Journal of Cellular Biochemistry 97:1151–1159 (2006)

Gene Expression in the Cell Cycle of Human T-Lymphocytes: II. Experimental Determination by DNASER Technology

Claudio Nicolini,[1,2]* Rosanna Spera,[1] Enrico Stura,[2] Stefano Fiordoro,[2] and Luca Giacomelli[1]

[1]Nanoworld Institute and Biophysics Division, University of Genova, Corso Europa 30, 16132 Genova (Italy), Genova Section of the INBB (National Institute Biostructures), Genova, Italy
[2]Fondazione Elba, Via delle Testuggini, 00100 Rome (Italy)

Abstract Human lymphocytes gene expression before and after PHA stimulation is monitored by DNASER technology, a novel bioinstrumentation entirely constructed in our laboratories as previously reported. The validity of the DNASER measurements is confirmed by standard fluorescence microscopy equipped with CCD. The human lymphocytes gene expression here experimentally probed using commercially available DNA microarrays such as Human Starter, appears compatible both with independent bioinformatic prediction and with existing experimental data, pointing to MYC as the key gene in the G_0–G_1 transition induced by PHA in resting lymphocytes. It does not escape our notice that in cell biology and cancer research DNASER technology based on microarray constructed with few leader genes identified from bioinformatics represents a meaningful cost-effective route alternative to massive frequently misleading molecular genomics. J. Cell. Biochem. 97: 1151–1159, 2006. © 2005 Wiley-Liss, Inc.

Key words: lymphocytes; PHA stimulation; DNASER; gene microarray

The key question in cell biology is the effect of changes in the overall state of the cell such as the phase of the cell cycle on gene expression and its regulation. In this study, we perform such a search to identify genes involved in the control of the cell cycle of human T-cell lymphocytes. Human T-lymphocytes constitute an ad hoc model due to the fact that their progression through the cell cycle is easily initiated by activation [Oosterwegel et al., 1999; Cantrell, 2002; Isakov and Altman, 2002] and, in particular, was quantitatively characterized time ago [Abraham et al., 1980]. Experimental investigations in this area used genome-wide measurements of gene expression levels with DNA microarrays from which it is possible to infer data on interactions between the genes and the resulting proteins [Butte, 2002]. To map such interactions directly from microarray experiments, researchers involved in the area compose sophisticated software [Jones et al., 2004, Troyanskaya, 2005]. This resulted in accumulation of immense amount of data on gene and protein interaction (about 800,000 interacting gene and protein pairs are currently known). Based on the existing databases built-up recently cluster analysis (Sivozhelezov et al., 2005) was recently used to determine the most important genes called "leader genes" in order to compile and to update maps of the major biological control systems, and to integrate them in a concise manner to discern common patterns of interactions between gene expression and their correlated coding of proteins during cell cycle progression.

In this communication, we probe experimentally gene expression during human lymphocyte cell cycle progression using commercially available DNA microarray in order both to test if a newly developed and simple technology called DNASER [Nicolini et al., 2002] is appropriate for monitoring gene expression during the human lymphocytes cell cycle and to have an

Grant sponsor: Organic Nanoscience and Technology (FIRB Grant); Grant sponsor: FISR–MIUR Grant.

*Correspondence to: Claudio Nicolini, Nanoworld Institute and Biophysics Division, University of Genova, Corso Europa 30, 16132 Genova, Italy.
E-mail: manuscript@ibf.unige.it

Received 28 July 2005; Accepted 20 September 2005
DOI 10.1002/jcb.20694

© 2005 Wiley-Liss, Inc.

independent validation of the accuracy of this bioinformatic prediction. The DNASER represents a novel bioinstrumentation for real-time acquisition and elaboration of images from fluorescent DNA microarrays developed in our laboratories [Nicolini et al., 2002; Troitsky et al., 2002].

