Background: Detection of fluorescent probes by fluorescence in situ hybridization in cells with preserved threedimensional nuclear structures (3D-FISH) is useful for studying the organization of chromatin and localization of genes in interphase nuclei. Fast and reliable measurements of the relative positioning of fluorescent spots specific to subchromosomal regions and genes would improve understanding of cell structure and function.

Methods: 3D-FISH protocol, confocal microscopy, and digital image analysis were used.

Results: New software (Smart 3D-FISH) has been developed to automate the process of spot segmentation and distance measurements in images from 3D-FISH experiments. It can handle any number of fluorescent spots and incorporate images of 4',6-diamino-2-phenylindole counterstained nuclei to measure the relative positioning of spot loci in the nucleus and inter-spot distance. Results from a pilot experiment using Smart 3D-FISH on ENL, MLL, and AF4 genes in two lymphoblastic cell lines were satisfactory and consistent with data published in the literature.

Conclusion: Smart 3D-FISH should greatly facilitate image processing and analysis of 3D-FISH images by providing a useful tool to overcome the laborious task of image segmentation based on user-defined parameters and decrease subjectivity in data analysis. It is available as a set of plugins for ImageJ software. © 2005 Wiley-Liss, Inc.
that the spatial proximity of genes could be a key factor in
the induction of translocations (14–18). Further, the differ-
ences in relative position of CTs in different cell lineages
could contribute to the large variety of translocations,
especially in hematoologic cancer (19).

To further address this relation, the spatial distribution
of CTs, or intergenic distances between genes potentially
involved in chromosomal translocations, must be accu-
rately analyzed. The 3D-FISH technique allows one to pro-
duce 3D images where real gene-to-gene distances can be
measured with better precision than with classic 2D-FISH
images (20). In a 2D-FISH experiment, such a task is quite
straightforward, but it often, if not always, leads to erro-
neous data due to the localization of spots (genes) on dif-
f erent focal planes (2D projection). In a 3D-FISH experi-
ment, the analysis is a heavy workload and complex task
for the biologist. Obtaining reliable data involves multiple
steps: visual inspection for selecting a region of interest,
manual segmentation, and detection of the center of spots
corresponding to the localization of genes. Because hun-
dreds of image stacks need to be analyzed for reaching sta-
tistical significance, the analysis is very time consuming.
Further, and importantly, it has the bias of the user’s sub-
jectivity and is not well reproducible or reliable.

The goal of the present work is to provide an user-
friendly software to improve the quality of rather high
throughput analysis of 3D-FISH images and the subse-
cquent positioning of genes and measurement of intergenic
distances. With the help of this software, named Smart
3D-FISH, image stacks corresponding to hundreds of cells
can be automatically analyzed overnight on a desktop
computer. The software can be used on any FISH studies,
including 2D studies. It can analyze images with a virtually
unlimited number of color channels (probes). The results
are saved as text files that can be directly incorporated
into standard spreadsheet software. In the first part, the
procedure of image processing and software are
described. The second part deals with the validation of
these procedures and software in a 3D-FISH experiment
involving three genes (three color probes) and 4',6-dia-
mino-2-phenylindole (DAPI) counterstained nuclei.

This new software has been successfully applied in a
3D-FISH experiment aimed at measuring the 3D radius dis-
tribution of genes involved in acute leukemia, namely
MLL, AF4, and ENL, in two lymphoblastic cell lines. Smart
3D-FISH is available as a set of plug-ins for the image analy-
isim software (21). It can be freely downloaded from

MATERIALS AND METHODS

Cell Culture and Karyotype

Two human cell lines were chosen due to their near
diploid karyotype. A human B-cell precursor leukemia
(NALM-6) was kindly provided by Lagneaux L. (Brussels,
Belgium) with the karyotype 46(43-47)<2n>XY,
t(5;12)(q33.2;p13.2); and we purchased a human B lym-
phoblastoid cell line (IM-9, DSMZ Braunschweig, Ger-
mannot) that has an apparently normal karyotype 46(44-
46)<2n>XX. These two lymphoblastic cell lines were cul-
tured in RPMI-1640 medium supplemented with 10% fetal
calf serum, 1% glutamine, in the absence of antibiotics, at
37°C in a humidified atmosphere containing 5% CO2.

