	S (0.80)	S (0.90)	S (0.92)
M_9	S (0.66)	S (0.79)	S (0.95)	S (0.96)

*Means difference with M_1 not significant. S means difference with M_1 significant. Significance $\alpha = 0.1$. Critical value of the statistic, for $\alpha = 0.1$, $c = 0.49$.

corrections (particular systemic correction and internal control) are shown in Table 5. The improvement introduced by these two correction steps is reported in the next section.

System Sensitivity

The hypothesis validation tests described earlier in ''Evaluation of the System Sensitivity'' were used to evaluate the significance of the obtained results. Their purpose was to prove whether a significant improvement had been brought about by the correction and control strategies described in ''Systematic Error Correction'' and ''Internal Control of the Experiments''. Tables 6 and 7 show the results. Values labeled with an S correspond to those distributions that the test found significantly different from the disomic population (significance level $\alpha = 0.1$). The KS test (summarized in Table 6), used to detect 'non-specific' significant differences between the ''control'' and ''test'' population, detected a significant difference between the population M_4 (25% monosomic) and M_1 (100% pure disomic population), and shows the improvement in sensitivity achieved after the different correction steps. The values of the statistic show the progressive improvement in sensitivity achieved by the different correction steps. The Z'max test, more powerful than the KS detected as significantly different the population M_2 (5% monosomic).

Table 7
Results of the Corrected Z'max Statistic

	Step	
Sample	Initial data	After control correction
M_2	S (2.20)	S (2.84)
M_3	* (0.79)	S (6.54)
M_4	S (6.37)	S (14.73)
M_5	S (10.06)	S (20.43)
M_6	S (14.93)	S (30.22)
M_7	S (19.02)	S (49.86)
M_8	S (14.50)	S (25.05)
M_9	S (21.24)	S (42.26)

*Means difference with M_1 not significant. S means difference with M_1 significant. $\alpha = 0.1$. Critical value of the statistic, for $\alpha = 0.1$, $c = 1.96$.

DISCUSSION

The purpose of this paper, besides briefly describing a system and a set of algorithms that can be used to automatically detect and score FISH signals in interphase nuclei, is mainly to show a way to correct errors and to interpret the data obtained from it. The system described is fully automatic, requiring only a minimum amount of manual interaction in every sample (to determine the area

where the system has to operate), and in every type of sample (to train the classification algorithms and the calibration method). The results obtained directly by the system before any correction are very close to the ones obtained by a skilled human operator: the test used to compare the manual and automatic counts proved equality of both distributions for all the nine samples used. That allows us to state that the selection of the instrumentation and the algorithms was correct, as so was the system training. The use of Castleman's correction with both the particular (sample based) and general matrices improves the system performance. The particular method outperformed the general one in seven of the nine samples. On average and measuring the performance in terms of the KS averaged statistic, the particular method was nine times better than the general one. The fact that in some cases the general matrix outperforms the particular matrix may be due to the small number of nuclei used to generate some of the matrices, as can be seen in Tables 2 and 3. Future work based on a higher number of cells analyzed may be needed to improve the accuracy of the general and mainly the particular matrices. The use of the general matrix, although in principle less accurate, is very useful as it does not need a per-slide system training, which is expensive in time. Therefore a general correction based on a sufficiently high number of cells, when associated with correct standardization of the protocols, should be preferred to the particular correction, due to the minimum effort demanded.

Another conclusion of this analysis is that the use of the internal, nucleus by nucleus, control, provided by the simultaneous use of two probes, greatly improves the system sensitivity, as can been seen in Tables 6 and 7. This outstanding improvement can be somewhat explained however by the specific (experiment related) type of control used. The normal control used in a real scenario, although improving the sensitivity, is not expected to provide such good results obtained in our ''artificial'' monosomies.

On the whole, both correction methods showed that the system is able to detect 5% monosomic populations with a significance level $\alpha = 0.1$. This figure proves, in our opinion, the adequacy of the applied strategies of correction and control. The proposed system for FISH spot counting in interphase nuclei is a step to improve and optimize the resources of clinical and research cytogenetic laboratories and to remove the main obstacle of time and resource consumption for the widespread use of FISH interphase analysis for the detection of many low-frequency numerical genetic aberrations.

Finally, a future suggestive work is the use of the actual mosaic proportions (known in these ''artificial'' samples) as the true values in Castleman's correction, instead of the human counts. This could be an attempt to correct for all sources of error at the same time, even those that the humans are subject to. Theoretically, that would produce estimates even superior to human dot counts.