The lymphocytes are indeed cells of fundamental importance in the immune system and have a role in the antigen recognition, in the antibody production and in the cytotoxic response to tumor cells. A global approach to the study of the gene expression, as the one supplied from DNA "microarray," can thereby reveal the gene expression profiles involved in the cell cycle regulation of lymphocytes.

MATERIALS AND METHODS

All reagents and materials were purchased from Sigma Aldrich (St.Louis, MO) if not otherwise specified.

Cell Systems

To obtain human lymphocytes a sample of heparinized peripheral blood specimens were diluted 1:1 with sterile phosphate buffer saline (PBS, pH 7.2). The sample was centrifuged on a Ficoll–Hypaque gradient (specific gravity 1.080) at 1,500 rpm for 40 min at room temperature without brake [Abraham et al., 1980]. The lymphocytes rich interface was collected and washed twice with PBS, and the number of cells counted in presence of trypan blue. The 2×10^6 cells were seeded into each well containing 2 ml of RPMI 1,640 culture medium supplemented with 10% heat-inactivated fetal bovine serum, penicillin 100 U/ml, and streptomicin 50 μg/ml

The cells have then been incubated at 37°C in a humid chamber with 5% of CO_2 for 24, 48, 72 h. The cells have been then collected and counted and then it has been proceeded to RNA extraction.

RNA Extraction

In order to obtain total RNA extraction the cells pellet (minimal 1×10^7 cells) has been dealt with an extraction kit from Amersham Biosciences containing LiCl, CsTFA, and an extraction buffer. The samples thus obtained have been conserved at −80°C. For the estimation of the extracted RNA they have been used 200 μl RNAsi free cuvettes and the following formula has been used:

$$[\mathrm{RNA}] = \mathrm{A260} \times \mathrm{D} \times 40\ \mu\mathrm{g/ml}.$$

In order to avoid contaminations of genomic DNA, the RNA samples have been subordinates to digestion with the enzyme DNase I.

cDNA Syntesis and Microarray Hybridization

The CyScriptRT enzyme, together to 25 μg of RNA (for each sample) and to dCTP Cy3 and dCTP Cy5 nucleotides, has been employed in a synthesis reaction lead to 42°C for 1 h and ½ in order to obtain the labeling by retrotranscription. The synthesis product of this reaction, marked cDNA, has been then subordinate to a purification step. Total RNA was extracted and amplified using T-7 in vitro transcription.

The cDNA marked samples have been purified employing a purification kit and chromatographic columns supplied by Amersham Biosciences. The cDNA obtained has been precipitated and resuspended in bidistilled water to quantify, by the employing of the spectrophotometer, the samples and to verify the labeling. The cDNA marked samples have been subsequently lyophilized and resuspended in an opportune volume (120 μl) of hybridization buffer (Salt-Based ibridation).

For the array hybridization 1 μg of cDNA marked with Cy3 and 1 μg of cDNA marked with Cy5 they have been mixed in a tube and resuspended in the hybridization buffer together to the control sample (Arabidopsis control). The sample thus obtained has been denatured and spotted on the array. The matrices have been then put in a hybridization chamber at 42°C for 20 h. To eliminate the aspecific binding the array has been washed with SSC buffers of decreasing concentration.

Human Starter Array

The employed array is the Human Starter Array by MWG Biotech, choose on the base of the gene we are interested about. It contains 161 oligonucleotides (designed to be specific for the respective human gene sequence), 32 replicas, and 7 gene specific Arabidopsis control oligonucleotides, for 200 total spots, disposed in 10 columns and 20 rows, more an exact copy, in total 400 spots for array. The diameter of one spot is 100 μm and the distance between two near spots is 250 μm.

DNASER Apparatus

The DNASER, thanks to the particular opto-mechanical structure joined to the high-sensitivity CCD camera, generates wide area images in a single shot using samples with very low fluorescence intensity. That is possible because the sample is illuminated with a beam produced by a white light source. The beam, through a specific optical path, is accurately filtered and focused on the sample and then on the CCD camera sensitive area. A supervisor PC controls DNA microarray images acquisition and elaboration.

