COMMITTEE CERTIFICATION OF APPROVED VERSION

The committee for Peter Szaniszlo certifies that this is the approved version of the following dissertation:

GENE EXPRESSION MICROARRAY ANALYSIS OF SMALL, PURIFIED CELL SUBSETS

Committee:

James F. Leary, Ph.D., Supervisor

Istvan Boldogh, Ph.D., Supervisor

Bruce A. Luxon, Ph.D.

William R. Widger, Ph.D.

David W. Niesel, Ph.D.

Victor E. Reyes, Ph.D.

Dean, Graduate School

GENE EXPRESSION MICROARRAY ANALYSIS OF SMALL, PURIFIED CELL SUBSETS

by
Peter Szaniszlo, M.D.

Dissertation
Presented to the Faculty of The University of Texas Graduate School of
Biomedical Sciences at Galveston
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

Approved by the Supervisory Committee

James F. Leary, Ph.D.
Istvan Boldogh, Ph.D.
Bruce A. Luxon, Ph.D.
William R. Widger, Ph.D.
David W. Niesel, Ph.D.
Victor E. Reyes, Ph.D.

April, 2007
Galveston, Texas

Key words: gene expression profile, genomics, cell purification, multistage magnetic sorting, laser enabled analysis and processing, RNA amplification, microgenomics

To my family

ACKNOWLEDGEMENTS

I wish to thank the members of my committee, Dr. J.F. Leary for his guidance as my mentor, Drs. I. Boldogh, B.A. Luxon, W.R. Widger, D.W. Niesel, and V.E. Reyes as members of my Ph.D. Thesis Supervisory Committee. Their scientific advises and critical comments have greatly improved the quality of my work. I would also like to thank the former committee members, Drs. W.A. O'Brien, A. Simmons, and A. Nakeff for their support. I an grateful to the former and present chairmen of the Microbiology and Immunology Department, Drs. S. Baron, S.M. Lemon, and D.W. Niesel for providing the necessary facilities for the successful completion of this project. I also thank the former and current Graduate Program Directors, Drs. W.R. Fleischmann, D.W. Niesel, and T.K. Hughes for their assistance through the graduate school process. I am very thankful to all the members of Dr Leary's and Dr. Boldogh's laboratories for being such great colleges and friends. I also owe my gratitude for the administrative staff of the department, S. Daniel, M. Lewis, and others who have always helped me when I asked them. I would also like to thank Drs. P. Todd, M.R. Koller, and B.O. Palsson for trusting me with parts of their projects. My sincere gratitude to Dr. N.J. Garg for helping me succeed in earning my degree. Very special thanks to Dr. J.I. Rosenblatt for sharing some of his wisdom with me and for not letting me forget that science should be fun. I could not have completed this work without the encouragement, support, and forbearance of my mother, Erzsebet Szaniszlo, my father, Jozsef Szaniszlo, my sister, Zsuzsanna Szaniszlo, and my dear wife, Marta Lorinczi; I owe them my eternal gratitude and love. Drs. I. Boldogh, J.F. Leary, and the Department of Microbiology and Immunology provided financial assistance for this study.

GENE EXPRESSION MICROARRAY ANALYSIS OF SMALL, PURIFIED CELL SUBSETS

Publication No._____

Peter Szaniszlo, Ph.D.
The University of Texas Graduate School of Biomedical Sciences at Galveston, 2007

Supervising Professor: James F. Leary

Gene expression microarray technology is potentially capable of examining all of the cellular processes at the mRNA level at a given moment. One major challenge in its applicability is that most biological samples are cell mixtures and the cell type of interest is often a minor cell subset. Using well-defined mixtures of model cell types with different cell ratios we found that the overall gene expression profile (GEP) of mixed cell populations was the weighted average of the GEPs for each cell subpopulation in the cell mixture. Thus, without applying any cell separation the cell type in majority dominated the overall GEP of the sample while the GEPs of minor cell subsets were diminished. We showed that the functional threshold for the necessary purity of a cell type in a sample to produce virtually identical overall GEP to a pure sample was 75% and this could be achieved by conventional cell sorting methods without altering the overall GEP in the process. For the purification of small, biohazardous samples, we tested the applicability of multistage magnetic sorting (Magsort) and laser enabled analysis and processing (LEAP). We developed optimized sample labeling and sorting protocols for both technologies and demonstrated that while the maximum purity we could achieve with Magsort was 75-80%, with the LEAP instrument we could purify fluorescently labeled cell subsets to 80-100%. The purified cells from biological samples often do not provide enough RNA for direct microarray studies without RNA amplification. We found that both linear and exponential amplification was capable of producing enough RNA for microarray analysis even from a single cell. Both methods distorted the GEP, however, with linear amplification much fewer genes were affected and only this method preserved the GEP differences between samples. Further studies are needed to analyze and possibly eliminate all GEP distortion. In conclusion, the purification of minor cell subsets from biological samples prior to microarray analysis is not only necessary, but also achievable without GEP distortion. Using linear RNA amplification of small purified samples, meaningful microarray data can be produced about the GEP of even a few cells.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A2780 cells	human epithelial cell line (ovarian carcinoma)
ABC	antibody binding capacity
aRNA	antisense ribonucleic acid
BE	beam expander
CBMCs	cord blood mononuclear cells
CCD	charge-coupled device (camera)
CD	cluster of differentiation
cDNA	complementary deoxyribonucleic acid
CEM cells	human T-cell line (acute lymphoblastoid leukemia)
cRNA	complementary ribonucleic acid
CTO	CellTracker Orange
DNA	deoxyribonucleic acid
ds	double stranded
dT	deoxythymidine
EGF	epithelial growth factor
FCS file	flow cytometry standard format file
FITC	fluorescein isothiocyanate
FSC	forward scatter
galvo	galvanometer
GEP	gene expression profile
GFP	green fluorescent protein
GUI	graphical user interface
HeLa cells	human epithelial cell line (cervix adenocarcinoma)
HiReCS	high resolution cell sorter
HIV	human immunodeficiency virus
IR	infrared
kD	kilodalton
KG-1a cells	human stem-cell line (bone marrow acute myelogeneous leukemia)
LCM	laser capture microscopy/microdissection
LEAP	laser enabled analysis and processing
LPC	laser pressure catapulting
MA	microarray
MACS	automated magnetic cell sorting (Miltenyi)
Magsort	multistage magnetic sorting
MAT	microarray technology
MM	mismatch
mRNA	messenger ribonucleic acid
mT	millitesla
MW	molecular weight
NDF	neutral density filter
PBMCs	peripheral blood mononuclear cells
PBS	phosphate buffered saline

PCR	polymerase chain reaction
PE	R-phycoerythrin
PM	perfect match
poly-A	poly-adenine
RNA	ribonucleic acid
QMS	quadrupole magnetic sorting
R^2-value	multiple correlation coefficient
SSC	side scatter
TMR	tetramethylrhodamine
TS	tag sequence
UV	ultraviolet

CHAPTER 1.

INTRODUCTION

GENE EXPRESSION MICROARRAY ANALYSIS

In the past decade a paradigm shift has been taking place in molecular biology. Pioneered by the Human Genome Project (1,2) and other full-genome sequencing studies (3-5) researchers started to use high throughput technologies to grasp the complexity of biological systems. These breakthrough efforts signaled the arrival of the "Omics Era" where in the new fields of genomics, transcriptomics, proteomics, cytomics, and many other 'omics' a comprehensive view at entire species and networks of biomolecules started to facilitate the step by step discovery process (6). For the field of transcriptomics, gene expression microarray technology (MAT) became a rapidly developing analytical tool in basic and clinical research.

The first microarray (MA) was published in 1995 (Figure 1.1.) (7), and the number of publications using MAs had been growing exponentially from 1995 to 2000 (8,9). Since the new millennium this number has kept increasing in a linear fashion reaching a total of over 42,000 by April, 2007. In the 1990s MAT was considered to be one of the most promising new approaches for the science of the 21st century, where fast, high throughput technologies would be used to solve the intricate puzzles a living cell or organism presents (7,10,11). MAT is potentially capable of analyzing the entire transcriptome (all of the cellular processes at the mRNA level) at a given moment. It delivers a comprehensive quantity of data about the transcriptional level, and it is expected to reveal novel processes, pathways and molecular interactions in the living cell. MAT data, providing "freeze-frame" views of the transcriptome, could help us understand the role and weight of known and yet to be discovered molecular mechanisms in the 'big picture' (12-14).

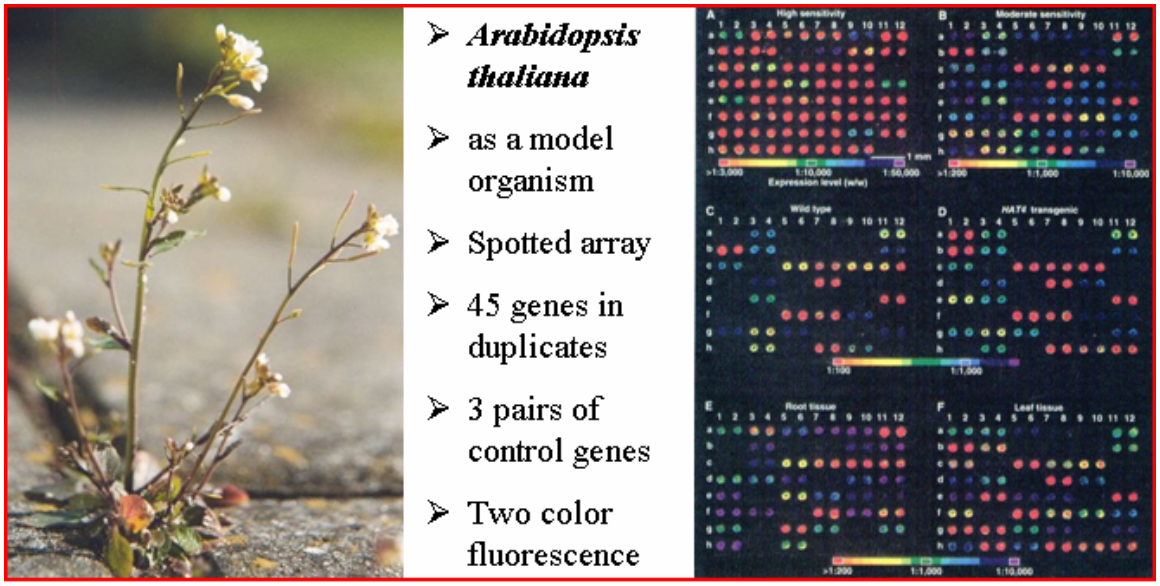

Figure 1.1. The first microarray ever published

Left: Arabidopsis thaliana, a small flowering plant. Right: Spotted microarray results of the expression levels of 45 *A. thaliana* genes and 3 control genes all in duplicates. Two fluorochromes (fluorescein and lissamine) were used. After intensity readings for both fluorochromes the images were pseudocolored. (Adapted from Schena, M., D. Shalon, et al. (1995). "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." Science. **270**(5235): 467-7

In clinical research MAT is expected to help us better understand the molecular basis of diseases, the difference between healthy and diseased cells, tissues, and organs, giving us clues about potential treatment (15-17). At the same time MAs can be used to monitor the effects of these treatments, especially of new (and old) drugs (18-20). The unique view at the cells and tissues that MAT offers also generated high hopes in the field of pathology. It was expected to revolutionize and automate the analysis of tissue sections and the classification of diseases, disease subtypes and stages. Surprisingly, in this latter field (especially in the analysis and classification of different types, subtypes and stages of tumors) MAT has proven to be immediately useful delivering very promising and convincing results (21-23). In the last few years gene expression profiling led not only to accurate sub-classification of tumors, but offered predictions for the tumor's response to specific treatments revolutionizing the concept of patient-tailored therapy (24-28).

However, in basic research, initial MA studies often caused disappointment, failing to generate or confirm new hypotheses (24,29,30). In many occasions MAT has generated more confusion than comprehension, more questions than answers (20,31,32). Despite the fact that most of the early problems with MAT (low reproducibility and sensitivity, high background, standardization of sample preparation and data analysis) has improved greatly in recent years, several studies found poor correlation between corresponding data sets from different laboratories and questioned the reliability of microarray data in general (30,33). The underlying problem is that different laboratories use different MA platforms, RNA preparation and hybridization protocols, and they interpret their results using different data analyses (13,34-36). Although the general principle of microarray technology is the same for all platforms, the practical differences might lead to different results. A gene expression microarray is an ordered array of immobilized nucleic acids. MAs operate as a "reverse Northern blot" where the unlabeled probe sequences are pre-attached to a membrane or glass surface, and the sample RNA gets labeled and hybridized to these immobilized probes. Although various commercial and custom MA platforms have been developed most of them belong to one of two basic

types (37-40). Arrays of the first type, the short oligonucleotide arrays, are manufactured exclusively by Affymetrix Inc. MAs of the second type, the spotted arrays, include a wide variety of platforms (37,38). Spotted array technology (Figure 1.2.) produces the probe sets first (usually by PCR reactions) and immobilizes them afterwards (usually by robotic printing) onto a nitrocellulose membrane or glass surface. Short oligonucleotide arrays (37,39) are produced by a process called photolitography where a microscopic array of 25-mer oligonucleotide probes are synthesized onto the surface of a silicon chip nucleotide by nucleotide (Figure 1.3.). In addition, there are several differences between the platforms in sample preparation, hybridization and data interpretation, some of which are listed in Table 1.1. To produce comparable data across platforms, array types, and laboratories, recent studies suggest that researchers should standardize protocols as much as possible, carefully choose methods for sample preparation and data analysis based on the biological problem they investigate, and document experiments accurately (41,42).

