Cytometry
PART A
Journal of the International Society for Analytical Cytology

Automated Scoring of Multiprobe FISH in Human Spermatozoa

Melissa J. Perry,[1]* Xing Chen,[1] Xin Lu[2]

[1]Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts

[2]Department of Family and Preventive Medicine, Department of Medicine, University of California, San Diego

Received 3 January 2007; Accepted 13 August 2007

Grant sponsor: NIEHS; Grant number: KOES10959; Grant sponsor: NIOSH; Grant number: T42/CCT110421.

*Correspondence to: Melissa J. Perry, Department of Environmental Health, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA.

Email: mperry@hsph.harvard.edu

Published online 2 October 2007 in Wiley InterScience (www.interscience.wiley.com)

DOI: 10.1002/cyto.a.20468

Abstract

In the case of chromosomal aneuploidy in sperm wherein the incident rate is low and a large number of cells require scoring, automated methods that rely on computer software to segment and to count fluorescence signals are particularly necessary due to countless hours spent in reading slides and to the potential for interoperator differences. The purpose of this pilot experiment was to determine whether there were significant differences in the estimates of disomy frequency produced by automated versus manual scoring of signals for chromosome X, Y, and 18 in human sperm. The frequency of X18, Y18, XX18, YY18, and XY18 were determined in four separate normozoospermic samples. Slides were hybridized using a standard sperm FISH protocol for centromere-specific probes. Between 500 and 564, DAPI positive nuclei were captured from each sample and scored using the automated system, and the same slides were scored by a trained cytogeneticist, who was blind to the purpose of the study and the automated system results. None of the estimated frequencies was significantly different between manual and automated methods, regardless of whether individual slides or pooled results across all samples were compared. To our knowledge, this is the first report examining the validity of automated cell scoring in human spermatozoa. The results from this pilot exploration of sperm FISH suggest the comparability between automated and manual methods for estimating sex chromosome disomy and provide evidence that automated laser scanning of multiprobe sperm FISH should be explored further.

Key terms

fluorescence in situ hybridization; chromosomal aneuploidy

THE utility of fluorescence in situ hybridization (FISH) to identify chromosome-related events has been demonstrated in a variety of cell types over the past two decades (1). Progressive improvement of visualization since the early 1980s included first the indirect labeling of antibodies onto which stains and then fluorochromes were attached, followed by direct labeling of DNA probes. The need for automated scoring arises due to the countless hours spent in reading slides and to the potential for interoperator differences. Automated scoring or the use of computer software to segment and to count fluorescent signals holds particular promise for advancing the studies of human sperm aneuploidy wherein scoring 10^4 cells per person is often the target due to the rareness of numerical aberrations. Manual scoring of 10^4 cells can take up to 20 h whereas using an automated system can take 1–2 h including review of noninclusive events. Automated spot counting can solve the problem of fatigue due to many hours of microscope viewing, which can lead to misinterpretation and scoring inconsistencies. Statistical probability studies suggest automated scoring is the only reasonable way to reliably detect sporadic or low rate aberrations with small error rates (2–4).

Automated Scoring of Chromosome Abnormalities in Nonsperm Cells

Application of automated systems for scoring numerical and structural chromosome aberrations in different human cell types has advanced in recent years,

ISAC
International Society for Analytical Cytology

particularly for use in uncultured amniocytes (5), peripheral lymphocytes (6–9), and fibroblasts (9). Most of these systems have in common a controllable motorized microscope with an automated stage, image segmentation and acquisition software, and event counting software. Schunck et al. (8) reported that reliable and reproducible results were found for automated analysis using the Metafer™ system to score dicentric chromosomes, micronuclei, and Comet assay results in peripheral lymphocytes. In considering these reliable results for examining structural chromosome aberrations, authors pointed out the promise of automated scoring for fluorescence spot counting in detecting numerical chromosome aberrations. Lev et al. (5) examined numerical chromosome abnormalities using the Duet™ (BioView, MA) automated system in comparison to manual signal counting. When FISH was performed on 56 amniotic fluid samples and cells were examined for aneuploidy in chromosomes 21, 18, 13, X and Y, automated and manual results showed a high correlation ($r = 0.9$, $P < 0.0001$).