The illumination system, the optical filter system and the motorized holder blocks constitute the optomechanical components. The light source, a 150-W Xenon lamp, generates a collimated beam through an elliptical reflector. Two filters eliminate IR and UV components. In this way, one avoids excessive heating and ultimately damage, of the sample. Two interchangeable interference filters are placed between two achromatic doublets so as to work in parallel illumination conditions. These two filters are different because are tuned, respectively on the excitation and emission frequency of the DNA microarray fluorescent spots.

The controlled system for the acquisition of microarray "DNASER" [Nicolini et al., 2002] is defined in three logical blocks:

- Hardware part
- Software part
- Optical part

The DNASER hardware part is based on a three-dimensional controlling system of a mover, a metal structure intended to be used as sample holder, with three stepper motors connected to a screw system allowing the translational movement on all of the three axes for the sample to be explored. This system is controlled by an electrical alimentation circuit for the two phases of each single stepper motor in both the directions for the current flow, receiving pulses from the Centronics interface (controlled by the parallel port of the PC) and modulating the current intensity to have the correct value of the magnetic field in the stator of the stepper motors, by tuning the voltage on the base of four Darlington type transistors presents for each axis (12 transistor in total). The "out of range" automatic stop is controlled by a counter allowing to move only up to 255 steps for each direction for each axis.

The DNASER software is required to drive both the camera for the image acquisition and the movement circuits for the sample holder positioning. For the image acquisition, it is based on the HiPic application provided with the ORCA II camera, using both the serial port of the PC and a dedicated PCI board designed to acquire information in "parallel task" mode on different regions of the CCD. HiPic allows translating in a reduced range of video frequencies the information obtained directly from the CCD, using a proprietary control called LUT. This function is very useful working with "weak" samples. To control the stepper motors of the mover, a custom module of LabView (National Instruments) is used, controlling physically the serial port of the PC (for this reason Windows 95 or 98 are needed). The input data to the spot analysis are the raw DNA microarray images. The output data of this processing are all the spot features.

- Brightness features (spot foreground and background intensity).
- Geometric features (spot foreground and background area, circularity).

The optical part of the DNAser is based on an illumination system, a filter set, the reflection surface and the acquisition device, to acquire images of the sample marked with particular fluorophores (Cy3 and Cy5). The optical source is a xenon (or mercury) lamp, characterized by an emission spectrum covering all the visible, limited in the IR/UV zone by specific filters. The filtered light reaches the sample surface and then is reflected to the sensible surface of the ORCA II (Hamamatsu) camera, the CCD, crossing a specific filter for the frequency of interest for the experiment.

Fluorescence Microscopy

To test the results obtained by DNASER technology the microarray was analyzed by fluorescence microscopy; the analysis was performed with an Axioplan Zeiss microscope equipped with an objective of magnification 2× (in order to acquire the greater possible area) and Zeiss filter sets no. 15 and 26, chosen to visualize the Cy3 and Cy5 dye, respectively. Digital pictures were obtained by an air-cooled scientific grade CCD camera Zeiss, with a dynamic range of 14 its (grey levels from 0 to 16,383). The obtained images were analyzed by GenePix software.