Another main source of the remaining problems is that, while MA analysis requires 0.5-5 million cells per sample, biological systems (tissues, organs) with this number of cells are almost always mixtures of several different cell types (43-45) - all of which may behave differently in a given experiment. The two main exceptions are cell lines and tumors, where one can have more than enough cells of the same type for MA analysis. Not surprisingly, these two are the main fields of successful MA studies (46-50). However, tumors only represent a narrow segment of pathological processes in humans, and immortalized cell lines have been repeatedly shown to differ significantly from *in vivo* cells, even the ones from which they originated (26,51). Other MA studies of unsorted cells were succesful because most of the cells in the sample behaved similarly, thus their reaction to experimental condition changes could be detected (52-54). Nevertheless, even in these cases, many subtle effects might have remained undetected, overshadowed by the non-reacting cells.

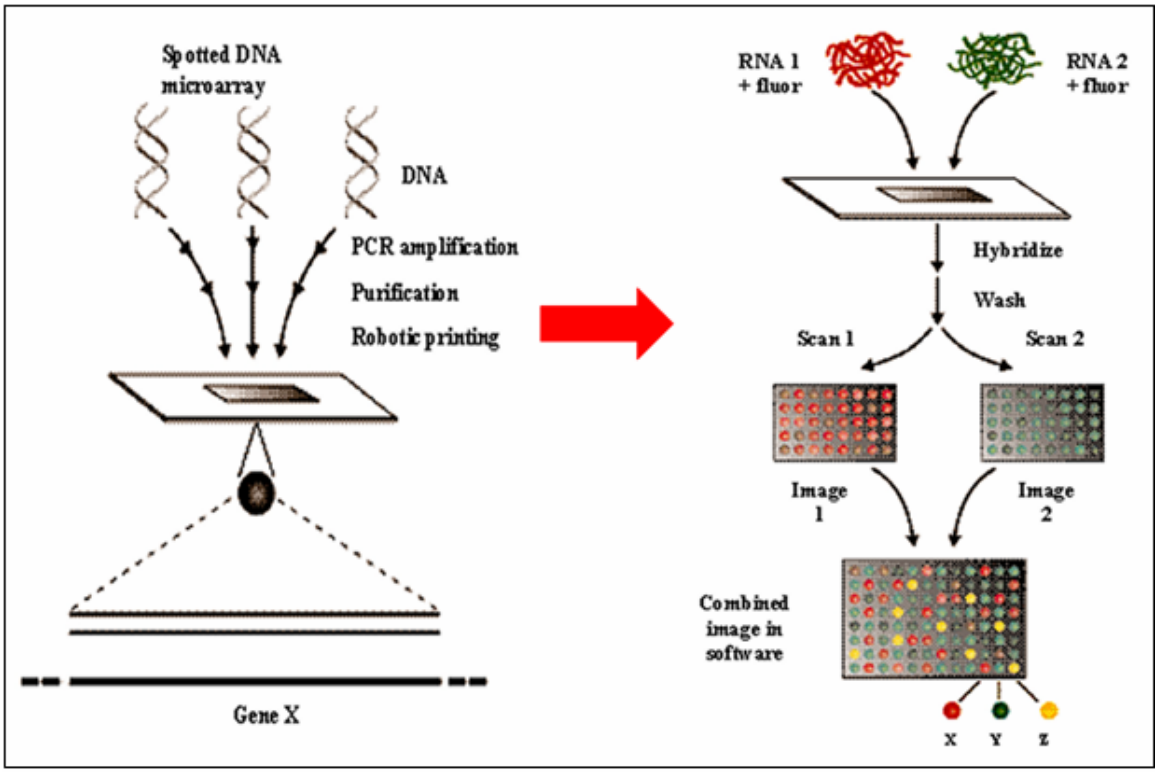

Figure 1.2. Spotted microarray technology

Left: Spotted microarray fabrication by robotic printing of PCR amplified probes onto a plastic or glass surface. Right: Two samples are labeled with different fluorochromes and hybridized to the same microarray. These samples are directly comparable by microarray analysis. (Adapted from Harrington, C. A., *et al.* (2000) *Current Opinion in Microbiology* 3:285-291)

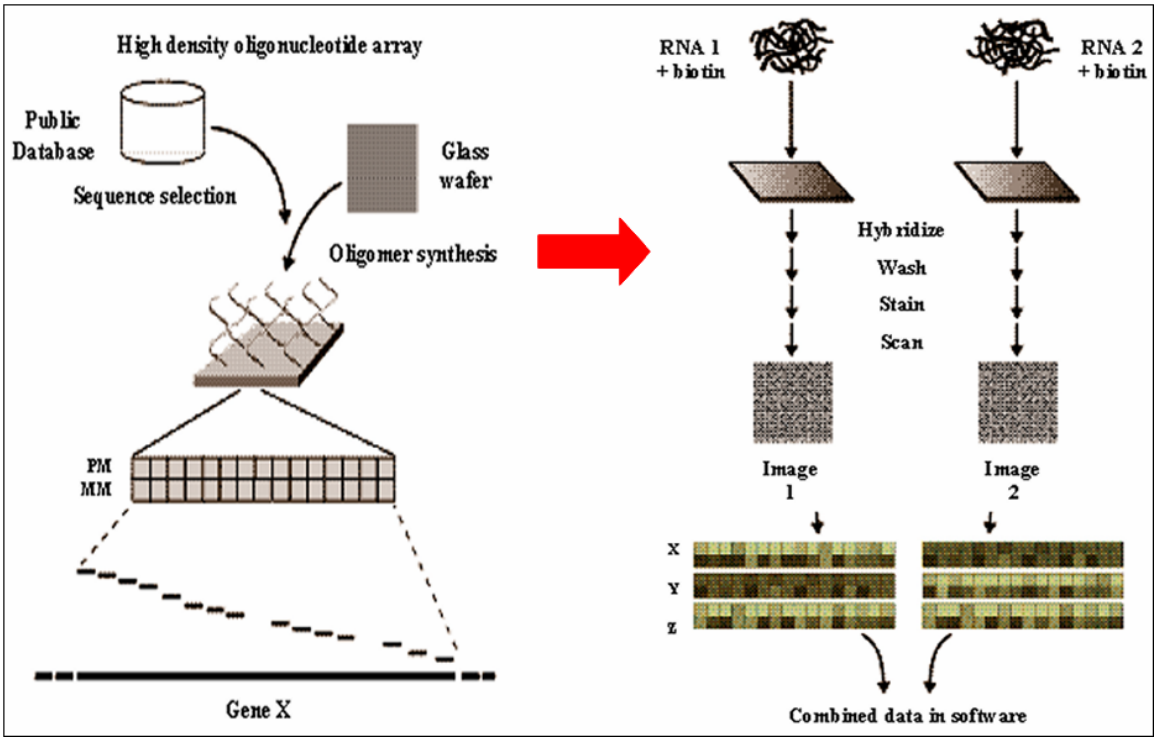

Figure 1.3. Affymetrix microarray technology

Left: Short oligonucleotide microarray fabrication by synthesizing 25-mer oligos onto a silicon surface. Perfect match (PM) and mismatch (MM) probes are used to differentiate between specific and non-specific hybridization. Right: Two samples are labeled with biotin and hybridized to two arrays of the same kind. The hybridized probes are fluorescently labeled and scanned. The data analysis is done by computer software. (Adapted from Harrington, C. A., *et al.* (2000) *Current Opinion in Microbiology* 3:285-291)

Array type	**Affy. Array**	**Spotted Array**
Array fabrication	photolithography	robot spotter
Probe	25-mer oligo	30-120-mer to 0.5-2.0kb
Mismatch probe	yes	no
Genes per array	**~38,500**	**100-40,000**
Number of probes/gene	11 + 11	1-3
Features per array	> 1,000,000	100-50,000
Starting material	**5-25µg total RNA**	**0.5-10.0µg total RNA**
Sample hybridized	aRNA (cRNA)	cDNA, mRNA, aRNA
Samples per array	1	1 or 2
Detection range	4 orders of magn.	3 orders of magn.
Reproducibility	**high**	**high**

Table 1.1. Comparison of Affymetrix and spotted arrays

The two types of gene expression microarrays differ in many aspects (features in black), but most importantly they are both capable of analyzing the entire human genome, they both require microgram amounts of RNA and they both generate highly reproducible results (features in red).

Several studies addressed this problem by purifying the cells of interest prior to microarray analysis and many of them have shown promising results (55-57). However, it is still not well understood how different cell sorting methods and sample handling protocols affect the gene expression profile (GEP). It is a reasonable concern that during cell purification the original profile of the purified cells gets altered due to the vulnerability and transient nature of the mRNA molecule (31,58-61). It is also not well known how much effect different cell types have on each other's GEP when they are mixed together. Another unanswered question about cell sorting for microarrays is how much purification is needed when 100% purity is not a feasible option. These questions need to be addressed to produce meaningful MA data.

NEW CELL SEPARATION TECHNOLOGIES

Multistage Magnetic Sorting

The most commonly used cell purification technologies are column-based immuno-magnetic cell separation and flow cytometric cell sorting. Traditionally, in immuno-magnetic cell separation a mixed suspension of cells in which specific cell types have been labeled with immuno-magnetic reagents (magnetite-containing microparticles or nanoparticles) has been added to a column of steel beads surrounded by a strong permanent magnet. A high magnetic field gradient is created around the beads, to which cells attach. Non-magnetized cells do not adsorb, and, after washing the column, magnetized cells, which do adsorb, are removed by removing the magnetic field and washing the column (62). In alternative methods, steel beads might be replaced by wires or needles that also provide steep magnetic field gradients. This technology is capable of very high throughput and - as another major advantage - it sorts cells in a closed system; therefore it can be used to purify biohazardous samples. This method constitutes binary magnetic sorting. All magnetized cells, regardless of amount of label, are collected, and all other cells are not. In a typical application from a 10-50% pure sample one round of magnetic bead sorting raises the purity to 70-80%, an additional round increases it to or above 90%, and three consecutive rounds usually results in 95-99% sample purity (63).

Besides its relatively low sort efficiency another disadvantage of column-based immuno-magnetic cell separation is that it is challenging to purify very small samples with it without serious cell loss.

The other most commonly used method, flow cytometry can be used to select cells on the basis of fluorescent label content, so that proportional, and multicolor sorting, rather than just binary, and single label sorting is available. Since collected cells are selected on the basis of a "sort window" most flow cytometric sorts are also ultimately binary; that is, one cell fraction is kept while labeled cells outside the sort window are most often discarded. Although for most traditional applications flow cytometric cell sorting is the method of choice, sorting very large or very small samples is still a challenge with this method. Recently, high speed cell sorting brought the throughput of this technology closer to that of immuno-magnetic sorting. A major disadvantage of flow cytometric cell sorting is that it aerosolizes the sample therefore it should not be used to sort biohazardous samples.