LITERATURE CITED

1. Anastasi J, Thangavelu M, Vardiman JW, Hooberman AL, Bian ML, Larson RA, Le Beau MM: Interphase cytogenetic analysis detects minimal residual disease in a case of acute lymphoblastic leukemia and resolves the question of origin of relapse after alogenetic bone marrow transplantation. Blood 77:1087–1091, 1991.
2. Anastasi J, Le Beau MM, Vardiman JW, Fernals AA, Larson RA, Rowley JD: Detection of trisomy 12 in chronic lymphocytic leukemia by fluorescence in situ hybridization to interphase cells: A simple and sensitive method. Blood 79:1796–1801, 1992.
3. Beucher S, Meyer F: The morphological approach to segmentation: The watershed transformation. In: Mathematical Morphology in Image Processing, Dougherty E (ed). Marcel Dekker, Inc., pp 433–481, 1992.
4. Boddeke FR, van Vliet LJ, Netten H, Young IT: Autofocusing in microscopy based on the OTF and sampling. Bioimaging 2:193–203, 1994.
5. Bradley JV: Distribution-Free Statistical Tests. Prentice-Hall, Englewood Cliffs, New Jersey, 1968.
6. Carothers A: Counting, measuring, and mapping in FISH-labelled cells: Sample size considerations and implications for automation. Cytometry 16:298–304, 1994.
7. Castleman KR: Color compensation for digitized FISH images. Bioimaging 1:159–165, 1993.
8. Castleman KR, White BS: Dot count proportion estimation in FISH specimens. Bioimaging 3:88–93, 1995.
9. Cremer T, Landegent J, Bruckner A, Scholl H, Schardin M, Hager H, Devilee P, Pearson P, van der Ploeg M: Detection of chromosome aberrations in the human interphase nucleus by visualization of specific target DNAs. Human Genetics 74:199–200, 1988.
10. Ewens WJ, Griffiths RC, Ethier SN, Wilcox SA, Graves JAM: Statistical analysis of in situ hybridization data: Derivation and use of Z-max test. Genomics 12:675–682, 1992.
11. Hopman AHN, Ramaekers FCS, Raap AK, Beck JLM, Devilee P, van del Ploeg M, Vooijs GP: In situ hybridization as a tool to study numerical chromosome abnormalities in solid bladder tumors. Histochemistry 89:307–316, 1988.
12. Kibbelaar RE, Kok F, Dreef EJ, Kleiverda JK, Cornelisse CJ, Raap AK, Kluin PhM: Statistical methods in interphase cytogenetics: An experimental approach. Cytometry 14:716–724, 1993.
13. Krijtenburg PJ, Alers JC, Bosman FT, van Dekken H: Comparison of automated and manual analysis of interphase in situ hybridization signals in tissue sections and nuclear suspensions. Cytometry 25:99–103, 1996.
14. Malpica N, Ortiz de Solórzano C, Vaquero JJ, Santos A, Vallcorba I, García–Sagredo JM, del Pozo F: Applying watershed algorithms to the segmentation of clustered nuclei. Cytometry 28:289–297, 1997.
15. Moore DH, Epstein L, Reeder J, Wheeless L, Waldman FM, NCI Bladder Tumor Marker Network: Interlaboratory variability in fluorescence in situ hybridization analysis. Cytometry 25:125–132, 1996.
16. Netten H, Young IT, van Vliet LJ, Tanke HJ, Vroljik H, Sloos WCR: FISH and chips: Automation of fluorescent dot counting in interphase cell nuclei. Cytometry 28:1–10, 1997.
17. Ortiz de Solórzano C: Análisis y caracterización de imágenes de microscopía de muy baja luminosidad: Automatización del Análisis Citogenético. Ph.D. Thesis. Universidad Politécnica de Madrid, December 1996.
18. Otsu N: A threshold selection method from gray level histogram. IEEE Trans. System Man Cybernetics. SMC-8:62–66, 1978.
19. Pinkel D, Straume T, Gray JW: Cytogenetic analysis using quantitative, high sensitivity, fluorescence hybridization. Proc. Natl. Acad. Sci. USA 83:2934–2938, 1986.
20. Poddighe PJ, Moesker O, Smeets D, Awwad BH, Ramaekers CS, Hoopman AHN: Interphase cytogenetics of hematological cancer: Comparison of classical karyotyping and in situ hybridization using a panel of eleven chromosome specific DNA probes. Cancer Res 51:1959–1967, 1991.
21. Poddighe PJ, Ramaekers FCS, Hopman AHN: Interphase cytogenetics of tumours. J Pathol 166:215–224, 1992.
22. Serra J: Image Analysis and Mathematical Morphology. Academic Press, New York, 1982, pp 436–437.
23. Vázquez–Mazariego Y, Cabello P, García–Sagredo JM, Lopez–Yarto A, Vallcorba I, Resino M, Muñoz R, Pérez I, Mayayo M, Ferro MT, SanRomán C: Burkitt lymphoma with a duplication of der(8) t(2;8) derivative: Interpretation of a complex caryotype by chromosome painting. Cancer Gen Cytogen 76:136–139, 1994.
24. Vincent L: Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. IEEE Trans. on Image Proc. 2:172–201, 1993.
25. Vollath D: Automatic focussing by correlative methods. J Microscopy 147:279–288, 1987.
26. Vrolijk H, Sloos WCR, van de Rijke FM, Mesker WE, Netten H, Young IT, Raap AK, Tanke HJ: Automation of spot counting in interphase cytogenetics using Brightfield microscopy. Cytometry 24:158–166, 1996.