Other molecular applications have demonstrated the successful use of automated methods for distance measurements in three-dimensional nuclear structures (10), and for measuring telomere length in proliferating cells (11,12). However, none of these methods has yet been widely adapted for studying numerical chromosomal aberrations in human sperm.

Computerized Scoring of Sperm Cells

Baumgartner et al. (13) has published one of the only sperm FISH studies using automated scoring to detect numerical chromosome aberrations. They examined disomy frequency in chromosomes 8 and X and reported that the frequencies of genotypically abnormal sperm were not significantly different from the automated results using laser scanning cytometry (LSC) and manual scoring. LSC provides data equivalent to flow cytometry, but instead of passing the laser beam through a volume flow, cells are analyzed on slides. Because the position of scanned cells and their fluorescence signals are recorded electronically in the automated system, cells can be easily relocated. Study authors concluded that the rapid LSC reduced time-consuming manual scoring of fluorescence spots among 10,000 sperm per slide from an average of 8 h to ~30 min. However, their LSC was only able to detect one color fluorochrome at a time, representing a single chromosome; their report evaluated disomy in chromosome 8 and X. In this pilot study, we sought to expand the application of automated fluorescence signal detection using a system capable of scoring three fluorochromes per slide. Of particular concern for this initial study was whether computerized methods that relied on the automated segmentation of fluorescence signals would be comparable to manual results in scoring three-probe FISH for chromosomes X, Y, and 18 in human sperm.

Materials and Methods

Sperm Slide Preparation for FISH

We used blinded normozoospermic (concentration $\geq 55 \times 10^6$ spermatozoa per milliter) samples from a hospital-based fertility clinic to compare manual and automated scoring using LSC. We used X, Y, and 18 centromere-specific probes for determining X, Y, and XY disomy, and chromosome 18 was used as an autosome control. The hybridization was performed using a combination of probes provided for this experiment by Vysis (Downers Grove, IL). 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).

The semen sample was diluted with T/S buffer (14) of pH 8 and the concentration determined. Approximately 10×10^6 cells were used in the initial washing steps. After 3 wash cycles of centrifugation and resuspension, the sperm cells were finally resuspended in 1 ml of T/S buffer (14) and the concentration determined again. Sample (10 μl) was horizontally spread on a glass slide, allowed to dry, and then placed in a −20°C freezer until the hybridization process was started.

Hybridization

The cells were swollen in a cold DTT buffer for 30 min on ice followed by incubation in LIS buffer (10 mM diiodosalicylic acid-lithium salt (LIS), 50 mM Tris–HCl, pH 8.0) for 1 h at room temperature in the dark. During the LIS swelling, the slides became light-sensitive, and all further treatments took place in the dark. To denature the DNA, slides were placed in a 70% formamide/2× SSC solution, warmed to 73°C, for 5 min, and then suspended in a series of ethanol (70, 85, and then 100%) baths for 1 min each at room temperature. Orange (Y), aqua (18), and green (X) probes were mixed according to Vysis protocol. Briefly, 1 μl of each probe was added to 7 μl of CEP hybridization buffer and 1 μl of each of the probes for every slide being treated. The mixture was both pulsed on the centrifuge and vortexed three times then denatured in the 73°C water bath for 5 min directly before application to the slides. Concurrent with the probes being heated, the slides were placed on a slide-warmer at 47°C for 2 min to evaporate any remaining ethanol. Ten microliters of the probe mixture were applied immediately to the dense area of cells on the slide. This area was then fitted with a coverslip and sealed with rubber cement to prevent evaporation of the probe mixture. The cells were allowed to hybridize overnight at 42°C in a light-resistant, humidified box. The next day, the coverslips were removed, and the slides were subject to three washes with 50% formamide/2× SSC for 10 min each, followed by one wash with pure 2× SSC for 10 min, and finally one wash with 2× SSC/0.1% NP-40 for 5 min. All washes took place in the dark at 46°C. After air drying, 10 μl 6-diamino phenylindole dihydrochloride (DAPI) II counterstain was applied to the target area of the slide. A coverslip was placed over the area and sealed with clear nail polish. The slide was placed in the freezer at −20°C for at least 30 min prior to viewing under the microscope.