DNA Probes

Bacterial artificial chromosome/phage artificial chromo-
some (BAC/PAC) clones were obtained from the Resources
for Molecular Cytogenetics (RMC) database (http://
www.biolologia.uniba.it/rmc/) in Italy: CTD-217A21 for the
MLL gene (22) and RP11-476C8 for the AF4 gene. RP11-
2344B19 for the ENL gene was purchased from Invitrogen
(Carlsbad, CA, USA). Clones were grown in LB medium
with appropriate antibiotics: kanamycin (50 μg/ml) for PAC
clones or chloramphenicol (12.5 μg/ml) for BAC clones.
DNA extraction was carried out according to the protocols
in the RMC database. The probes of MLL, AF4, and ENL
genes were labeled by nick translation by incorporating
fluorophore-tagged nucleotides dUTP-Alexa 488 (Molecular
Probes Europe, Netherlands), dUTP-Cy5, and dUTP-Cy3
(Amersham, Buckinghamshire, UK), respectively. Four hun-
dred nanograms of labeled probes, 10 μg of human cot-I-
DNA, and 5 μg of salmon sperm DNA were mixed in med-
ium containing 50% deionized formamide, 2× standard sal-
cite citrate (SSC), 10% dextran sulfate, and 0.5 M of sodium
phosphate dibasic:sodium phosphate monobasic (5:1).

3D-FISH Experiment

A dense cell suspension in 1× phosphate buffer saline
(PBS) was applied to slides coated with poly-l-lysine for
10 min to allow cell adhesion. Cells were then fixed in
buffered 4% paraformaldehyde in 1× PBS for 10 min to
preserve the native 3D structure of the nuclei and then
washed 3 times for 5 min in 1× PBS. During all proce-
dures, air drying was carefully avoided. Cells were per-
meabilized with 0.5% Triton X-100 and 0.5% saponin in
1× PBS for 15 min. After a bath in 20% glycerol in 1× PBS
for 20 min at room temperature, cells were freeze-thawed
by briefly dipping the slides three times in liquid nitrogen.
Cells were treated with 400 μg/ml of RNase A (Roche
Diagnostics, Myelan, France) for 15 min at 37°C. After a 5-
min bath in 0.1 N HCl, cells were washed in 2× SSC for 5
min and incubated in 50% formaldehyde and 2× SSC, pH
7.5, for 1 h. The probe preparation was then dropped onto
slides. Cells and probes were simultaneous dena-
tured at 75°C for 8 min. Cells were incubated overnight at
37°C in a humidified chamber. Post-hybridization washes
were performed a first time in 50% formaldehyde and 2×
SSC, pH 7.2, at 42°C, 3 times for 15 min and a second time
in 0.1× SSC, pH 7.2, at 60°C, 3 times for 15 min. Nuclei
were counterstained with DAPI at 0.2 μg/ml. Slides were
mounted in Vectashield medium.

Image Acquisition and Measurement

Confocal microscopy was carried out using a TCS con-
focal imaging system (Leica Instruments, Heidelberg, Ger-
many) equipped with a 63× objective. For Alexa 488,
Cy3, and Cy5 excitations, an argon-krypton ion laser and
a helium-neon ion laser were adjusted to 488, 568, and 647 nm, respectively. A biphotonic device was used for DAPI excitation. For each optical section, four fluorescence images were acquired in a sequential mode (i.e., Cy5, Cy3, Alexa 488, and DAPI). The confocal pinhole was adjusted to allow a minimum field depth. The focus step between sections was generally 0.35 μm (which corresponds to the optimal optical resolution) and the XY pixelization was set to 100 nm. Focal series were then processed to produce a single composite image file (stack).

Typically, a stack of 40 confocal planes was acquired.

Total DNA is counterstained by DAPI dye. Due to the biphotonic acquisition mode on the microscope, images from the DAPI channel are shifted. To register them with other images from other color channels, DAPI images were translated. The XY translation was then implemented to automatically shift the area of interest for the DAPI channel.