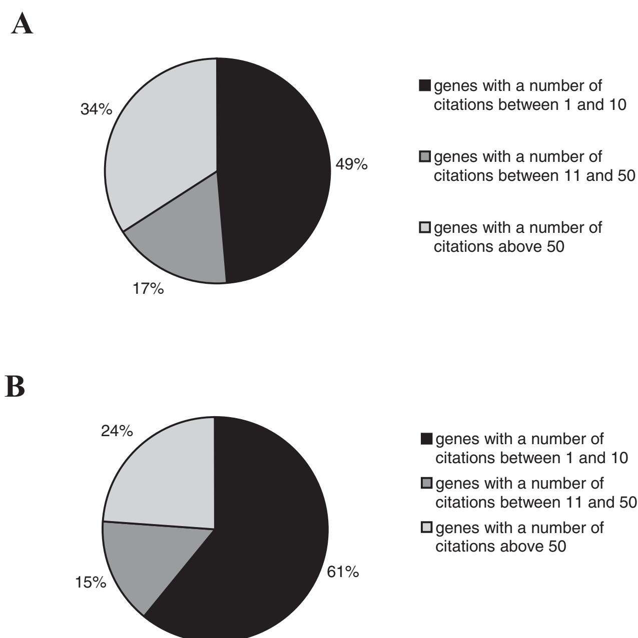

Fig. 1. **A**: Citations in PubMed of human lymphocytes genes. **B**: Citations in PubMed of human lymphocytes cell cycle genes.

RESULTS

Significance of the Used Arrays

The Human Starter Array by MWG Biotech here utilized to analyze the changes in the genetic expression of lymphocytes after PHA stimulation is confirmed by a search in the PubMed database to be quite significant in probing human cell cycle (Fig. 1 and Tables I and II).

For each gene, we performed a research in the PubMed database, in order to identify its importance in the cellular system we used, human T-lymphocytes. We evaluated the importance of every gene as the number of citations reported in the database. The search in PubMed was automated with the preparation of an interrogation script. In this way, it is possible to perform complex searches in a relatively short time. Moreover, a manual search on a very large number of genes would lead to a high error frequency. Input of the search is a text file, containing every gene of interest. The search was performed using the following queries: genename and (human or sapiens) and lymphocytes; genename and (human or sapiens) and lymphocytes and (cell cycle or (G_1 or G_2 or G_0 or S or M) and phase).

Among the 161 genes displayed on the Human Starter Array, 76 (47.2%) have citations in PubMed in relation to the study of human lymphocytes. We divided genes into four classes, according to their number of citations and thus we found that 26 genes showed great importance in this field (number of citations >50). Then, we realized a deeper query in order to identify the genes studied in relation to the human lymphocytes cell cycle. We found that 46 genes out of 161 (28.5%) satisfied this query, and that 11 genes can be considered of great importance.

Extraction of the Total RNA and Labeling

For every experiment we obtained a good RNA total extraction yield, as summarized in

TABLE I. Number of Genes Displayed on Human Starter Array

Human starter array		
		[%]
Total number	161	
Studied concerning human lymphocytes	76	47.2
Studied concerning human lymphocytes and cell cycle	46	28.5
Studied concerning human lymphocytes and microarrays	23	14.2

TABLE II. Genes Displayed on the Human Starter Array Dealing With Cell Cycle in Human Lymphocytes and Studied With Microarray Technology

Gene	Number of citations
Myc	10
Tnfrsf6	6
Adprt	3
Cdk4, Cxcr4, Pcna, Ubiquitin	2
Alpha tubulin, Annexin a2, Ask, Atf4, Cd14, Cd44, Cdc2, Fos, Gapd, Hsp70, Ik, Ldha, Mig, Mpo, Smarca4′ Tra1	1

the below tables. Moreover the RNA spectrophotometrical analysis has evidenced a high purity degree, being the ratio 260/280 nm always more then 1.9 (Table III).

DNASER Versus Fluorescence Microscopy Image Acquisition

DNASER acquires the image containing the DNA microarray (repeated in two copies to have a useful redundancy of information about the hybridization) in two different conditions:

I) Single shot
II) Composite shots

The first solution gives the possibility to obtain an image containing the whole array, but with a low number of pixels for each spot. The second solution allows the acquisition of more than one image with different regions of the same microarray, assuring a high number of pixels for each spot, but with the difficulty of assembling the final image using the partial ones.