Quadrupole magnetic flow sorting (QMS) works in almost exactly the same way as flow cytometric sorting; however, the throughput is significantly higher (up to 10^7 cells/s). The concept of proportional magnetic sorting was derived from QMS technology. It has been shown by direct measurements (64), that cells labeled with subsaturating concentrations of immunomagnetic reagent become labeled in direct proportion to their Antibody Binding Capacity (ABC). These studies demonstrated that the magnetophoretic mobility (velocity of a cell in a given magnetic field and field gradient) could be calculated from the volumetric susceptibility of the labeling particles, the cell diameter, the dynamic viscosity, and the magnetic permeability of free space. Multistage magnetic sorting (Magsort) has been described previously and has - as a unique feature - the ability to sort cells and particles according to their degree of magnetization and to collect every fraction sorted (65-68). Figure 1.4. illustrates the general principle of this new technology. Magsort is capable of dividing a magnetically

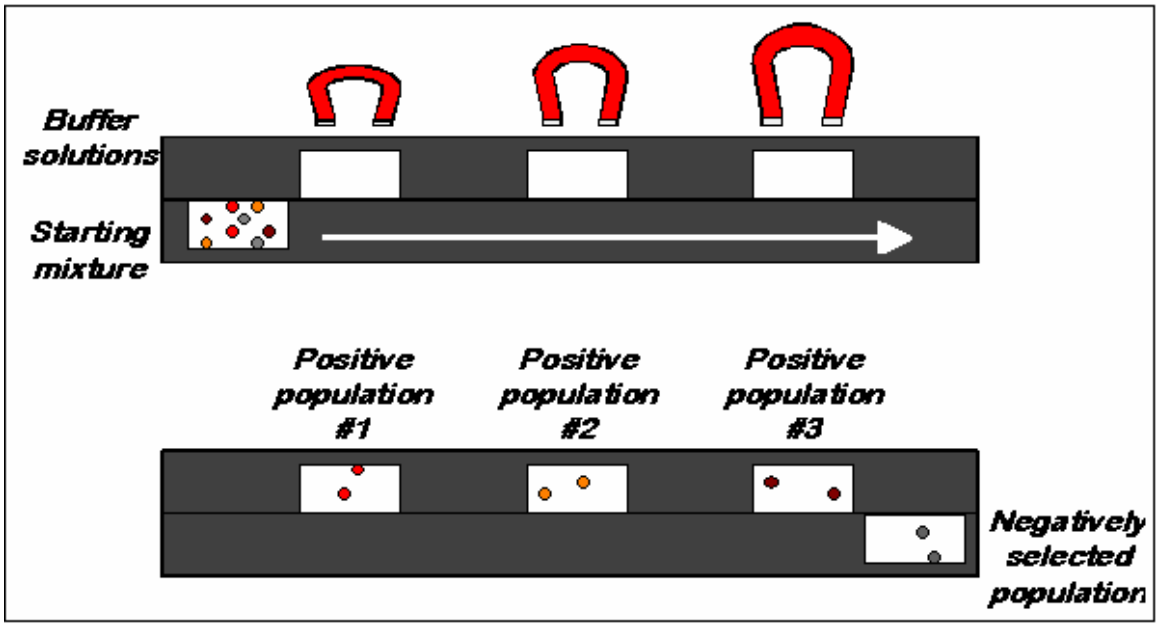

Figure 1.4.. General principle of multistage magnetic sorting.

A suspension of magnetically labeled cells is exposed to increasingly stronger magnetic fields in a stepwise fashion. Magnetically labeled cells are pulled up from the sample well into the sort chamber. By placing increasingly stronger magnets above the sample well, fractions of cells with decreasing magnetic susceptibility are collected. Unlabeled cells are negatively selected by remaining in the sample well.

labeled cell population into up to 15 individual fractions on the basis of magnetophoretic mobility. Cells are separated based on the amount of magnetic material attached to the cell surface, which is directly proportional to the surface antigen content of immuno-magnetically labeled cells. Receptor or cell-specific antibodies can therefore be used to not only separate the cells of interest from unwanted (unlabeled) cells, but to further separate those selected cells into more specific fractions. This system is designed for small samples and it works in a closed environment, therefore it may be suitable for purifying small, biohazardous samples for consecutive microarray analysis. Its feasibility for such experiments still needs to be tested.

Laser Enabled Analysis and Processing

Laser scanning cytometry uses laser-based opto-electronics and automated analysis capabilities to simultaneously and rapidly measure biochemical constituents and evaluate cell morphologies. While scanning cytometry has been with us for many years, the modern era of scanning cytometry began with the pioneering efforts of Kamentsky and others to produce a user-friendly commercial version through CompuCyte, Inc. (69,70). The new generation of instruments has the capability of fluorescence-based measurements which make use of the rich diversity of molecular probes now available. These molecular probes can bear different color fluorescent tags which can then be combined in Boolean combinations to distinguish cell subpopulations; much as has been done with flow cytometry (71). While excellent work was done extracting features from Feulgen and other non-fluorescent stains, one could only extract information that was there. Application of multicolor molecular probes greatly adds to the information content of each image pixel, essentially extending the analysis from two-dimensional to multidimensional. The result of these advances is that scanning cytometry now has many of the features (and power) of multiparameter flow cytometry while keeping its own advantages as a technology (72). These advantages of scanning cytometry include the ability to: (A) perform fluorescence and light scatter imaging measurements on either attached or suspension cells with a minimum of cell manipulation, (B) measure not only cell and nuclear morphology but also the spatial distribution of fluorescent molecular

probes, (C) return to the same spatial location for additional measurements of the same probes in time (kinetic measurements) or for re-staining with different probes. Other sophisticated instruments with capabilities for high-throughput screening have been developed (73,74) and sold by Beckman-Coulter, Inc. (73) and Cellomics, Inc. (75). A wide variety of applications have been developed by several key laboratories (76-85).

More recently, new instruments have been developed which combine some of the capabilities of scanning cytometry with the ability to manipulate cells. One of these, Laser Capture Microdissection (LCM) was developed by researchers at the National Cancer Institute (86) and is now marketed through Arcturus, Inc. This is widely used by hundreds of research laboratories around the world and while a powerful technology, it has the technical limitation that the cells need to be either fixed or frozen. LCM has been an important tool in the new field of microgenomics which purifies small cell subpopulations prior to gene expression analysis. Measurements, however, cannot be made in a conventional aqueous environment. A competing technology of "laser pressure catapulting" (LPC) has been developed by PALM Microlaser Technologies AG to handle this problem. Another instrument has the ability to manipulate single cells with laser tweezer technology (87) and is now commercially available through Cell Robotics, Inc. A reliable technique to manipulate cells, manual microinjection, is an extremely tedious process requiring both skill and patience. Even more recent automated microinjection techniques by Eppendorf AG in Germany, while representing an important advance, still cause considerable cell injury and are still comparatively slow.

A new technology, called Laser Enabled Analysis and Processing (LEAP), has been developed by Cyntellect, Inc. (88-91). Figure 1.5. illustrates the general principle of this new technology on a simplified model. The model shows the main components of the instrument: a flat platform holding the cells, broad-band light source with a fast moving galvanometer (galvo) mirror, a laser and a camera (88-90). (The computer that controls

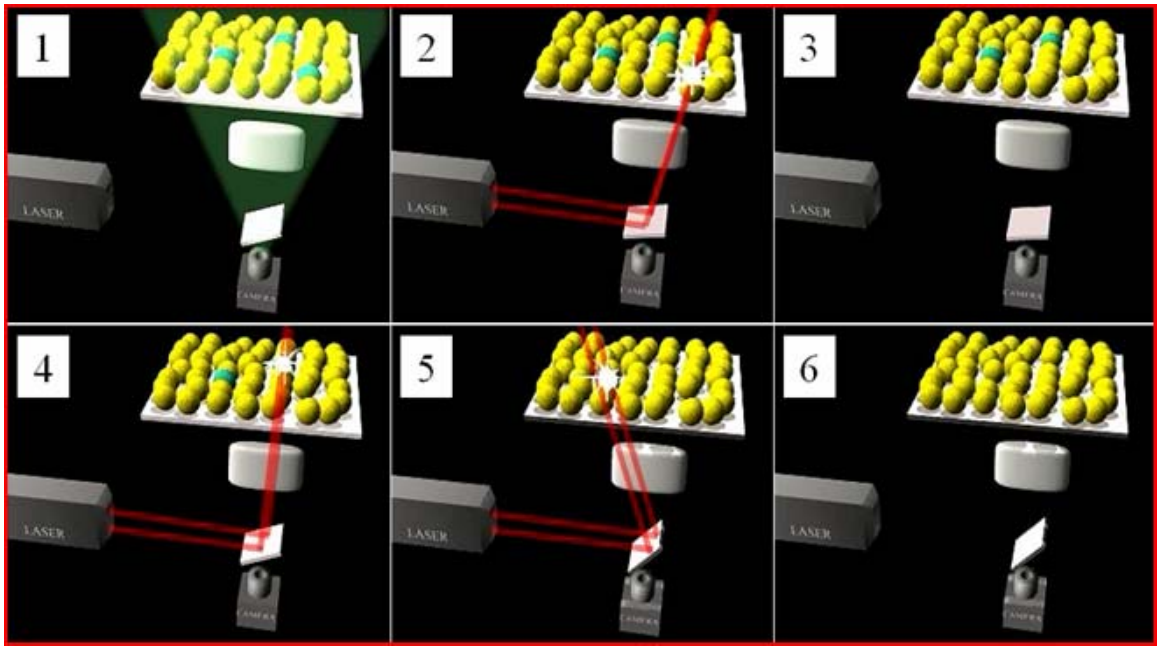

Figure 1.5.. General principle of LEAP sorting.

Fluorescently labeled target cells are illuminated and their fluorescent signal is detected by the camera (step 1). Based on the image sent by the camera, the computer registers the position of each target cell and by adjusting the position of the galvano mirror it directs the laser to start firing at the first target cell (step 2). The laser energy eliminates the target cell from the surface either by ablation or by laser catapulting (step 3). The laser keeps firing at each individual target cell targeted by the high speed galvano mirror (steps 4 and 5). At the end of this process all target cells of the field are eliminated while the purified cells remain unharmed (step 6).

all processes and acquires all data is not shown.) The target cells in the sample are specifically labeled with a fluorescent marker. Their fluorescence under illumination is detected by the camera. Based on the image sent by the camera, the computer registers the position of each target cell and by adjusting the position of the galvo mirror it directs the laser to start firing at the first target cell. The laser energy eliminates the target cells one by one from the surface either by ablation or by laser catapulting. At the end of this process all target cells of the field are eliminated and the camera takes an image of the next field starting the same process over. Using lower laser energy settings the target cells can be induced to go into apoptosis instead of being ablated (90). Shooting with even lower laser energy can result in optoinjection of the targeted cells by macromolecules added to the medium prior to LEAP (88). The key to the speed and accuracy of LEAP technology is the combination of real time image analysis, custom F-theta optics that provide a large field-of-view, and high-speed galvo mirrors to both scan the surface (to obtain images) and to steer a laser beam (to hit target cells). When fully automated, LEAP technology could potentially yield throughputs of greater than 100,000 events per second (88-91).

LEAP technology provides many of the advantages of LCM and LPC, as well as the ability to manipulate single cells by a variation of laser tweezer technology. Besides providing a method of live cell sorting based on laser ablation or laser catapulting, LEAP also offers a unique method for high-speed microinjection of macromolecules into living cells using a pulsed laser ("laser optoinjection") set at sub-lethal energies. This optoinjection can be very fast (hundreds of cells/sec) and, unlike electroporation, has a very low rate of injury to cells which can be individually selected on the basis of multiple fluorescent probes in an automated molecular imaging process. LEAP can be performed in a totally sealed environment in a process similar to inverted fluorescence microscopy. With all its advantages LEAP could be especially suitable to purify homogeneous cell subpopulations prior to microarray analysis from small, biohazardous samples that could not be easily handled by conventional cell sorting technologies (72). However, the effects

of laser-mediated cell purification and optoinjection on the purified or injected cells, especially on their gene expression profile, are not yet known.

MICROGENOMICS

Microarray analysis requires microgram amounts of total RNA. This constraint narrows down the field of applicability for this powerful technology considerably, because without RNA amplification such amounts are obtainable only from millions of cells (44,45,92,93). Biological samples of this size are usually heterogeneous cell mixtures (with the exceptions of tumors and immortalized cell lines) and after the necessary cell purification (63) they often yield much fewer cells of the purified cell type. In other cases, like in small biopsies, fine needle aspirates, rare- or micromanipulated cells, the original tissue sample is already too small for direct microarray analysis (94-96). Nevertheless, it would be of great value to analyze the GEP of these small biological samples to extend the benefits of microarray technology to a wide range of pathologic processes in defining sub-classes of diseases (other than tumors), predicting responses to different treatments and, ultimately, designing patient-tailored therapies (24-28).

Single cell biology is a rapidly growing field of research (44,45,93,97-99). Most tissues are complex mixtures of heterogeneous cells, all of which respond to physiological, pathological, and experimental conditions in a unique fashion characteristic to the cell type. Many of these responses are reflected in the transcriptional activity of the given cell. In the context of the cellular diversity of tissues the ability to study the entire transcriptome of single cells using gene expression microarray technology could yield valuable insight into the biochemistry, physiology, and pathology of biological systems (97). For many diseases the GEP alterations in single cells may be more informative about the underlying pathomechanism than regional expression patterns of tissues (96). In an effort to apply gene expression microarray technology to small samples and single cells, several methods have been developed to amplify picogram amounts of RNA available from these samples to microgram quantities required by microarrays (45,93,100-103). New methods for single cell identification, separation, and

handling have been developed as well as RNA extraction and preservation techniques adapted to single cell biology (44,45,93,94,98,101,104-106). Collectively, these efforts created the rapidly developing new field of microgenomics.

Most RNA amplification techniques belong to one of two basic method types, they can either be exponential (also called logarithmic) or linear technologies (93,96). The principle of exponential RNA amplification is illustrated by Figure 1.6. First, the carefully isolated and preserved RNA undergoes a reverse transcription reaction primed by oligo-dT primer to amplify only the mRNA species from total RNA. At the same time, this reaction is also used to label the produced cDNA with a universal tag sequence for the next step. The resulting cDNA is then amplified by PCR reaction using the universal tag sequence for priming. The resulting product is double stranded cDNA. The method is named exponential because the amplifying enzyme (usually Taq polymerase) uses the products of one cycle as templates in the next cycle. In contrast, linear amplification applies RNA polymerases (usually T7 RNA polymerase) as the amplifying enzyme which does not use the products of one round as templates in the next round directly. The principle of exponential RNA amplification is shown in Figure 1.7. The isolated RNA is also reverse transcribed by oligo-dT primer, but the additional sequence tag introduced by this method is a T7 promoter sequence. In the next step the T7 polymerase uses this sequence as starting point for transcription. The resulting product is antisense RNA (usually called aRNA or cRNA). Currently, the advantages and disadvantages of these technologies, especially their effect on the amplified GEP, are not fully understood. Avoiding GEP distortion by RNA amplification is of major importance for the field of microgenomics, and needs to be further investigated (63,93,96).