Slide Scoring

Data were captured to make comparisons between automated and manual results. The experiment was designed to

Table 1. Comparison of Normal and X and Y Disomy Frequencies in Sperm Nuclei Using Manual Versus Automated Scoring Methods of Triple-Probe FISH

			X18				Y18				XX18				YY18				XY18			
SLIDE	METH	*N*	*N*	%	*p*	*P*	*N*	%	*p*	*P*	*N*	%	*p*	*P*	*N*	%	*p*	*P*	*N*	%	*p*	*P*
1	auto	500	271	54.20	0.95	0.43	223	44.60	0.85	0.67	0	0.00	1.00	0.77	0	0.00	1.00	0.57	3	0.60	0.25	1.00
	manual	500	273	54.60			227	45.40			0	0.00			0	0.00			0	0.00		
2	auto	510	231	45.29	0.67		225	44.12	0.81		1	0.20	1.00		0	0.00	1.00		2	0.39	0.50	
	manual	505	221	43.76			218	43.17			0	0.00			0	0.00			0	0.00		
3	auto	561	165	29.41	0.06		208	37.08	0.33		3	0.53	0.72		15	2.67	0.55		4	0.71	1.00	
	manual	561	196	34.94			225	40.11			5	0.89			11	1.96			5	0.89		
4	auto	561	290	51.69	0.98		247	44.03	0.98		1	0.18	1.00		1	0.18	1.00		4	0.71	0.55	
	manual	564	293	51.95			247	43.79			2	0.35			1	0.18			7	1.24		

Meth, automated or manual scoring; *p*, *p* value for comparison on each slide; *P*, *P* value for pooling all four slides together.

capture and to score at least 500 DAPI positive nuclei using the automated system, and then to score the same 500 or more cells by a trained cytogeneticist, who was blind to the purpose of the study and the automated system results. Strict scoring criteria as described by Baumgartner et al. (15) were followed, and the automated classifier was programmed to operationalize these criteria. Nuclei needed to be intact with a clearly defined border. Nuclei that were not scored: disrupted nuclei with indistinct margins, overlapped nuclei, or sperm heads without well-defined boundaries, very large nuclei with diffused chromatin, and very small nuclei with no signals. Nuclei that were scored: only cells consistent with sperm size and shape or with a visible tail, clear hybridization signals, similar in size, separated from each other by at least one signal domain, and clearly positioned within the sperm head. One domain was considered to be the diameter of the signal, which is generally round. Nuclei were regarded as abnormal if they presented: two (or more) distinct hybridization signals for the same chromosome, each signal was equal in intensity to the single signal found in normal monosomic nuclei, each signal was equal in size to the single signal found in normal monosomic nuclei. In the case of disomy, all signals of the same color needed to have the same intensity and be separated from each other by a distance of at least one domain (the diameter of each signal). Disomy was scored when the two fluorescence domains of the same color occurred within the sperm head, comparable in brightness and size.

The instrumentation used in this study was similar to the system optimized for Vysis centromere-specific probe imaging, reported by Netten et al. (16), and included the use of a Zeiss Axioscope Microscope with five filters including a dual band-pass excitation filter, single excitation filters, and a dual band-pass emission filter. The automated laser cytometery system proceeds by (a) autofocusing the microscope on every field of view (segmentation) by moving the stage first in the X- and Y-axis and then into the Z-direction, capturing images in different focus planes and analyzing the focus quality based on local contrast criteria. During the subsequent scan, the slide is automatically kept within the plane of best focus, (b) acquiring the counterstained and FISH labeled images, (c) analyzing the images to define the position of the nuclei and the number of FISH signals inside of them, and (d) performing all the stage movements and filter changes required to scan the area. The scoring of nuclei by the image classifier was operationalized using the guidelines detailed earlier, and nuclei were selected based on area, contour ratio, sphericity, and eccentricity (16).

Data Analysis

Patterns of X, Y, and XY disomy were examined, and frequencies from the automated system and manual system from the same 500 or more cells were compared. Because the coordinates for scanned cells positions are recorded by the automated system, the same cells can be relocated easily and scored manually. A slide grid was used to locate the same five separate locations on the slide, from which 100 cells each were captured and scored. Four separate slides (one sample per slide), each with at least 98% hybridization efficiency were compared, and a total of 2,127 cells were scored using both systems. To compare the event frequencies using each method, statistical testing was performed using two sample tests of frequency difference with Pearson's chi-square or Fisher's exact tests when no more than five cells were encountered in a category. We calculated *p* values for comparing the difference between manual and automated methods on each slides, as well as a *P* value for pooling all four slides together, within disomy conditions.