The radius of each nucleus was estimated according to a spherical approximation by measuring the volume of the nucleus, counterstained by DAPI, in 3D-FISH images. The distance of genes to the nuclear center was then measured and expressed as a percentage of the nuclear radius. Statistical analyses were performed using Student’s t test with an α value of 0.05.

**Image Processing**

All algorithms were implemented in JAVA and can be run as plugins on the free multiplatform ImageJ software. A possible procedure for 3D processing is to apply 2D processing in a slice-by-slice manner. However, slice-by-slice processing tends to remove weak signals that could be considered pixels of noise in the 2D slice but that are part of a signal in 3D images. Hence, all implemented procedures were adapted to work with voxels (i.e., incorporating X, Y, and Z directions) in 3D images. The implemented algorithms can work with real 3D color images. A 3D color image is considered as a set of unlimited numbers of 3D gray-level images (each corresponding to a different color channel). For a set of images less than or equal to three channels, the standard red/green/blue (RGB) color code is used for display. For more than three colors, each 3D gray-level image is distributed into at least one red, green, or blue channel; a 3D RGB image is then calculated by summing the contribution of all the 3D single gray-level images.

**Noise Filtering and Spot Segmentation**

Figure 1 shows the flow chart of the overall algorithm. The first step in processing the images is to remove background noise. The median filter is well suited for salt-and-pepper noise; for a general reference on image processing, see Russ (23) and Castleman (24). This filter, adapted to 3D images, is used with a neighborhood of radius equal to 2 pixels. Then a 3D Top Hat filter is applied to enhance the spots against the background and decrease the level of background noise. This filter is defined by: $\text{Top Hat}(\text{Im}) = \text{Im} - \text{Max7}(\text{Min7}(\text{Im}))$, as previously described (25). The second step is the segmentation of the spots. The central pixels inside each spot are determined to correspond to pixels whose value is greater than 99.95% of the histogram, i.e., with a value greater than $v$ ($v$ is defined as 99.95% of all pixel values in the range $[0 \rightarrow v]$). This value of 99.95% of the histogram was computed by estimating the volume of one spot and considering that a stack could comprise up to four spots.

More precisely, the pixel whose intensity is maximal in the 3D stack, but greater than 99.95% of the histogram, is detected. This pixel is a seed from where a spot will be segmented. A local threshold is computed corresponding to mean + σ of the pixels belonging to three lines in the three directions (X, Y, Z) passing through the detected center of the spot. The seed is extended, by 3D connectivity, to adjacent pixels whose value is greater than this local threshold to form an object. If the final segmented object is touching one border of the image in the X-Y plane, then it is removed. These procedures (detection of maximum pixel and 3D connectivity) are applied until no more unsegmented pixels have a value greater than 99.95% of the histogram. The segmentation is performed on all the different stacks except for the DAPI channel.

For the DAPI channel, a special segmentation is automatically performed. The radius of the first median filter applied to reduce the noise is increased to a radius of 4 pixels to obtain a more homogeneous signal. The stack is then thresholded using the “Isodata” algorithm (26).

For spots and DAPI segmentation, a 3D mathematical morphologic closing procedure is applied followed by a 3D mathematical morphologic opening procedure (radius 2 for gene images, radius 4 for nucleus DAPI images) to remove very small objects, fill holes, and make shape look more compact.

**Analysis and Validation of Segmentation**

A first analysis of the validity of the segmentation is done after the application of the median filter (Fig. 1). If the computed signal-to-noise ratio (SNR; maximum − minimum value of gray level in the image stack) is lower than a threshold fixed by the user, further measurements will be discarded into an invalidated spots file. A second analysis of the validity of the segmentation is based on the volume and number of objects. The list of segmented objects is analyzed to detect whether or not the segmented objects correspond to gene spots. Segmented objects with volumes smaller than the minimal volume ($\text{minVol}_i$), fixed by the user, are removed. In the same way, objects with volumes larger than the maximum volume ($\text{maxVol}_i$) are removed. They are often due to an incorrect experimental procedure or acquisition mode. Then, if the number of spots exceeds the value fixed by the user (two in the present work), the segmentation is considered invalid and further measurements will be performed, but the results will be redirected into the invalidated spots file.