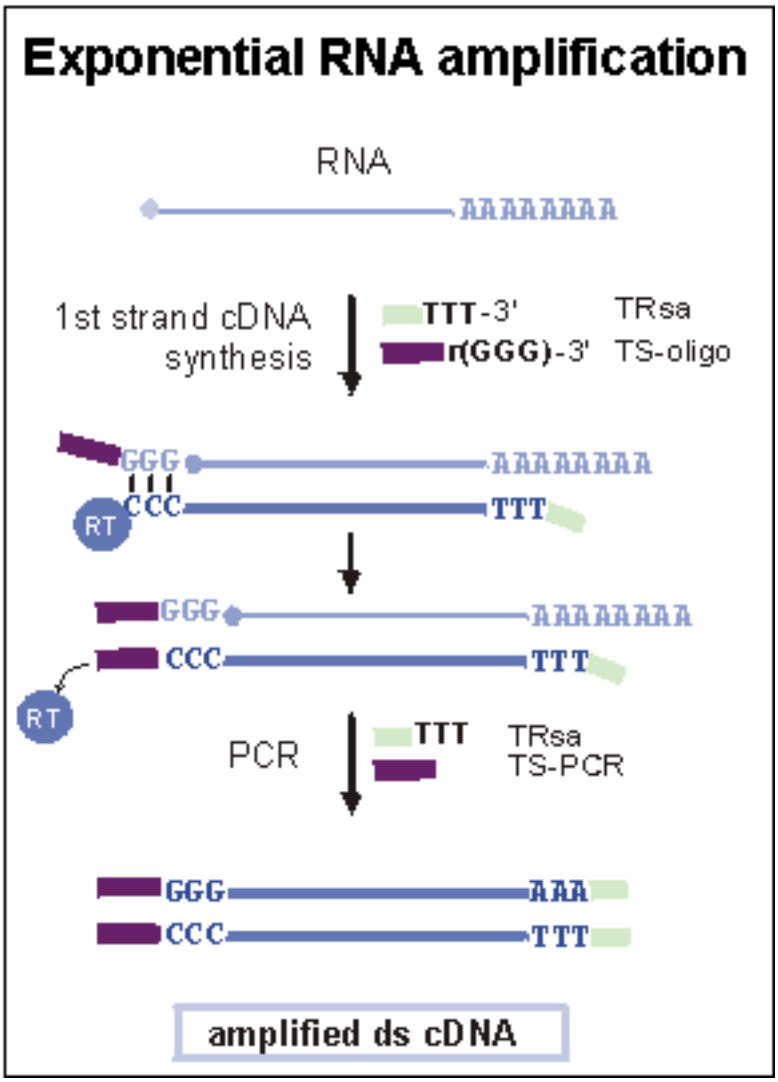

Figure 1.6. General principle of exponential RNA amplification

Reverse transription with oligo-dT primers only converts the poly-A tailed mRNAs into cDNA. The reverse transcription reaction is also used to label the cDNA with a universal tag sequence (TS-oligo). PCR reaction is used to amplify each reverse transcribed sequence using the universal tag olig sequence for priming. The resulting product is double stranded cDNA. (Adapted from www.clontech.com.)

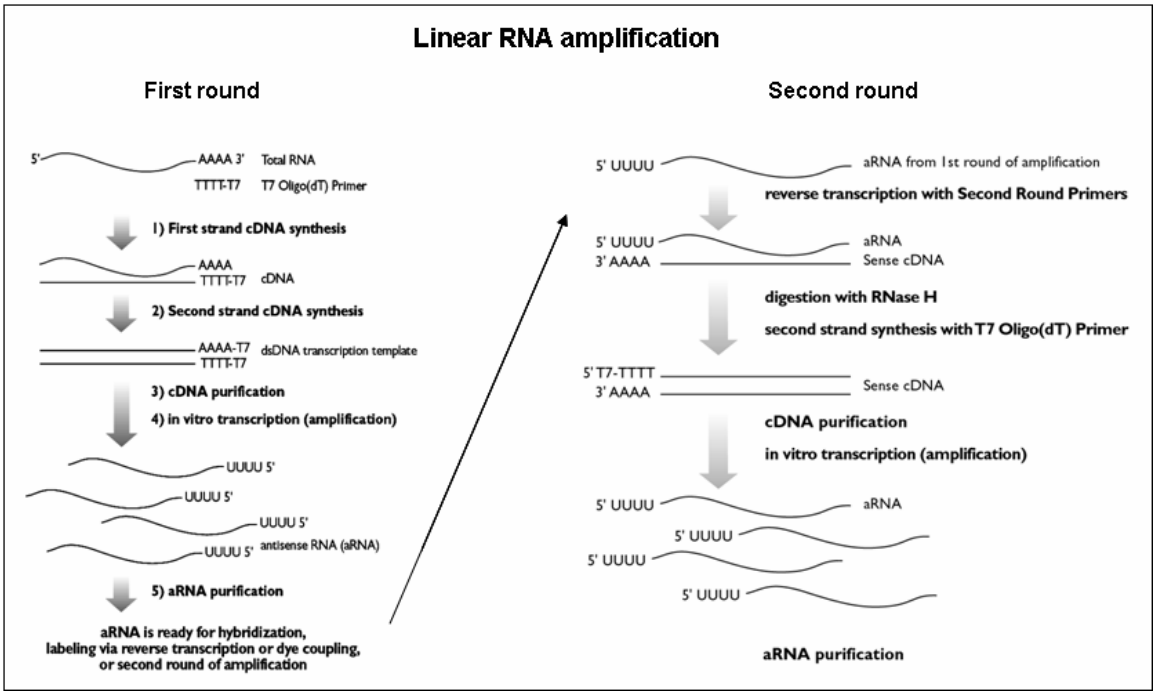

Figure 1.7. General principle of linear RNA amplification

Reverse transcription with oligo-dT primers only converts the poly-A tailed mRNAs into cDNA. The reverse transcription reaction is also used to label the cDNA with a T7 promoter sequence (T7 oligo dT-primer). In vitro transcription utilizing T7 RNA polymerase is used to amplify each reverse transcribed sequence starting transcription at the T7 promoter sequence. The resulting product is single stranded, antisense RNA (aRNA or cRNA). This product can be used for a second round of amplification if necessary. (Adapted from www.ambion.com.)

HYPOTHESIS AND OBJECTIVES

The overall hypothesis of this project is that it is possible to analyze the gene expression profile of small, purified cell subsets by microarray analysis after cell purification and RNA amplification. Our long term objectives are to develop a widely applicable method for the gene expression microarray analysis of a few cells or even of a single cell. Our immediate objectives are**:**

1. To test whether purifying the cells of interest prior to gene expression profile analysis will enhance the sensitivity and accuracy of microarray experiments. We compared the results of microarray experiments with and without prior sample purification. We studied how different degrees of purity affected gene expression profile (GEP) accuracy and sensitivity, and whether there was a practical threshold for general sample purity required for meaningful microarray analysis.

As discussed in the introduction, a cell's GEP is not static and is very sensitive to changing conditions around the cell as well as to the effects of the applied handling protocols. It has not been adequately studied how experimental conditions, especially purification techniques affect the cell's transcriptome. **2. To validate if cell purification procedures of mixed cell samples alter the cells' gene expression profile.** Here we studied the effects of different cell purification methods on the GEP of the purified cell subset. We prepared model cell mixtures with different cell subset ratios and examined how well the original "pure" GEP could be recovered after purification of a given subset. We tested our findings on real biological samples. We analyzed CD34+ hemopoietic stem/progenitor sells purified from umbilical cord blood for these studies. Actual experimental samples often provide obstacles that are difficult to overcome by traditional cell purification methods. Very sensitive cells (like hepatocytes) or biohazard samples could be difficult to purify using conventional flow cytometric cell sorting or column-based immuno-magnetic cell purification. For this reason we evaluated two new, cutting edge cell purification and cell manipulation technologies for microarray experiments.

One major obstacle in gene expression profiling the majority of clinical and experimental samples, especially after cell purification, is that they do not provide sufficient amounts of RNA for direct microarray analysis. To overcome this problem RNA amplification is necessary. **3. To test whether it is possible to reconstruct a meaningful gene expression profile for a few, purified cells after RNA amplification**. To this end we validated two commonly used RNA amplification methods representing the two major approaches (exponential and linear amplification) in today's microgenomics. We analyzed their effects on the GEP and developed new strategies to interpret the produced amplified GEP data.

CHAPTER 2.

MATERIALS AND METHODS

CELLS AND CELL CULTURES

CEM Cells

Human, CD4+ T-cell line (acute lymphoblastoid leukemia); obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: CEM-T4 from Dr. J.P. Jacobs (Foley et al., 1965). CEM cells were cultured using RPMI 1640 medium with 2 mM L-glutamine and 10% fetal bovine serum in the presence of 5% CO_2 at 37 ^{o}C.

A2780 Cells

Human, CD4- epithelial cell line (human ovarian carcinoma, ECECC 93112519); kindly provided by Dr. Istvan Boldogh (Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, TX) A2780 cells were cultured using RPMI 1640 medium with 2 mM L-glutamine and 10% fetal bovine serum in the presence of 5% CO_2 at 37 ^{o}C.

KG-1a Cells

Human, CD34+ stem-cell cell line (human bone marrow acute myelogenous leukemia, ATCC CCL-246.1); kindly provided by Dr. Brian R. Davis (Sealy Center for Molecular Hematology and Oncology, University of Texas Medical Branch, Galveston, TX). KG-1a cells were cultured using Iscove's modified Dulbecco's medium with 4 mM L-glutamine and 20% fetal bovine serum in the presence of 5% CO_2 at 37 ^{o}C.

HeLa Cells

Human, epithelial cell line (cervix adenocarcinoma) kindly provided by Dr. Kui Li (Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, TX). HeLa cells were cultured using Eagle minimum essential medium with 2mM L-glutamine and 10% fetal bovine serum in the presence of 5% CO_2 at 37^{o}C.

RPMI 1640 medium, Iscove's modified Dulbecco's medium, Eagle minimum essential medium, fetal bovine serum, trypsin, and glutamine were purchased from Invitrogen (Carlsbad, CA).

Primary Hepatocytes

Mouse primary hepatocytes were isolated by the collagenase perfusion from either wild-type C57Bl/6 mice (The Jackson Lab, Bar Harbor, ME.) as described previously (107). Hepatocytes were cultured in Attachment media consisted of Williams E (Sigma, St. Louis, MO), 5% Fetal bovine Serum (Hyclone, Logan, UT), 1% Penicillin/Streptomycin (Invitrogen, Carlsbad, CA), and 100nM Insulin (Invitrogen, Carlsbad, CA). After the 4 hours of incubation, cells were plated in Growth media consisted of Williams E, 1% Streptomycin/Streptomycin, 100nM Insulin, and 2ng/ml endothelial growth factor (EGF; Invitrogen, Carlsbad, CA).

For experimental purposes, liver cells were plated on plastic 6 well plates (Corning, Acton, MA). Cell density ranged from $1x10^5$ to $1x10^6$ /per well. During the 4-hour attachment period, cells were infected with the adenovirus AdGFP (108) at a multiplicity of infection of 100 (based on infection of 293 cells), resulting in infection of about 50% of the hepatocytes. Cells were supplemented with growth medium following attachment, and were either maintained overnight at 37˚C in a 5% CO_2 humidified incubator, or were immediately stained with CellTracker Orange (CTO; Invitrogen, Carlsbad, CA) for LEAP experiments.

Cell culture medium, serum and glutamine were purchased from Gibco BRL, Grand Island, NY.

Model Cell Mixtures

Cultured CEM and A2780 cells were counted and tested for viability. Calculated volumes for 3.0×10^7 cells each were pelleted and resuspended in PBS (Gibco BRL, Grand Island, NY). Both cell suspensions were re-counted and appropriate volumes of each cell type for each planned mixture were calculated. The cell mixtures were then prepared by mixing the calculated volumes of each cell suspension.

Cord Blood Cells

Human cord blood was obtained under informed consent from HIV-negative normal donors under IRB-approved protocols at the Department of Obstetrics & Gynecology, Maternal-Fetal Medicine, University of Texas Medical Branch, Galveston, TX. Cord blood was drawn into yellow-capped vacutainer tubes (Beckman Coulter, Inc. Fullerton, CA), containing ACD anticoagulant. 2-4 blood samples were pooled and cord blood mononuclear cells (CBMCs) were isolated using Ficoll-Paque density gradient (Pharmacia Biotech, Piscataway, NJ) following manufacturer recommended protocols. CD34+ stem/progenitor cells were purified from CBMCs by magnetic sorting using MACS CD34 Progenitor Cell Isolation Kit (Miltenyi Biotec, Bergisch Gladbach, Germany) following manufacturer's recommendations.