Results

Table 1 presents the results comparing the automated and manual methods in scoring the same 500 or more cells. The percentage estimated by different methods and from different slides varied. X18 and Y18 frequencies ranged from 29.41 (auto) to 54.60 (manual), X disomy frequencies ranged from 0.00 (manual) to 0.89 (manual), Y disomy ranged from 0.00 (both auto and manual) to 2.67 (auto), and XY disomy ranged from 0.00 (manual) to 0.89 (manual). However, none of the estimated frequencies was significantly different between manual and automated methods, regardless of whether individual slides or pooled results across all four samples were compared. Differences in frequencies of X18 approached borderline significance for sample 3 (auto: 29.4 vs. manual: 34.9; $P = 0.06$); however, this difference did not approach significance when

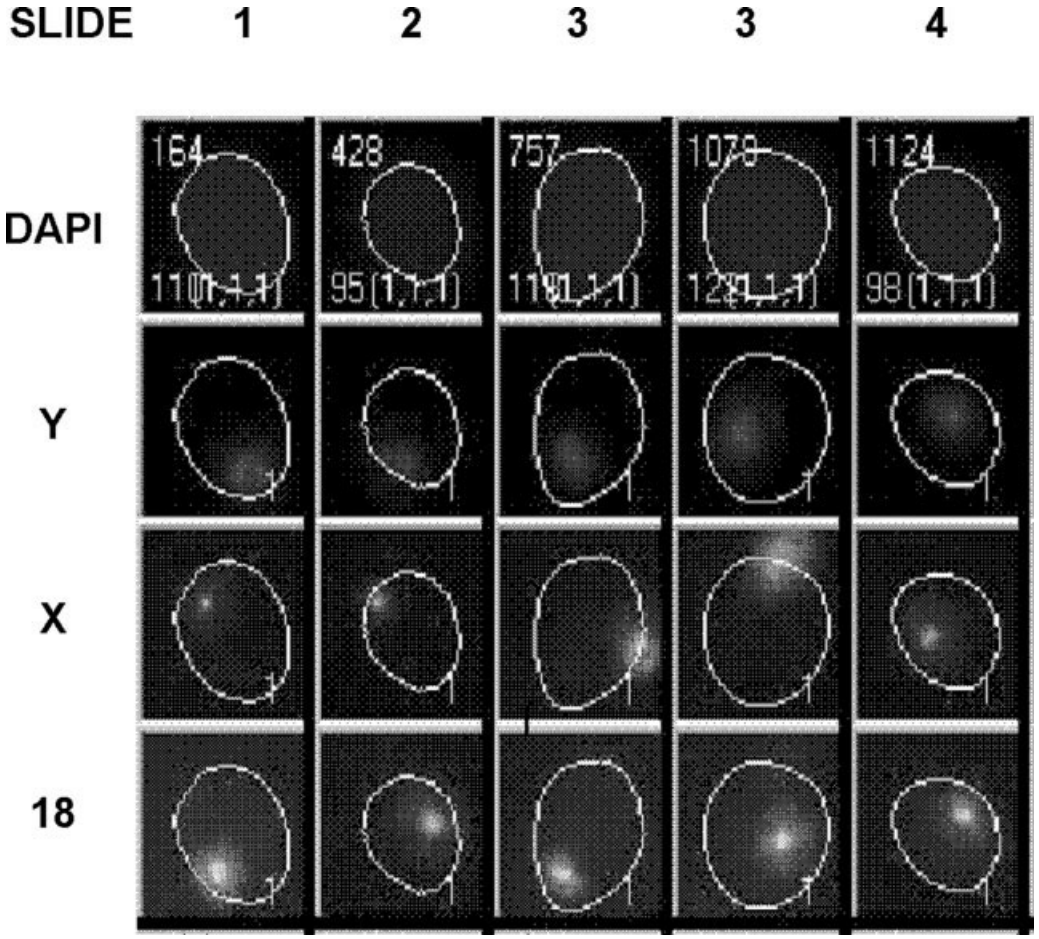

Figure 1. Five instances of XY18 disomy detected in sperm nuclei from four samples using automated image segmenting software, three fluorescent centromere-specific probes, and DAPI for nuclei staining. Signals were seen in the X, Y, and 18 chromosomes in the same nuclei.

the X18 results were pooled across all four samples (pooled $P = 0.43$). Image 1 depicts five instances of XY18 disomy in each of the four slides, including two examples from slide 3 (Fig. 1).