In all cases, results of measurements are saved as text files. Two text files are created; the first text file will store the results for correct segmentation, and the other one will store the results for incorrect segmentation, i.e., where one parameter (SNR, volume and number of spots) does not correspond to the values fixed by the user. In this last case, annotations will indicate the nature of the problem.
Spot Separation Procedure

Figure 2 illustrates the spot separation procedure. The mathematical morphologic procedures may not be sufficient to separate two closely located spots. Close spots may be merged into one big spot with a volume roughly twice the average volume of one spot. The user can fix the minimum volume that may correspond to a merged spot ($minV_2$). The slices of the stack containing the object are projected along the Z axis onto a 2D plane to find local maxima that may correspond to the center of the two merged spots. The obtained projection is then smoothed so that local maxima are more easily detected. If two or more local maxima are encountered, they are separated into two clusters using a k-means algorithm. If the center of the two clusters are farther apart more than a fixed distance, corresponding approximately to the diameter of one spot ($dist$), the object can be separated into two smaller objects. The centers of these two clusters are the new $x$ and $y$ coordinates. The $z$ value of the two new objects is determined as the center of the primary segmented object along the $Z$ axis for the $x$ and $y$ coordinates of each new spot. Pixels of the primary big object are then resegregation.
Results

Image Processing and Software Description

From the confocal microscope, a stack is obtained where cells are present for all channels (corresponding to probe colors). An integrated utility was developed to help biologists select an area of interest delimiting one cell and save all the various channels in different directories. Stacks corresponding to ENL, AF4, and MLL genes were saved on the red, yellow, and green channels, respectively. An additional blue channel for DAPI counterstaining of nuclei was also used to determine the relative positioning of different probes to the center of nucleus.

The segmentation procedure was inspired by detection of microcalcification in mammograms using a RH-maxima procedure (27), with the exception that all the procedures were performed in 3D and not in a slice-by-slice manner.

The segmentation procedure was inspired by detection of microcalcification in mammograms using a RH-maxima procedure (27), with the exception that all the procedures were performed in 3D and not in a slice-by-slice manner.

FIG. 2. Spot separation process. A: 2D projection along the Y axis of the segmented image stack. White arrow indicates one object whose volume is between minV and maxV (Fig. 1) before the spot separation process. B: The same 2D projection proves that the object detected as one object is, in reality, composed of two colocalized spots. C: To distinguish these two spots on the segmented image, a 2D projection is performed along the Z axis to enhance the signal corresponding to each spot. Then local maxima are detected to find XY coordinates of each spot. If two local maxima are detected farther apart than the dist (see text), the two spots are separated. D: On the segmented stack, Z coordinates of each new spot are then detected along the Z axis (from XY coordinates previously detected) as the middle of the segmented object.
manner. One important issue is the threshold setting to find a seed for a spot. A high threshold corresponding to the 99.95% of the histogram was chosen. An automatic determination of this value was tested based on histogram analysis (supposing the presence of two or three classes correspond to the spots, the hybridization procedure noise, and the background) but was found to depend strongly on the image acquisition; we then chose to keep the fixed value of 99.95% as the threshold. However, to control the quality of the segmentation procedure, a set of parameters entered by the user is used to control the validity of the segmented objects (Fig. 1).

The software is comprised of several utilities that can be used separately or run consecutively thanks to a configuration file. A configuration file is created to process many directories, each containing as many subdirectories as number of colors. A certain number of parameters must be entered by the user such as the resolution in X-Y and Z. For each color, two SNR values are to be entered, namely $SNR_1$ and $SNR_2$. The SNR reflects the difference between the maximum and minimum of stack intensity, computed after 3D median filtering procedure, and indicates the presence of spots. However, a high SNR does not necessarily reflect the presence of bright spots, as would be the case if the offset of the microscope was not optimal. Nevertheless, a low SNR implies a small range of values and, in most cases, the absence of spots or spots insufficiently bright to be detected. If the computed SNR is below $SNR_1$, the results will be directed into an invalidated spots file. If the image SNR is between $SNR_1$ and $SNR_2$, the stack is labeled as “visual analysis” for further in-depth analysis. Then for each channel, parameters are to be filled such as the maximum volume for one spot ($maxV_1$), the minimum volume for one spot ($minV_1$), and the minimum volume that corresponds to the possible colocalization of two spots ($minV_2$). These parameters are used to check the quality of segmented objects. When the microscope setting is not optimal, the software may detect a large object raised from background pixels. In this case, the segmented object will be too large and will not be considered a spot. The next parameter, the number of spots, checks the validity of the segmentation by directing the results into an invalidated spots file when the number of spots is too large. The value of this parameter can be changed by the user according to the context of investigation (e.g., cell lines with nondiploid karyotype). The last parameter corresponds to the minimum distance between two colocalized spots ($dist$).