CELL SORTING AND CELL PREPARATIONS FOR SORTING

Column-based Magnetic Cell Sorting

CD34+ cord blood stem/ progenitor cells were purified from CBMCs with MACS CD34 Progenitor Cell Isolation Kit (Miltenyi Biotec, Bergisch Gladbach, Germany) following manufacturer's recommendations. An aliquot of CD34+ KG-1a cells was processed with the MACS CD34 Progenitor Cell Isolation Kit exactly like the cord blood cells to monitor the effects of purification on the cells' GEP. For this purpose these "purified" cells were compared to unprocessed KG-1a cells. CD4+ CEM cells were purified from model cell mixtures with MACS CD4 Microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany) following manufacturer's recommendations. After three consecutive rounds of column purification, the purity (typically around 95-99%) of eluded cells was analyzed by subsequent flow-cytometric analysis of an aliquot labeled with anti-CD34-PE (for CD34+ cord blood cells and KG-1a cells) or anti-CD4-PE (for CD4+ CEM cells) antibodies (Caltag Laboratories. Burlingame, CA).

Flow Cytometry Analysis and Cell Sorting

CD34+ cord blood cells and KG-1a cells were labeled with phycoerythrin (PE)-conjugated, murine, anti-CD34 antibody (Caltag Laboratories. Burlingame, CA) using

factory recommended protocols. CEM cells and CEM/A2780 cell mixtures were similarly labeled with PE-conjugated, murine, anti-CD4 antibody (Caltag Laboratories. Burlingame, CA). Cells were analyzed and sorted on our custom-built High Resolution Cell Sorter (HiReCS) system (109) set up for standard fluorescence analysis. A tunable argon-ion laser tuned to the 488 nm wavelength was used in all analyses with optical filters that were optimal for PE excitation and emission. Samples were acquired on three parameters: PE, FSC (forward scatter), and SSC (side scatter) and stored as listmode data in standard FCS 2.0 format for subsequent analysis. The program WinList 5.0^{TM} (Verity Software House, Topsham, ME) was used for flow-cytometric data analysis.

Combined Magnetic and Flow Cytometric Cell Sorting

A CEM/A2780 cell mixture containing 10% CEM cells was pre-sorted by one round of magnetic sort using MACS CD4 Microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany) as described above. The resulting CEM-enriched cell mixture was labeled with PE-conjugated, murine, anti-CD4 antibody (Caltag Laboratories. Burlingame, CA) and flow-sorted for PE-positive cells. This sort enhanced CEM purity from 70% to 95% as shown by subsequent flow-cytometric analysis of an aliquot. The resulting 95% pure sample was processed by microarray analysis as purified CEM cells.

Cell Fixation

Fluorescent antibody-labeled cells were washed once in PBS (Gibco BRL, Grand Island, NY) and resuspended in 100-200 μL PBS. The sample was mixed with 500-1000 μL (5x the volume of PBS), -20^{o}C-cold methanol (Sigma, St. Louis, MO) and incubated at -20^{o}C for 5 minutes in the dark. Cells then were pelleted and resuspended in PBS for further processing.

LEAP Instrument

The Laser Enabled Analysis and Processing (LEAP) instrument (Figure 2.1.) platform has been previously described by Koller et al (90)**.** Briefly, a Q-switched, diode-pumped, solid-state, Nd:YAG laser (JDS Uniphase, San Jose, CA) was coupled with a novel, custom designed fluorescence imaging system (91). The average power output of this laser at 532nm is about 50mW. It pulses at a 1kHz frequency with a pulse width of 0.75ns, and peak power output of above 50kW at 532nm. The instrument was designed

with a custom achromatic F-theta scanning lens with a 12 mm field of view. Brightfield imaging is provided by light-emitting diodes, and epifluorescence excitation is provided by a halide/xenon hybrid lamp. Imaging is provided by two mega-pixel-γ-intensified CCD cameras. Custom software is used to direct the laser beam pulses at user-selected or auto-selected targets. The LEAP instrument is combining several cutting edge technologies (brightfield and fluorescent microscopy, high speed, broad-field optical imaging, high precision laser targeting, laser mediated cell-ablation, -catapulting, and -manipulation, a stepping motor system with microstepping capability, and multithreding computer technology) to be capable of cell visualization, cell elimination, and cell optoinjection. The general principal of LEAP is described in Chapter 1. The beta-prototype instrument housed in our laboratory included three interacting computers with three monitors, running several custom software packages, an optical imaging and targeting path with a broad band illuminating lamp, six filter wheels, several lenses, mirrors and prisms, two CCD cameras, a high energy, pulsed laser providing high peaks of power output in the UV, visible, and IR ranges of the spectrum, two high-speed galvanometer mirrors alternatively targeting the laser beam to the sample and the reflected light from the sample to the cameras, 16 stepping motors accurately handling all required movements of the sample holding platform. As shown on Figure 2.2. the key moving elements are the pair of fast-moving galvanometer mirrors that direct both the visual and the laser light to predetermined positions at a highly accurate fashion using alternative optical paths. These high-speed galvo mirrors, when combined with F-theta optics allow for large surface area imaging (up to 1 mm^2) without the need to move the stage for every view.

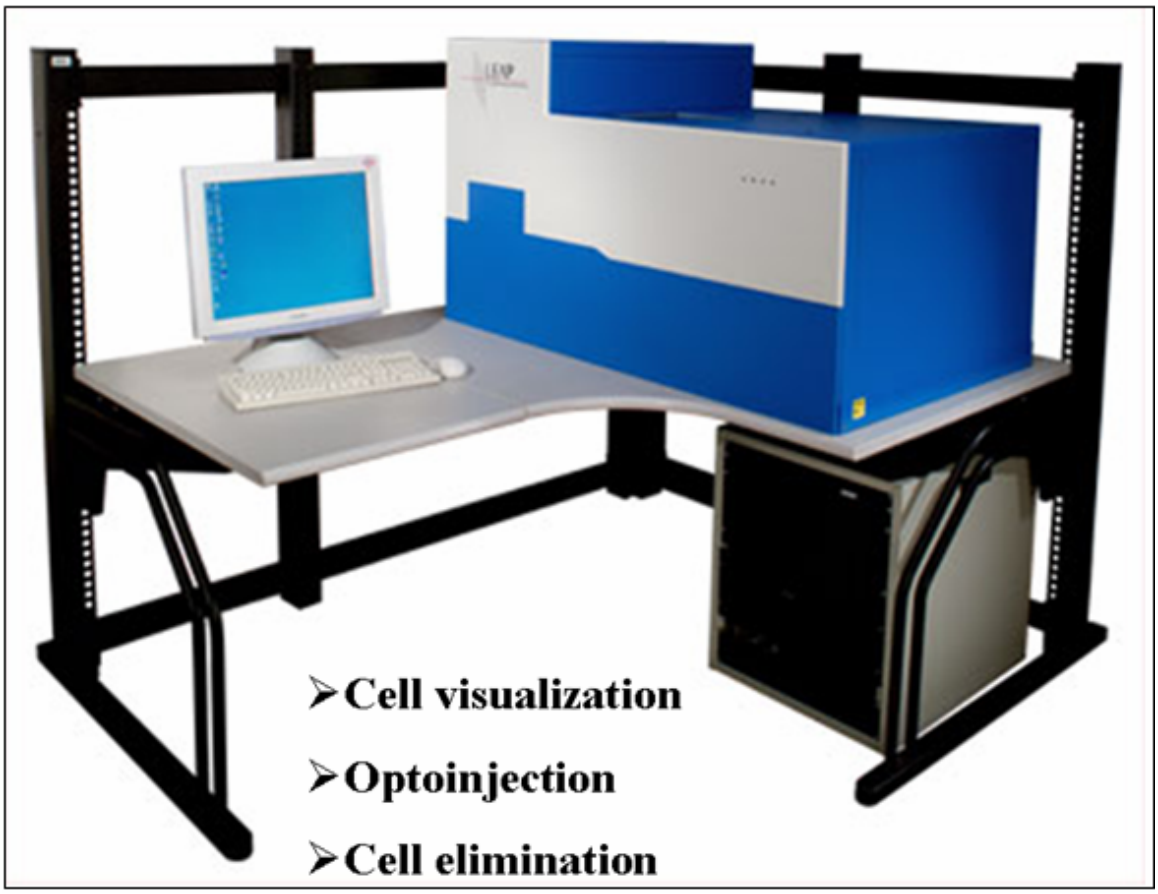

Figure 2.1. The commercial LEAP instrument

The LEAP instrument as it was envisioned by the manufacturer. The three main capabilities of the system are listed on the illustration. (Adapted from www.cyntellect.com)

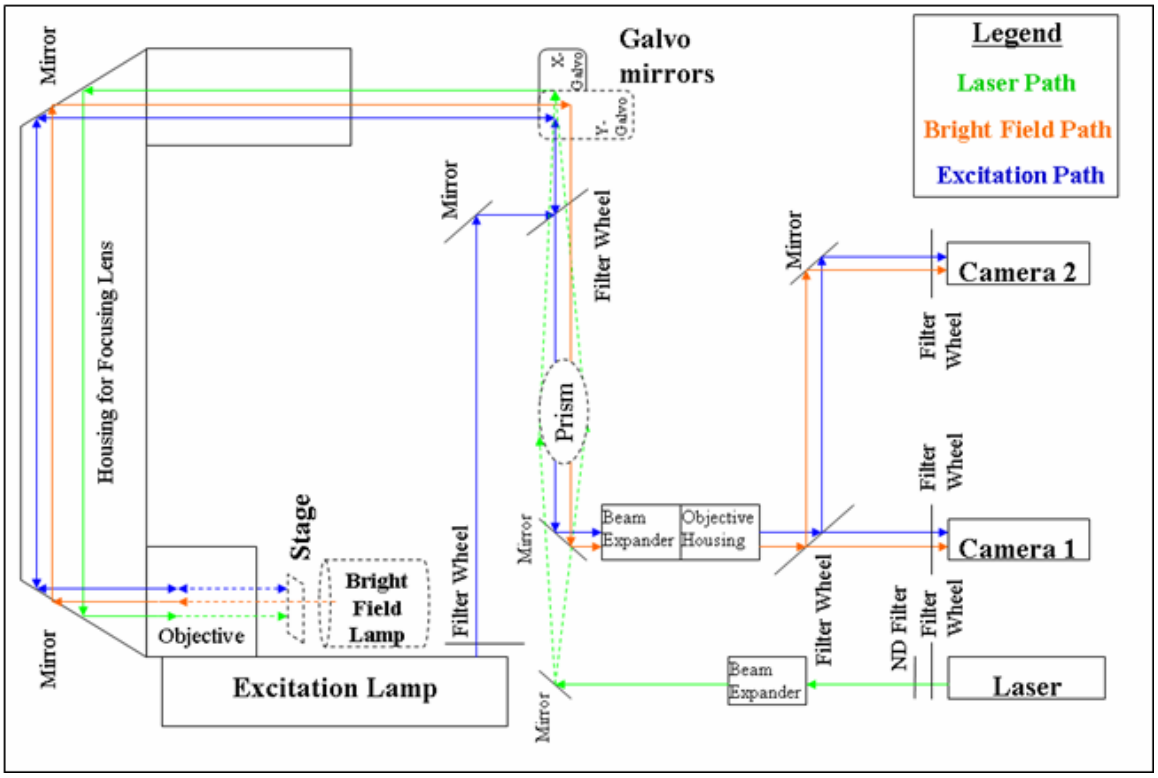

Figure 2.2. LEAP instrument optical paths

The orange arrows represent the optical path for brightfield imaging. Blue arrows represent the fluorescent imaging path. The laser beam follows the path represented by the green arrows. The sample is held on the Stage that can be moved by high accuracy stepping motors.

Cell Preparation for LEAP Processing

HeLa Cells for Ablation and Optoinjection Experiments

HeLa cells were trypsinized and plated to confluency in a Lab-Tek® II Chamber Slide™ System 8 Well Glass Slide (Nalge Nunc, Naperville, IL) using HeLa culture medium. Cells were incubated overnight in the presence of 5% CO_2 at 37°C. For ablation experiments, the cells were washed once with serum-free 199 medium (Invitrogen, Carlsbad, CA) and left in serum-free 199 medium. For optoinjection experiments the cells were washed once with serum-free 199 medium, then the medium was changed to serum-free 199 medium that contained Tetramethylrhodamine (TMR)-conjugated dextrans of MW=3kD, 10kD, 40kD, or 70kD (Invitrogen, Carlsbad, CA) at a concentration of 100µg/mL.

Hepatocytes for Optoinjection Experiments

Primary hepatocyte cultures comprising 50% adenoviral infection (GFP positive cells) were washed once and maintained in serum-free 199 medium in the original six-well tissue culture plate (Corning, Acton, MA) for LEAP. Following LEAP, cells were returned to Growth media for subsequent incubation and analysis.