The surface of the sample acquired in single shot mode is 9×9 mm^2, with a lower optical efficiency in the outer regions of this area (15%–20% calculated in mean value over a 3×3 pixel matrix taken from the outer region of the acquired image) with respect to the central region, and with a contrast ratio also lower in these outer regions.

Using the composite shots mode with four different images having an effective area of 6.5×6.5 mm^2, the optical efficiency in the outer region reduces of only 10% with respect to the central region. Images acquired in composite shots mode require digital filtering to be connected in an assembled image, to avoid perimetral variations that could introduce noise in the signal analysis. For this reason an algorithm for the equalization of the luminance and the reduction of the video noise was designed. This algorithm considers a threshold value for contrast and brightness discriminating the valid information from the signal coming from impurities present over the glass (dust, water, etc.). In this way the "point like" disturbs are removed from the image.

Starting from the raw images the equalization is performed by mean of the predefined filter (not shown) increasing the brightness in the outer regions so that the mean brightness in every sufficiently wide region is similar in value. It is now possible to compose the images, and pass them to the elaboration system dedicated to the extrapolation of numeric information and the conversion in tabular format.

A meaningful spot definition algorithm, taking advantage of freeware routines and library for the bitmap (bmp) importation, was written and used to obtain the brightness table, setting a threshold value (th) over which the spots are validated.

Based on the above image acquisition of a region of Human Starter Array hybridized

TABLE III. RNA Extraction Yield for the Different Lymphocytes Samples

Time, in hours, after PHA stimulation	Number of cells	% of quiescent cells (in $G_0 + Q$)	% of proliferant cells (in G_1,S,G_2)	Total RNA pg/cell	Fluorochromes utilized
0	10^7	90	8,1,1	30	Ethidium bromide
0	5×10^6	90	8,1,1	24,6	cy3 green
48	5×10^6	36	24,37,3	39,4 (+60 %)	cy5 red
0	5×10^6	—	—	24,6	cy5 red
48	5×10^6	36	—	39,4 (+60 %)	cy3 green
0	5×10^6	90	8,1,1	22,4	cy3 green
24	5×10^6	53	51,8,3	30,4 (35,7 %)	cy5 red
0	5×10^6	—	—	22,4	cy5 green
72	5×10^6	17	45,36,2	38,4 (+72 %)	cy5 red

For every experiment we obtained a good RNA total extraction yield, as synthesized in the below tables. Moreover the RNA spectrophotometrical analysis has evidenced a high purity degree, being the ratio 260/280 nm always more then 1.9.

by reversed labeled mRNA of resting human T-lymphocytes, Figure 2 shows the perfect agreement between our automated DNASER technology and lengthy Fluorescence Microscopy.

Gene Expression of Human Lymphocytes in the Cell Cycle

Here we present the results relative to the 48 h stimulation experiments with respect to the resting counterpart (Figs. 3 and 4).

The images obtained by DNASER analysis were analyzed using GenePix software, the same was did with the images of the same array obtained by Fluorescence Microscopy (so to compare the results). The resulting image obtained by the superimposition of the sample images of lymphocytes at 0 h (in red) and 48 h (in green) was generated using the GenePix software. The image was acquired above an intensity threshold considered significant. The identification of the spots using GenePix is usually realized by comparing the image with a grid, which contains every information needed (e.g., coordinates and gene names) to recognize