Parameters implemented in the software (SNRs, volume of each spot, number of spots, and distance between two spots close to each other) can be adjusted for each channel to take into account variations of fluorescence characteristics of each probe and their hybridization quality, post-hybridization washes, microscope parameter settings, and karyotype of cell lines. In the present work, adequate setting of these parameters was obtained by averaging values tested to give satisfactory results on more than 100 stacks: SNR values of 80 and 120, minimum spot volume of 2.5 \(\mu m^3\), minimum spot volume for two “colocalized” spots of 5 \(\mu m^3\), distance of 7 pixels between colocalized spots, and maximum spot number of 2.

Separation of two closely located spots (Fig. 2) was the most complicated process. Nevertheless, we were able to correctly split merged spots in most cases (70% to 80%), except when the spots were one on top of the other along the Z axis (<2% of cases).

The results of the distance measurements are saved in two different files. One file corresponds to segmentations that were labeled as correct and saved in a validated spots file, and the other corresponds to segmentations that were labeled as incorrect and saved in an invalidated spots file. The results are available as a text file where all the information of the spots is indicated, such as the center position of the spot, its area, and its volume. The distances between homologous and nonhomologous genes, between genes and center of nuclei, and angles between homologous genes with center of nuclei were also computed.

**Validation and Statistical Significance**

A pilot 3D-FISH study was performed to validate the above described procedures and software by determining with statistical significance the positioning of $ENL$, $AF4$, and $MLL$ genes to the center of nuclei and their intergenic distances in two lymphoblastic cell lines. Only their positioning (distance) to the center of nucleus is reported here. Intergenic distances will be reported in a future publication.

A 3D-FISH experiment was performed in NALM-6 and IM-9 leukemia lymphoblastic cell lines. These two human B-cell lines were chosen because they have near diploid, normal karyotype. Figure 3 shows an example of consecutive central slices out of a 40-image stack with different color channels corresponding to $ENL$ (red), $AF4$ (yellow), and $MLL$ (green) genes in NALM-6 cells. Perfect correlation was obtained for FISH spots (corresponding to labeled genes) and DAPI counterstaining in blue (corresponding to nucleus) on the original stacks compared with the segmented stacks. Statistical analysis, performed on more than 150 cells that corresponded to a total number larger than 450 image stacks, showed good sensitivity (probability to have a correctly segmented image in the validated spots file) and good specificity (probability to have an incorrectly segmented image in the invalided spots file) with values of 98% (427 of 435) and 99% (101 of 102), respectively. Note that the five of eight stacks that were not correctly segmented on validated spots file displayed two spots one on top of the other along the Z axis.

For all distances (gene to center of nucleus and gene to gene), comparison between 3D and 2D distances, which were computed on a 2D projection, was performed (Table 1), and more than 93% of measurements changed. In 44% of cases, measurements were underestimated by at least 10% and, in some cases, the ratio changed by more than 90%. The same comparison was performed on the angle between homologous genes and the center of the nucleus. More than 50% of angles were changed by a value greater than 10 degrees. Average distances for each cell lines were computed for 3D and 2D distances (Table 2). Results
showed significant differences between 3D and 2D gene-to-center measurements in both cell lines. Distribution of 3D distances between loci of each gene (MLL, AF4, and ENL) with the center of nucleus in NALM-6 and IM-9 cell lines was analyzed. Results are shown in Figure 4 and Table 2. Radius distribution of the AF4 and ENL genes was significantly similar in the two cell lines ($P < 0.0001$). However, a small difference in radius distribution was noted for the MLL gene. It was noted that the distribution along the nucleus radius showed that the ENL gene has a more central nucleus localization than the MLL and AF4 genes. The AF4 gene appears to be more peripheral.