CEM/KG-1a Model Cell Mixtures for LEAP Purification Experiments

CEM and KG-1a cells were counted and tested for viability using Trypan blue (Invitrogen, Carlsbad, CA) exclusion. 3.0 x 10^6 CEM cells were labeled with phycoerythrin (PE)-conjugated, murine, anti-CD4 antibody (Caltag Laboratories, Burlingame, CA) using factory recommended protocols. 3.0 x 10^6 KG-1a cells were labeled with FITC-conjugated, murine, anti-CD34 antibody (Caltag Laboratories, Burlingame, CA) using factory recommended protocols. Labeled KG-1a and CEM cells were pelleted and resuspended separately in serum-free 199 medium. Both cell suspensions were re-counted and appropriate volumes of each cell type for a 50% mixture were calculated. The 50% cell mixture was then prepared according to the calculations and plated at different densities in a Lab-Tek® II Chamber Slide™ System 8 Well Glass Slide. The plated slide was centrifuged at 400g for 10 minutes using an ALC PM140 centrifuge (ALC, Winchester, VA).

KG-1a Cells for Optoinjection Experiments

For optoinjection KG-1a cells were washed once with serum-free 199 medium then they were resuspended in serum-free 199 medium that contained TMR-conjugated dextrans of MW=3kD, 10kD, 40kD, or 70kD at a concentration of 100µg/mL. The cells were plated in this dextran-containing medium at high density in a Lab-Tek® II Chamber Slide™ System 8 Well Glass Slide. The plated slide was centrifuged at 400g for 10 minutes.

LEAP-Mediated Cell Purification

Samples were loaded into the LEAP instrument via the sample-loading platform utilizing the sample loading software. A small section of the bottom right corner of each well was used for empirically determining the proper laser settings for ablation and/or cell removal. The position of the beam waist of the laser was adjusted by changing the beam expander (BE) settings so that it was in the optimal position for ablation/cell removal for the given experiment. Similarly, the optimal laser power was adjusted using the LEAP neutral density filter (NDF). The optimal laser power was found when targeted cells in the calibration area were destroyed or moved from their original position, but untargeted cells were not affected. The optimal number of pulses and repeats was determined in earlier experiments and used here. For these experiments, a BE position of 0.6-1.0 (arbitrary units), NDF position 100-200 (arbitrary units), 1-5 pulses, and 1-3 repeats were determined to be the best conditions.

LEAP-Mediated Optoinjection

The same method was used as described for cell purification with the following modifications: once the optimal BE position was determined; the NDF setting was adjusted to decrease the laser power until no cell damage could be observed. The optimal conditions for these experiments were found to be a BE position of 0.8-1.0, NDF position 200-240, 1-5 pulses, 1-3 repeats, with a 4-second delay between repeats.

Analysis of LEAP Results

After LEAP processing, the cells were washed three times with the appropriate media for the cell type. In the ablation and optoinjection experiments of HeLa cells, samples were viewed utilizing a custom Diaphot Inverted Fluorescent Microscope (Nikon, Garden City, NY) and photographs were taken with a CoolPIX 990 Digital Camera (Nikon, Melville, NY). Dextran-containing HeLa cells were visualized utilizing a custom microscope fluorescent filter setup for TMR (Excitation: 525nm /Dichroic: 570nm /Emission 605nm). In the ablation/purification experiments of hepatocytes, and the KG-1a/CEM cell mixtures, the samples were visualized, pictures were taken, and images of the same view were overlaid using the LEAP instrument's optics, cameras, and image analysis software. PE-, FITC-, CTO-, and GFP-labeled cells were visualized using manufacturer recommended LEAP filter configurations. Single channel overlay images that were used to monitor cell movement due to LEAP-shooting, were re-colored using Paintshop Pro 6.00 (Jasc Software, Minneapolis, MN). Confocal images of optoinjected KG-1a cells were viewed and photographed using a Zeiss LSM 510 UV-META confocal imaging system. (Carl Zeiss, Oberkochen, Germany).

Multistage Magnetic Sorter (MagSort Instrument)

A prototype Multistage Magnetic Sorter was designed to implement the principle discussed in Chapter 1. A photograph of the prototype sorter used for these experiments is shown in Figure 2.3. with principal features called out. Central to the design is a pair of circular plates. The lower plate contains a well for sample insertion, and the upper plate contains a multitude of inverted, liquid-filled cavities, each of which forms, in its own turn, the top half of a shear cell when articulated with the sample well. The volume of each cavity is 0.48 ml. Fluids are inserted into each cavity through a fill port on the edge of the circular plate. This can be accomplished with the plates removed from the separator and can be done under a sterile hood, for example. The sample chamber is

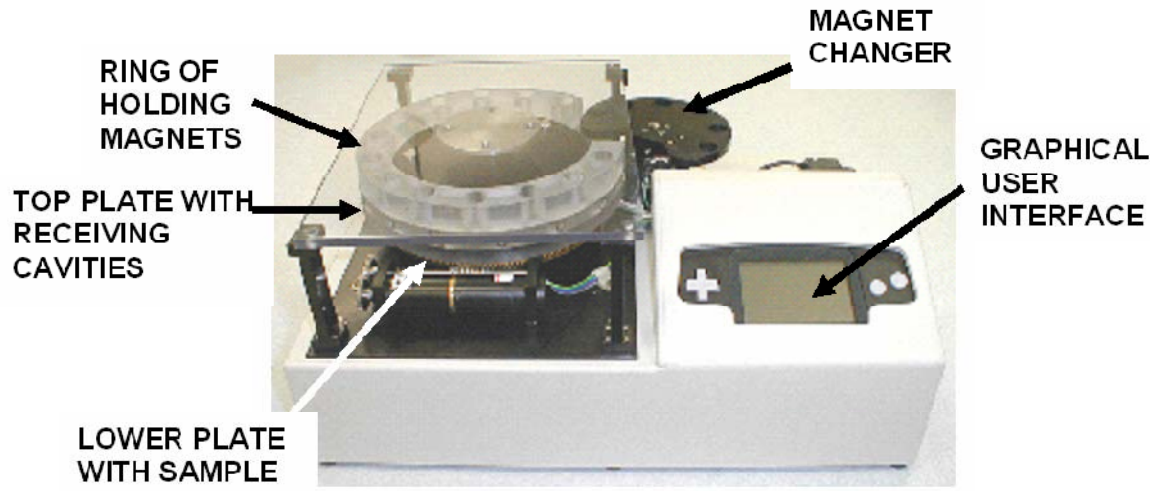

Figure 2.3. Key features of the multistage magnetic sorter

The key elements for cell sorting are a pair of horizontal circular plates, and a set of earth magnets above them. The upper plate rotates above the lower plate. The lower plate contains a well for sample insertion, and the upper plate contains the collection chambers, a multitude of inverted, liquid-filled cavities, each of which forms, in its own turn, the top half of a shear cell when articulated with the sample well of the lower plate. The volume of each cavity is about 0.5 ml. At each turn of the upper plate increasingly stronger magnets are placed above the magnetically labeled cells pulling them up from the sample well into the collection chambers. The footprint of the system is 10.5” x 14.0” (26.7 x 35.6 cm).

located in a single position under the magnet changer. The positions of the upper plate and the magnet changer are controlled by stepper motors driven by the on-board computer. The magnet changer rotates the selected magnet into position for each separation step. In the separator version used in our laboratory there were 15 separation chambers and 6 separation magnets providing the following magnetic fields: 68 mT, 172 mT, 277 mT, 377 mT, 445 mT, and 495 mT.

The multistage magnetic sorter is controlled by a Graphical User Interface (GUI). The graphics are shown on a color LCD screen where the user can access menus and program protocol timelines. The interface can store up to 10 unique timelines, each with up to 99 steps. Through the GUI, the user programs a protocol by using the navigation buttons (Figure 2.4.). Once saved, the programmed timeline is stored in non-volatile memory that is not affected by a loss of power.

Magnetic Particles for Sorting Experiments

Several samples of magnetic particles were obtained from vendors. Estapor particle types with different diameters (2.39 μm and 0.7 μm), both containing magnetite were purchased from EMD Biosciences Inc., Estapor Microspheres, Naperville, IL. Polysciences particles with 1-2 μm diameter were purchased from Polysciences, Inc., Warrington, PA, and Magsphere particles with 5.56 μm diameter were obtained from Magsphere, Inc., Pasadena, CA. Approximately 50% mixtures of two different particles were prepared for each sort experiment, the mixture was loaded into the sample insertion chamber of the Magsort instrument, and a preset sort program utilizing all 6 magnets was applied. Each fraction was collected after the sort and examined by microscopy for particle count. The individual experiments differed in slight details as described in the Results section of Chapter 4. The data presented represent, in each case, the results of a single reproducible experiment. Error limits in particle counts are assumed to follow a Poisson distribution. In all cases presented the relative standard deviation of particle counts was 10% or less.

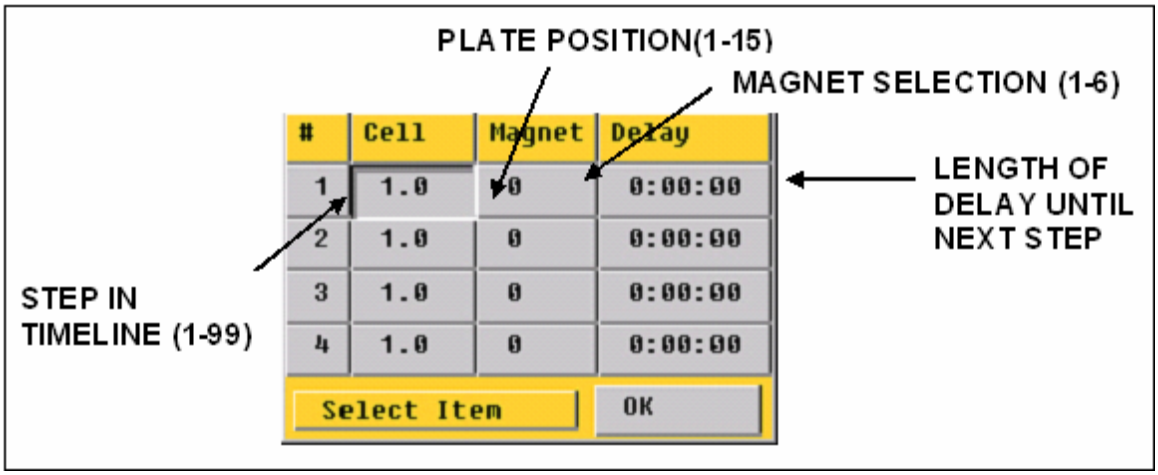

Figure 2.4. Screen shot of Graphical User Interface (GUI) for input of protocol timeline

The Magsort instrument is programmed through an LCD screen where the user can access menus and program protocol timelines to determine how long each magnet will be used for magnetic separation and which collection chamber (cell) will collect that particular fraction. The interface can store up to 10 unique timelines, each with up to 99 steps. The programmable features are labeled.

Model Cell Mixture Preparations for Multistage Magnetic Sorting Experiments

Four different types of immuno-magnetic particles were obtained from vendors. Two different Dynal beads with 4.5 μm and 1.0 μm diameter were purchased from Invitrogen Inc., Carlsbad, CA, and described in Chapter 4 (Figure 4.4.A and B). Bangs beads with and average diameter of 0.86 μm (Bangs Laboratories, Inc., Fishers, IN), and Miltenyi beads with 50 nm diameter (Miltenyi Biotec, Inc., Auburn, CA) are also further discussed in Chapter 4 (Figure 4.4.C and D). For cell mixture experiments approximately 50% mixtures of CEM and KG-1a cells were prepared with one of the cell types immuno-magnetically labeled and the other unlabeled. CEM cells were labeled with one type of immuno-magnetic bead for each experiment according to manufacturer's recommendations and anti-CD4, Alexa-488 conjugated antibodies (Invitrogen Inc., Carlsbad, CA) for visual confirmation of sort efficacy. KG-1a cells were also labeled with one type of immuno-magnetic bead for each experiment aimed to sort this cell type, according to manufacturer's recommendations and with anti-CD34, PE-conjugated antibodies (Caltag Laboratories. Burlingame, CA) for visual confirmation of sort efficacy. Further details of cell preparations for individual Magsort experiments are given in Chapter 4.

Multistage Magnetic Sorting of Model Cell Mixtures

In a typical application receiving culture medium was added to each cavity in the upper plate expected to receive sorted cells. The sample was then introduced into the sample chamber, the plates were installed on the sorter, and the sorter was assembled completely in preparation for a separation, including the placement of the selected magnets in selected positions on the automatic magnet changer. The cells in the sample were allowed to settle typically for 1 hour, so that the magnetized cells in the mixture would initiate their upward motion from the same starting position. Cells were allowed to settle either to the bottom of the sample chamber or to the top of a cushion of commercially available Percoll solution (Sigma Inc., Saint Louis, MO) upon which they floated at a common level above the physical bottom of the sample well. The 1-hour settling time was programmed into the Graphical User Interface, and the sorter was programmed so that each receiving cavity spent 10 or 11 minutes above the sample well

with a specified magnet above it. Typical automatic sample changing time was about 30 seconds, and a complete separation run would require about 2 hours. After a protocol was completed the plates were removed, and the collected fractions containing cells were unloaded for subsequent analysis by hemacytometer counting, flow cytometric or microscopic analysis. A typical sort program is displayed in Table 2.1.