Discussion

The results from this initial exploration of automated scoring of sperm FISH provide evidence that the automated laser scanning of multiprobe sperm FISH should be explored further. However, these results should be interpreted with caution due to the small number of cells scored. In one sample, the difference between X18 frequencies for manual and automated approached borderline significance with a difference of almost 5%. The promise of this technology lies in the ability to conduct large sample studies investigating thousands of sperm nuclei per sample, and further quantification of this type of variability needs to be explored in larger investigations. Accurate estimates of sperm disomy are hard to achieve due to their rarity. Results did not show evidence of major discrepancies in automated and manual estimates for disomy in either the individual or the pooled samples. Prior reviews of sperm FISH results in normal men have reported the following ranges (%): X disomy: 0.03–0.37; Y disomy: 0.04–0.21; XY: 0.06–0.42 (17). Another review recently reported that the percent total disomy for the sex chromosomes averaged across 23 studies of normal men using multicolor FISH was 0.26 (18). The disomy frequency estimates from this study were higher than these previously reported ranges. However, considerable differences in disomy frequency for specific chromosomes have been identified (19) and are potentially attributed to interdonor heterogeneity and/or different methodologies used among laboratories (primarily the type of DNA probes used, the scoring criteria applied, or the number of sperm analyzed) (18). For example, a recent report of over 70,000 sperm scored in healthy men reported higher disomy ranges per 10,000 sperm than reported previously: (%)-X disomy: $0.4 \pm 0.5 - 3.1 \pm 1.7$; Y disomy: $0.7 \pm 1.1 - 9.8 \pm 4.7$; XY: $0.7 \pm 0.8 - 14.1 \pm 1.3$ (20).

Automated cell imaging, segmentation, and scoring systems with multichannel capabilities have been developed for studying chromosomal aberrations using FISH in normal lym-

phocytes (6–9) and amniocytes (5) and in cancer cells (21), and for determining telomere length (11,12). However, application to FISH in human spermatozoa has posed a unique set of challenges. Human sperm nuclei are small and highly condensed in comparison to other cell types and therefore sample preparation and spread is critical to achieve adequate decondensation, hybridization, and uniform distribution of nuclei for automated image segmentation. Nuclei that are overlapping, with indistinct margins, abnormally large or small, or with weak fluorescent signals cannot be scored manually or by the automated system and these noninclusive events can result in considerable cell loss. Capturing a large number of nuclei to accurately estimate disomy or aneuploidy is therefore difficult if the slide preparation and hybridization protocol do not uniformly achieve these criteria. Therefore, optimizing automatic scoring of sperm nuclei has the technical demand of proper programming of the image classifier to adhere to strict-scoring criteria, the procedural demand of uniform sample preparation, decondensation, and hybridization, and the high throughput demand of maximizing the number of scored nuclei.

The results from this pilot study support further optimization and testing in more samples using larger numbers of sperm cells, so that the true promise of automated scoring, mainly efficient and accurate high throughput analysis in research studies can be realized. We are currently conducting a larger validation study comparing automated versus manual estimates from 10 separate sperm samples of 10^4 cells each, the results of which should be available in the coming year.