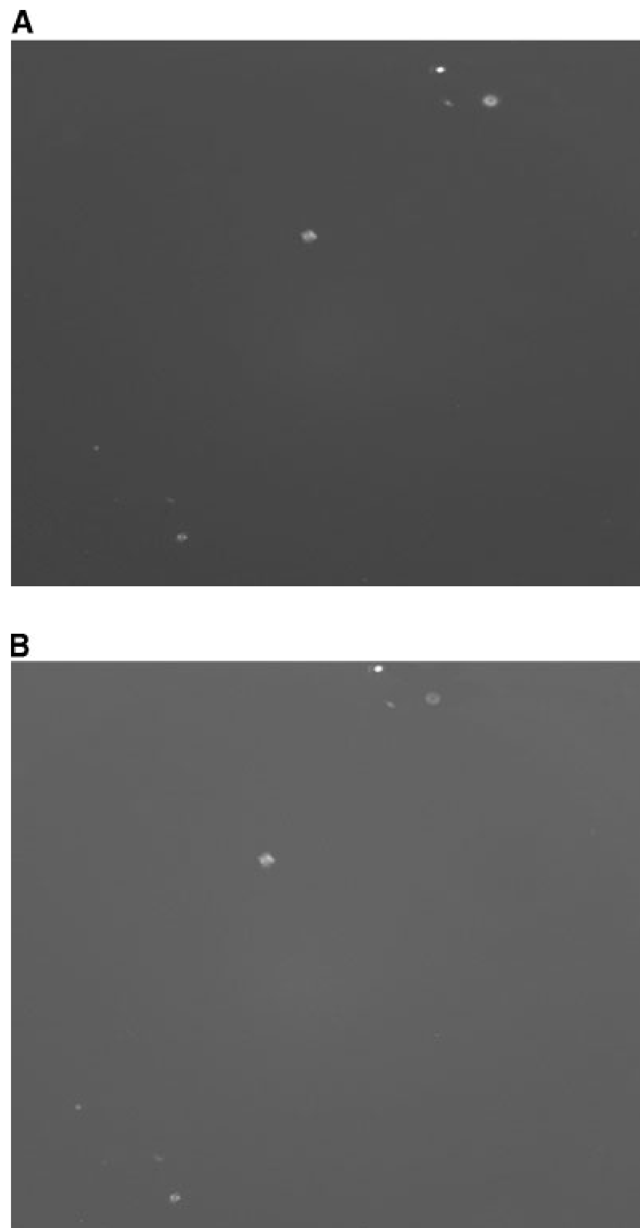

Fig. 2. Comparison of CCD Microscopy (**A**) versus DNASER (**B**) green fluorescence acquisition of the same DNA microarray (STARTER array) from resting human T-lymphocytes.

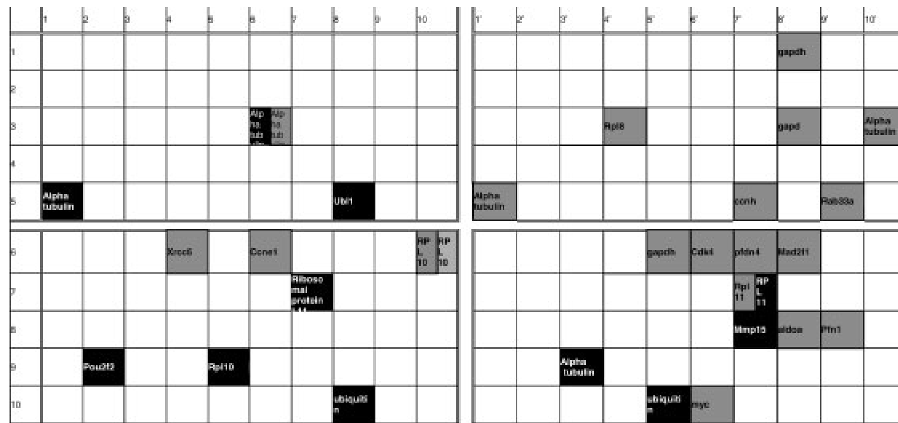

Fig. 3. Localization of gene expression in the Starter Array, measured with GenePix, on the resting (black, 0h) and stimulated (grey, 48h) lymphocytes.