Step #	Cell #	Magnet #	Delay Time
1	1	0	1 hour
2	2	1	11 min.
3	3	2	11 min.
4	4	3	11 min.
5	5	4	11 min.
6	6	5	11 min.
7	7	6	11 min.
8	8	0	0
9	1	0	0

Table 2.1. Magsort program for sorting immuno-magnetically labeled cells

The program is selected and modified by the Graphical User Interface of the Magsort instrument. Cell# stands for the number of sort chamber. Delay time is the time allowed for each sort chamber and the selected magnet to spend above the sample insertion chamber. Steps #8 and #9 are the required code for the instrument to end a sort run.

MICROARRAY ANALYSIS

Preparation of Labeled Probes and Microarray Analysis

Total RNA was isolated from all cell samples using RNAqueous™-4PCR RNA isolation kit (Ambion Inc, Austin, TX) following manufacturer's recommendations. Within each experiment each sample was normalized by the amount of isolated RNA. For spotted microarray analysis of 82 genes an Atlas™ Array Trial Kit (Clontech Lab., Palo Alto, California) was used following factory recommended protocols. For signal detection a Storm 860 phosphorimager (Molecular Dynamics, Sunnyvale, CA) was used. The created array images were analyzed by Scanalyze software (Stanford University, Stanford, CA) to quantitate microarray data and produce the raw data files. For short-oligo microarray analysis of over 12,000 genes a GeneChip® Human Genome U95Av2 (Affymetrix, Santa Clara, CA) was used. RNA labeling, hybridization, and scanning to produce the raw microarray data files were done by the Molecular Genomics Core

Facility of the University of Texas Medical Branch at Galveston, following factory recommended protocols.

Microarray Images of Cell Mixtures

We selected two cell lines (CEM and A2780) with characteristic similarities and differences in their GEPs to be able to monitor their contributions to the overall GEP of each sample. First we looked at the direct microarray images of the samples to gain an overall impression of the situation. Throughout this study we always inspected the same selected segment of the Affymetrix array for each sample to visually confirm our findings. The original array images were magnified, pseudo-colored and cropped using the Affymetrix Microarray Suite software. Each small square (feature) represents a gene sequence with millions of identical 25-mer oligonucleotides attached to the array surface. The segment we used for visual analysis contained a total of 165 features. Different colors represent different expression levels increasing in the following order: black, violet, blue, green, yellow, orange, red, and white.

Microarray Data Analysis

Images of Affymetrix arrays were generated directly from the raw image files. The original array images were viewed, magnified, pseudo-colored and cropped using the Affymetrix Microarray Suite 5.0 software (Affymetrix, Santa Clara, CA). For array-to-array image comparison always the same selected segment of the Affymetrix array was inspected and compared. Images of Clontech arrays, generated by the Storm 860 phosphorimager (Molecular Dynamics, Sunnyvale, CA) were directly used for visual array-to-array image comparison. Bar graphs, scatterplots and regression analysis results were generated from the raw data files using Microsoft Excel 2000 software (Microsoft, Washington, United States). Trellis plots were generated using S-Plus 6.2 software (Insightful Corp., Seattle, WA). Hierarchical clustering was generated using Spotfire 7.2 data analysis software (Spotfire Inc., Cambridge, MA) For generating the heat map, raw Affymetrix data files were trimmed by excluding 2% of all genes with the highest and 2% with the lowest intensity readings. The array data from the four samples were normalized by trimmed mean overall expression level. Genes that were called “absent” in all four arrays by the Affymetrix software and all genes with less than 2-fold expression

level difference in all the possible pair-wise comparisons among the four arrays were excluded. The remaining genes were organized into a heat map by hierarchical clustering based on gene expression levels

RNA AMPLIFICATION

Exponential RNA Amplification

Total RNA was isolated from all cell samples using RNAqueous^{TM}-4PCR RNA isolation kit (Ambion Inc, Austin, TX) following manufacturer's recommendations. The isolated RNA was amplified by SMART mRNA amplification kit (Clontech Laboratories Inc., Mountain View, CA) following manufacturer's recommendations. The general principle of exponential RNA amplification is described in Chapter 1. Briefly, from the isolated total RNA the mRNA fraction was first reverse transcribed into double stranded cDNA using oligo-dT primers to select for polyA-tailed mRNAs. The reverse transcription was also used to attach a universal tag-sequence to each cDNA that served as an annealing sequence for the primers in the next step. The produced cDNA was PCR amplified using the universal tag-sequence for priming. After 15 PCR cycles and from this point on after every 3 additional cycles the amount of the produced cDNA was quantified to determine the optimal cycle number. From 10^4 cells 18 PCR-cycles generated equivalent amounts of cDNA to an unamplified, 10^6 cell sample. From 10^3 cells 24 PCR cycles were necessary. The amplified cDNA was purified and used for microarray analysis.

Linear RNA Amplification

Total RNA was isolated from all cell samples using RNAqueous^{TM}-4PCR RNA isolation kit (Ambion Inc, Austin, TX) following manufacturer's recommendations. The isolated RNA was amplified by MesageAmp aRNA kit (Ambion Inc, Austin, TX) following manufacturer's recommendations. The general principle of linear RNA amplification is described in Chapter 1. Briefly, from the isolated total RNA the mRNA fraction was first reverse transcribed into double stranded cDNA using T7-oligo-dT primers to select for polyA-tailed mRNAs. The reverse transcription was also used to

introduce a T7 promoter-sequence into each cDNA that served as signal sequence for the T7 RNA-polymerase in the next step. The produced cDNA was amplified by in vitro transcription utilizing T7 RNA-polymerase. The produced antisense RNA (aRNA) was purified and used for microarray analysis or as a starting sample for a second round of linear RNA amplification using the same MesageAmp aRNA kit when it was necessary (typically when the original sample contained less than 1000 cells).

CHAPTER 3.

GENE EXPRESSION PROFILE ANALYSYS OF CELL MIXTURES AND PURIFIED CELL SUBSETS

INTRODUCTION

The objective of the experiments described in this chapter was to test whether purifying the cells of interest prior to gene expression profiling enhances the sensitivity and accuracy of microarray analysis. Information collected from a sample can only be as good as the sample itself. If the purity of the target cell population in a sample is low, gene expression microarray analysis of this sample might provide more information about the contaminating cell types than about the investigated cell subset. We designed a set of experiments to study the reproducibility, sensitivity, and accuracy of both main types of microarray technology as well as to investigate the effects of cell purification and sample handling on the gene expression profile (GEP). Most of the results presented here have been published (63).

Gene expression microarray technology (MAT) is potentially capable of examining all of the cellular processes at the mRNA level at a given moment (10-12). This approach delivers a comprehensive quantity of data about the transcriptional level of the cell, providing "freeze-frame" views of the transcriptome (12-14). However, initial MA studies often caused disappointment, failing to generate or decisively support new hypotheses. In many occasions MAT has generated more confusion than comprehension, more questions than answers (20,31,32). Despite the fact that MAT has improved greatly in recent years and most of the early problems (reproducibility, sensitivity, high background, standardization, preparation of samples, data analysis) have been addressed, this new technology is still struggling to break out of a relatively narrow field of applicability. One of the main sources of the remaining problems is that, while MA analysis requires 0.5-5 million cells per sample, biological systems (tissues, organs) with this amount of cells are almost always mixtures of several different cell types (43) - all of

which may behave differently in a given experiment. Even major changes in gene expression levels of a minor cell subset of the mixture might be lost in the background of the more numerous unchanged cells.

The obvious solution to the problem that cell mixtures display mixed gene expression profiles (GEPs) is cell sorting. Delivering sorted, more homogenous cell samples to MAs is expected to produce much clearer results than studying cell mixtures. Several recent studies have taken this approach and many of them have presented promising results (55-57). However, it is still not well understood how different cell sorting methods and sample handling protocols affect the GEP (31,58,59). It is also not well known how much effect different cell types have on each other's GEP when they are mixed together. On the other hand, in the real world of both basic and clinical applications, 100 percent purity of a given cell type is not always achievable or feasible. So the question is, how pure is pure enough?

Any purification method takes time and mRNA is a 'moving target', with possible degradation during the experiment. RNAs can be produced very rapidly and some mRNAs may be degraded in a matter of minutes in live cells, while others may be stable over several hours (58,110). Is it possible to 'freeze the GEP in time' until the cells get purified and delivered to the microarray? Most purification methods require labeling of the target cells. How much will the labeling and the purification process alter the GEP of the sorted cells?

To address these very basic questions, we studied the overall GEP of defined cell mixtures to model heterogeneous biological samples. For the "overall GEP" we used the unprocessed microarray readout of the cell sample with no values excluded. We evaluated the effects of cell labeling, fixation and sorting on the overall GEP. We also analyzed how well we could recover the GEP of a pure cell type by sorting these cells from a mixture. Since different types of microarrays do not necessarily produce the same data (13,35,36), we used both spotted (Clontech, Palo Alto, CA) and short-oligonucleotide arrays (Affymetrix, Santa Clara, CA) to compare the results of the same experiments on these different platforms.

Microarray data analysis is not an obvious exercise; in fact it has developed to be a new field of its own with quite a few competing methods (111,112). The complexity of these methods often creates a communicational gap between the data-producing biologists and the data-analyzing mathematicians and bio-statisticians (13,14,113). To avoid this gap and to keep the presentation of our results as directly connected to the samples they represent as possible, we have used very simple, straightforward methods to compare the overall GEPs of different samples.

RESULTS

Microarray Images of Pure Cell Samples

We selected two cell lines (CEM and A2780) with characteristic similarities and differences in their GEPs. This enabled us to monitor the contribution of each cell line to the overall GEP of each mixed sample. First we developed a direct method to visually compare microarray results of different samples. We examined the magnified microarray images of the samples to gain an overall impression of the situation. We selected a well identifiable segment of the GeneChip® Human Genome U95Av2 microarray from Affymetrix. Throughout this study we compared the same segment of the microarray for each sample to visually confirm our findings. Figure 1A displays the Affymetrix microarray images of the two cell types. The original array images were magnified, pseudo-colored and cropped using the Affymetrix Microarray Suite software. Each small square (probe cell) represents a gene sequence with millions of identical 25-mer oligonucleotides attached to the array surface. The segment we used for visual analysis contained a total of 165 probe cells. Different colors represent different expression levels increasing in the following order: black, violet, blue, green, yellow, orange, red, and white.

The two images in Figure 3.1. exhibit very similar patterns (both samples being human cell lines), but several differences can be found between them. Some sequences are only expressed in one or the other (circles 1 and 2). Both in circles 1 and 2 we

included a reference gene sequence as well that is fairly evenly expressed in the two samples. Many genes are expressed in both samples but at different levels (square 3). Differentially expressed genes are displayed in different colors in the microarray image due to pseudo-coloring. Triplicate arrays of the same sample produced virtually identical images.

Figure 3.2. displays the spotted microarray images of the same two cell lines (CEM and A2780). Each pair of spots represents a gene sequence, 96 sequences in all on each spotted array. Larger and darker spots represent higher gene expression levels. Again, within the very similar gene expression patterns there are several characteristic differences with some genes only expressed in one of the cell lines (circles 1, 2, and 4), others showing differential gene expression (square 3). Spotted array images of replicate samples were virtually identical. These observations confirmed that the GEPs of the two chosen cell lines are indeed different for many selected genes, and the differences between those selected gene expression levels can be visually detected in the images of both array types.

The bottom row of spots represents a set of "housekeeping genes" on this array. These genes are thought to be expressed at a fairly consistent level throughout the cell cycle and from tissue to tissue. They are often used as standards for comparison when studying expression of other genes of interest. In our experiments even these genes display visible differences between their expression levels (Fig. 3.2., circle 4) when comparing CEM and A2780 cells, indicating that using a randomly selected "housekeeping gene" as internal standard in Northern blots or other experiments measuring and comparing RNA expression levels throughout different samples might lead to false results.