Literature Cited

1. Levsky JM, Singer RH. Fluorescence in situ hybridization: Past, present and future. J Cell Sci 2003;116 (Part 14):2833–2838.
2. Carothers AD. Counting, measuring, and mapping in FISH-labelled cells: Sample size considerations and implications for automation. Cytometry 1994;16:298–304.
3. Castleman KR, White BS. Dot count proportion estimation in FISH specimens. Bioimaging 1995;3:88–93.
4. Kibbelaar RE, Kok F, Dreef EJ, Kleiverda JK, Cornelisse CJ, Raap AK, Kluin PM. Statistical methods in interphase cytogenetics: An experimental approach. Cytometry 1993;14:716–724.
5. Lev D, Daniely M, Zudik A, Preisler E, Hoffmann N, Kaplan T, Raz U, Yanoov-Sharav M, Vinkler H, Malinger G. Automatic scanning of interphase FISH for prenatal diagnosis in uncultured amniocytes. Genet Test 2005;9:41–47.
6. Bielorai B, Golan H, Trakhtenbrot L, Reichart M, Toren A, Daniely M, Zilberstein Y, Amariglio N, Rechavi G, Kaplinsky C. Combined analysis of morphology and fluorescence in situ hybridization in follow-up of minimal residual disease in a child with Philadelphia-positive acute lymphoblastic leukemia. Cancer Genet Cytogenet 2002; 138:64–68.
7. Huber R, Kulka U, Lorch T, Braselmann H, Engert D, Figel M, Bauchinger M. Technical report: Application of the Matafer2 fluorescence scanning system of the analysis of radiation-induced chromosone aberrations measured by FISH-chromosone painting. Mutat Res 2001;492:51–57.
8. Schunck C, Johannes T, Varga D, Lorch T, Plesch A. New developments in automated cytogenetic imaging: Unattended scoring of dicentric chromosomes, micronuclei, single cell gel electrophoresis, and fluorescence signals. Cytogenet Genome Res 2004;104:383–389.
9. Truong K, Gibaud A, Dupont JM, Guilly MN, Soussaline F, Dutrillaux B, Malfoy B. Rapid prenatal diagnosis of Down syndrome using quantitative fluorescence in situ hybridization on interphase nuclei. Prenat Diagn 2003;23:146–151.
10. Gue M, Messaoudi C, Sun JS, Boudier T. Smart 3D-FISH: Automation of distance analysis in nuclei of interphase cells by image processing. Cytometry A 2005;67A:18–26.
11. Narath R, Lorch T, Greulich-Bode KM, Boukamp P, Ambros PF Automatic telomere length measurements in interphase nuclei by IQ-FISH. Cytometry A 2005;68A:113–120.
12. Potter AJ, Wener MH. Flow cytometric analysis of fluorescence in situ hybridization with dye dilution and DNA staining (flow-FISH-DDD) to determine telomere length dynamics in proliferating cells. Cytometry A 2005;68A:53–58.
13. Baumgartner A, Schmid TE, Maerz HK, Adler ID, Tarnok A, Nuesse M. Automated evaluation of frequencies of aneuploid sperm by laser-scanning cytometry (LSC). Cytometry 2001;44:156–160.
14. Wyrobek A, Alhborn T, Balhorn R, Stanker L, Pinkel D. Fluorescence in situ hybridization to Y chromosomes in decondensed human sperm nuclei. Mol Reprod Dev 1990;27:200–208.
15. Baumgartner A, Van Hummelen P, Lowe XR, Adler ID, Wyrobek AJ. Numerical and structural chromosomal abnormalities detected in human sperm with a combination of multicolor FISH assays. Environ Mol Mutagen 1999;33:49–58.
16. Netten H, Young IT, van Vliet LJ, Tanke HJ, Vroljik H, Sloos WC. FISH and chips: Automation of fluorescent dot counting in interphase cell nuclei. Cytometry 1997;28:1–10.
17. Egozcue J, Blanco J, Vidal F. Chromosome studies in human sperm nuclei using fluorescence in-situ hybridization (FISH). Hum Reprod Update 1997;3:441–452.
18. Templado C, Bosch M, Benet J Frequency and distribution of chromosome abnormalities in human spermatozoa. Cytogenet Genome Res 2005;111:199–205.
19. Moore DH II, Epstein L, Reeder J, Wheeless L, Waldman FM. Interlaboratory variability in fluorescence in situ hybridization analysis. The NCI Bladder Tumor Marker Network. Cytometry 1996;25:125–132.
20. Rubes J, Vozdova M, Robbins WA, Rezacova O, Perreault SD, Wyrobek AJ. Stable variants of sperm aneuploidy among healthy men show associations between germinal and somatic aneuploidy. Am J Hum Genet 2002;70:1507–1519.
21. Wang YP, Castleman KR. Normalization of multicolor fluorescence in situ hybridization (M-FISH) images for improving color karyotyping. Cytometry A 2005;64A:101–109.