the spots. Anyway, grid position and dimension must be interpreted in a non-rigid way, to be a better response to experimental reality. Some of the genes identified in the t = 0 h sample were not identified in the t = 48 h sample, not allowing the calculation of a ratio intensity. We decided thus to measure the absolute intensity of every spot, as the median of the intensity of every single pixel. This measure can give some information about gene expression level. We assigned to every spot an indicator of the intensity degree and thus of the expression: high, median, or low (see Figs. 3 and 4). Intensities were then normalized. In this way, we focused our attention on 22 genes: 7 active at t = 0 h and 16 at t = 48 h after the proliferative

Gene	Time (t)		Expression level	
	0	48	t = 0	t = 48 h
Alpha tubulin	■ (black)	■ (grey)	0,38	0,09
Rpl 8		■ (grey)	0	0,06
Gapd		■ (grey)	0	0,05
Ubl1	■ (black)		0,35	0
Ccnh		■ (grey)	0	1,00
Rab33a		■ (grey)	0	0,12
Xrcc5		■ (grey)	0	0,06
Ccne1		■ (grey)	0	0,10
Rpl 10	■ (black)	■ (grey)	0,06	0,45
Gapdh		■ (grey)	0	0,24
Cdk4		■ (grey)	0	0,08
Pfdn4		■ (grey)	0	0,55
Mad2l1		■ (grey)	0	0,09
Rpl 11	■ (black)	■ (grey)	0,08	0,89
Tgfa		■ (grey)	0	0,05
Mmp15	■ (black)		0,05	0
Aldoa		■ (grey)	0	0,08
Pfn1		■ (grey)	0	0,06
Pou2f2	■ (black)		0,10	0
Adprt	■ (black)		0,56	0
Ubiquitin	■ (black)		0,05	0
Myc		■ (grey)	0	0,08

Fig. 4. Expression of the different genes in the Starter Array, measured with GenePix, on the resting (black, 0h) and stimulated (grey, 48h) lymphocytes. Different intensities represent different gene expression.

stimulus. It is interesting to observe how, in response to the PHA stimulation, there is an activation of genes involved in the control and regulation of the cell cycle, as MYC, CDK4, CCNH, CCNE1, TGFA, XRCC5, GAPD. We found the maximum expression level in the CCNH gene. It is also possible to see that Rpl 10 and Rpl 11, coding for ribosomial proteins and the alpha tubulin genes were expressed before and after the stimulation, even if with different intensities. They are, in fact, gene of structural importance or belonging to the transcription machinery, and they are at the basis of every biological process.

Correlation With Bioinformatically Identified Gene Leaders

We compared our data with those deriving from a recent bioinformatic analysis of human T-lymphocyte cell cycle, which has led to the identification of six important genes or "leader genes" [8]. Of those six genes, two are displayed on the Human Starter Array: CDK4 and MYC. Both of them were found to be expressed after the PHA stimulation. MYC and CDK4 are known to lead the progression from G_0 to G_1 and to G_1 to S phases, respectively. After a mitogen stimulation, like the one with PHA, cells are known indeed to pass from the quiescent state to the proliferating state (Fig. 4) [Abraham et al., 1980]. The activation of MYC and CDK4 is a signal of this important transition. This represent a fundamental validation of the bioinformatics algorithm Anyway, the four other leader genes (CDK2, CDC2, CDKN1A, CDKN1B) are not present on the analyzed array. Moreover, in [Sivozhelezov et al., 2005], we also identified 27 unlinked or "orphan" genes, whose specific linkages in the complex gene network of cell cycle control are still to be investigated. None of them is present on the Human Starter Array.

CONCLUSION

In the DNASER a white light beam illuminates the target sample allowing the images grabbing on a high sensibility and wide-band charge-coupled device camera (ORCA II—Hamamatsu). This high-performance device permits to acquire images faster and of higher quality than the traditional systems. The DNA microarrays images are processed to recognize the DNA chip spots, to analyze their superficial distribution on the glass slide and to evaluate their geometric and intensity properties. Differently form conventional techniques, the spots analysis is fully automated and the DNASER does not require any additional information about the DNA microarray geometry.