**DISCUSSION**

Visual comparison between segmented image stacks and original image stacks showed no difference in position, size, and number of spots. These observations validated the segmentation procedure. The procedure of spot separation implemented in this software is quite complex but seems to give quite a satisfactory outcome. However, it strongly depends on the minimal distance between two spots ($dist$). The separation of two spots along the $Z$ axis could be implemented by projecting candidate spots onto $xy$ and $yz$ planes. However, due to the relative small number of colocalized spots in this position and the optical distortion along the $Z$ direction, this procedure was not implemented. The interest of this separation could be to study the position of all genes through the cell cycle, taking into account newly replicated spots. The comparison between 3D and 2D measurements showed an underestimation of distances in 2D condition, which could lead to erroneous interpretation. Further, thanks to real 3D measurements, exceptional events have a better probability to be detected than on 2D projections.

Our 3D-FISH experiments show that the ENL gene localized on chromosome 19 occupies the most central nuclear localization, whereas the AF4 gene (on chromosome 4) appears to be more peripheral, probably near the nuclear membrane. The MLL gene (on chromosome 11)
has an intermediate localization. These observations reflect dynamics of loci inside CTs. They are consistent with previous studies showing that chromosomes 19 in interphase nuclei are more central (28). The relative positioning of chromosomes in nuclei is organized according to their size (10): the smaller chromosomes (e.g., chromosomes 19) are more centrally localized than the larger ones (e.g., chromosome 4). No significant difference in radial distribution was observed for the AF4 and ENL genes between these two lymphoblastic cell lines. As expected, these results corroborated a relative conservation of CT organization for a given cell line (12). However, the radial distribution of the MLL gene seems to be slightly different on these two lymphoblastic cell lines. The MLL gene seems to be more centrally located in the NALM-6 than in the IM-9 cell line. NALM-6 is a human B-cell precursor in contrast to the IM-9 cell line. Cell differentiation may cause a subtle difference in nuclear organization of the MLL gene, as has been shown for other genes in the literature (29,30).

In conclusion, it seems that the image processing procedures and parameter settings implemented in the software allow satisfactory image data processing and analysis. Other 3D-FISH experiments carried out in different cell lines with additional probes are also satisfactory (data not shown). Smart 3D-FISH is new software that has been developed to automate the process of spot segmentation and perform distance measurements in 3D-FISH. It can handle virtually any number of spots and color channels for intergenic distance measurements. It can also incorporate the images from DAPI channel (total DNA of nucleus) to measure distances of genes to a nuclear center. A 3D-FISH experiment carried out on ENL, MLL, and AF4 genes in two lymphoblastic cell lines has shown self-consistent data and corroborated previously reported data in the literature concerning the organization of chromosomes inside nuclei. Visual comparison of original image stacks with segmented image stacks by Smart 3D-FISH provided satisfactory detection of all spots, with a sensitivity of 98%. The software is user friendly and robust in use. Smart 3D-FISH is available as a set of plug-ins for ImageJ software at http://www.snv.jussieu.fr/~wboudier/softs.html.

It should greatly facilitate image processing and analysis by providing a useful tool to overcome the laborious task of 3D image measurements based on user-defined parameters and decrease subjectivity in data analysis.

ACKNOWLEDGMENTS

The authors thank Maïté Coppey and Christophe Chamot at the Institut Jacques Monod (Paris, France) for providing access and assistance to confocal microscopy facility. Thanks also go to Jean Soulier (Hospital Saint Louis, Paris, France) and Mickaël Durand-Dubief for providing assistance in the FISH technique, Mariorno Rocchi (University of Bari, Italy) for providing BAC/PAC FISH clones, Laurence Lagneaux (Belgium) for providing NALM-6 cell line, and Louise Anderson for reading the manuscript.
LITERATURE CITED