It is the first time that the DNASER technology is employed on commercial DNA array rather than on our novel surface-patterned microarray [Troitsky et al., 2002]; the results obtained with this technology then have been validated by comparison with that obtained by a traditional fluorescence microscope equipped by a CCD camera. Absolute fluorescence intensities as reliably measured by DNASER technology appear also of interest per se as quantitative indicator of mRNA level.

We verified that, in response to the PHA stimulation, there is an activation of genes involved in the control and regulation of the cell cycle and, by a bioinformatic analysis of human T-lymphocyte cell cycle, we identified six important genes or "leader genes" [Sivozhelezov et al., 2005]. Two of them are displayed on the Human Starter Array and interestingly both were found to be expressed after the PHA stimulation.

Thus, it becomes evident that a bioinformatic analysis must be the first step in the choice of an array targeted on a precise cell system. Commercial arrays often displays a large amount of genes, which covers almost every aspect of cell life and biology. They can be used for general purposes, but, if a deeper view is requested, an appropriate array, displaying the interesting genes emerging from a bioinformatic analysis, must be developed. Obviously microarray experimental data can provide quantitative information, which the bioinformatic analysis lacks. Theoretical and experimental analysis, thus, are complementary and give a precise and detailed vision of a particular biological process.

By the employment of bioinformatics, in the future, we will be able to project and to construct array addressed to the analysis of cell cycle gene expression that we think will be resolutive in the understanding of these complex mechanisms.

ACKNOWLEDGMENTS

We are grateful to Victor Sivozhelezov. This work was supported by FISR–MIUR grant on DNA Nanotechnology to Genova Section of the INBB (National Institute Biostructures) and to

Nanoworld Institute-CIRNNOB of the University of Genova and by a FIRB–MIUR grant on Organic Nanoscience and Nanotechnology to the Nanoworld Institute, namely to the Fondazione EL.B.A and to the Centro Interuniversitario di Ricerca sulle Nanotecnologie e Nanoscienze Organiche e Biologiche of Genova University.

REFERENCES

Abraham S, Vonderheid E, Zietz S, Kendall FM, Nicolini C. 1980. Reversible (G0) and nonreadily reversible (Q) noncycling cells in human peripheral blood. Immunological, structural, and biological characterization. Cell Biophys 2:353–371.

Butte A. 2002. The use and analysis of microarray data. Natl Rev Drug Discov 1:951–960.

Cantrell D. 2002. Protein kinase B (Akt) regulation and function in T lymphocytes. Semin Immunol 14: 19–26.

Isakov N, Altman A. 2002. Protein kinase C(theta) in T cell activation. Annu Rev Immunol 20:761–794.

Jones A, Hunt E, Wastling JM, Pizarro A, Stoeckert CJ, Jr. 2004. An object model and database for functional genomics. Bioinformatics 20:1583–1590.

Nicolini C, Malvezzi AM, Tomaselli A, Sposito D, Tropiano G, Borgogno E. 2002. DNASER I: Layout and data analysis. IEEE Trans Nanobiosci 1:67–72.

Oosterwegel MA, Mandelbrot DA, Boyd SD, Lorsbach RB, Jarrett DY, Abbas AK, Sharpe AH. 1999. The role of CTLA-4 in regulating Th2 differentiation. J Immunol 163:2634–2639.

Sivozhelezov V, Giacomelli L, Tripathi S, Nicolini C. 2005. Gene expression in the cell cycle of human T lymphocytes: I. Predicted gene and protein networks. J Cell Biochemistry this volume.

Troitsky V, Ghisellini P, Pechkova E, Nicolini C. 2002. DNASER II. Novel surface patterning for biomolecular microarray. IEEE Transaction on Nanobiosciences 1:73–77.

Troyanskaya OG. 2005. Putting microarrays in a context: Integrated analysis of diverse biological data. Brief Bioinform 6:34–43.