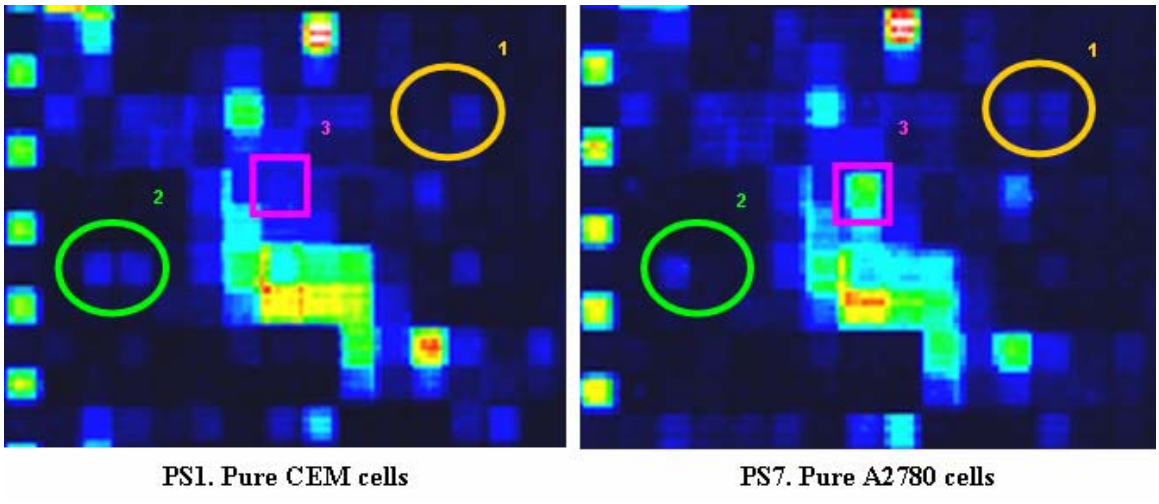

Figure 3.1. Affymetrix microarray images of CEM and A2780 cells.

Each small square represents the expression level of an individual gene sequence. Examples for gene sequences expressed only in one of the cell types marked with ovals 1 and 2. The sequence to the left in oval 1 is only expressed in A2780 cells, the sequence to the right in oval 2 is only expressed in CEM cells. A differentially expressed gene sequence with much higher expression level in A2780 cells than in CEM cells is visible in square 3. (To visualize the raw image data of microarray analysis, corresponding segments of Affymetrix images were cropped, magnified and pseudo-colored using the Affymetrix Microarray Suite 5.0 software.)

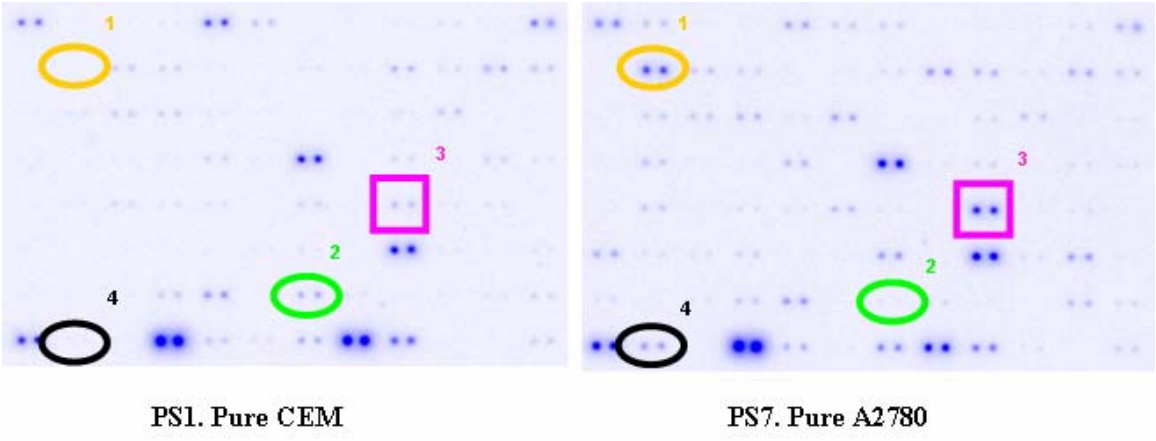

Figure 3.2. Clontech spotted array images of CEM and A2780 cells.

Each pair of spots represents the expression level of an individual gene sequence. Examples for gene sequences expressed only in one of the cell types are marked with ovals 1, 2 and 4. The sequences in ovals 1 and 4 are only expressed in A2780 cells; the sequence in oval 2 is only expressed in CEM cells. A differentially expressed gene sequence with much higher expression level in A2780 cells than in CEM cells is visible in square 3. Gene sequences in the bottom row – including the one in oval 4 – are representing "housekeeping genes". (The entire image of the Clontech Atlas Trial microarray membrane containing duplicates of 96 different gene sequences is displayed for each cell type.)

Microarray Images of Cell Mixtures

We prepared a set of cell mixtures from CEM and A2780 cells with gradually changing cell ratios. Figure 3.3. displays the same Affymetrix array image segments of seven cell mixtures. Sample PS1 contained pure CEM cells, the CEM/A2780 ratio in PS2 was 90:10, 75:25 in PS3, 50:50 in PS4, 25:75in PS5, 10:90 in PS6, and PS7 had pure A2780 cells. These gradual changes in the mixture ratio could be detected as gradually appearing and strengthening genes (circle 1), or progressively weakening and disappearing genes (circle 2) in the consecutive Affymetrix images corresponding to the CEM/A2780 cell ratio changes throughout the samples (CEM: 100%-90%-75%-50%-25%-10%-0%). Genes that were differentially expressed in the two cell lines also displayed gradually changing expression levels throughout the seven samples (square 3).

Figure 3.4. displays that the same tendencies could be detected using the spotted array images. This direct, visual analysis of the seven microarrays representing samples PS1-PS7 suggested that the GEP of a cell mixture quite accurately mirrored the ratios of the participating cell populations.

Scatterplot Analysis of Replicate Pure Cell Samples

After the above described 'intuitive approach' of visually inspecting expression levels of a few single genes throughout all samples we examined the 'overall GEP' of the same samples. For a comprehensive overall comparison of two GEPs we used simple scatterplots where each sample's GEP was represented on one axis, and each gene (small squares) was positioned using its expression levels in the two samples as coordinates. The two samples' GEPs were considered to be more similar when the genes lined up closer along a linear trendline using multiple linear regression analysis. To calculate this GEP similarity for each scatterplot we used, as is commonly done, the R^2-value (multiple correlation coefficient) as a measure of 'goodness-of-fit'. Higher R^2-value was considered to indicate higher similarity between two GEPs.

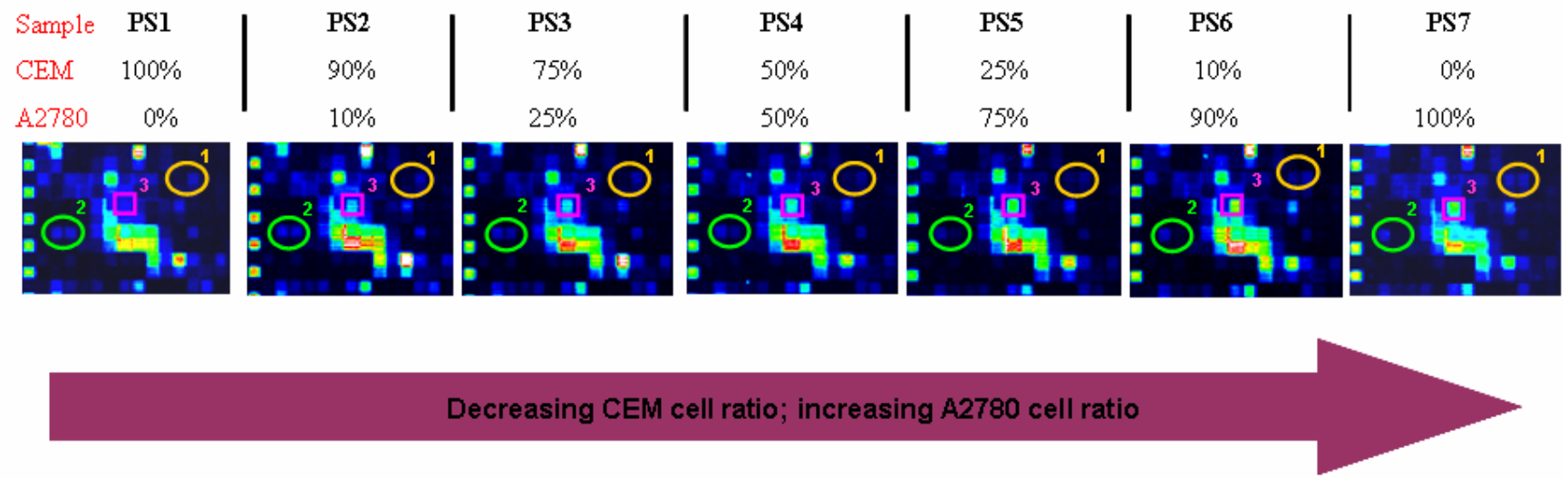

Figure 3.3. Affymetrix microarray images of CEM / A2780 cell mixtures.

Corresponding segments of Affymetrix microarray images for seven samples with different CEM / A2780 cell ratios are displayed. Each small square within the images represents the expression level of an individual gene sequence. Examples for gene sequences expressed only in one of the cell types are marked with ovals 1 and 2. A differentially expressed gene sequence with much higher expression level in A2780 cells than in CEM cells is visible in square 3. As the CEM cell ratio decreases and the A2780 cell ratio increases a gene sequence in oval 1 appears and gradually strengthens. Oval 2 marks a gene sequence that disappears as the CEM cell ratio decreases despite the increasing A2780 cell ratio. The signal intensity for a gene sequence in square 3 gradually increases.

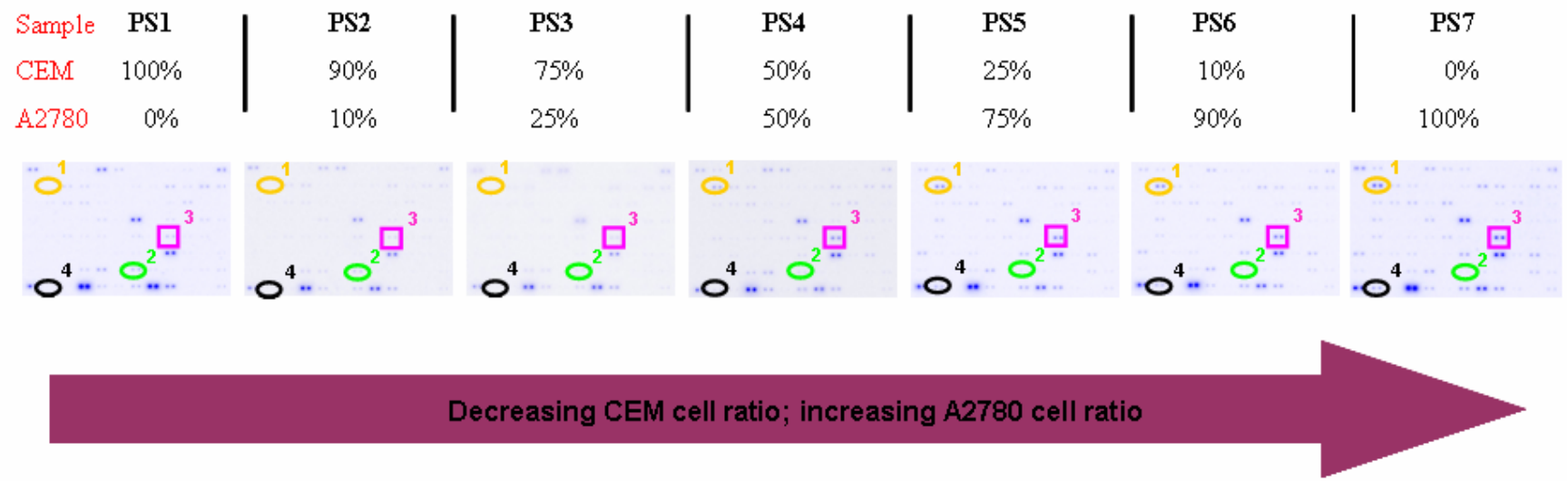

Figure 3.4. Clontech spotted array images of CEM / A2780 cell mixtures.

Clontech Atlas Trial microarray images for seven samples with different CEM / A2780 cell ratios are displayed. Each pair of spots represents the expression level of an individual gene sequence. Examples for gene sequences expressed only in one of the cell types are marked with ovals 1, 2, and 4. A differentially expressed gene sequence with much higher expression level in A2780 cells than in CEM cells is visible in square 3. As the CEM cell ratio decreases and the A2780 cell ratio increases a gene sequence in ovals 1 and 4 appear and gradually strengthen. Oval 2 marks a gene sequence that disappears as the CEM cell ratio decreases despite the increasing A2780 cell ratio. The signal intensity for a gene sequence in square 3 gradually increases.

Figure 3.5. displays pair-wise scatterplots of triplicate samples (pure A2780 cells). GEPs were produced by Affymetrix arrays. As expected, triplicate GEPs of over 12,000 genes demonstrated high similarity. Repeated triplicate experiments consistently produced R^2-values of 0.97 to 0.995. Replicate spotted array results for 96 genes per sample were similar. Based on these results, we considered the overall GEP of samples with R^2-values over 0.97 identical since this difference is within the error range of microarray analysis.

Scatterplot Analysis of Cell Mixtures

After establishing the approximate range of experimental error we examined the overall GEPs of our model cell mixtures. Figure 3.6. compares samples PS1 and PS2, a pure CEM sample to a 90% pure one. With an R^2-value of 0.9842 the overall GEPs of these two samples seemed to be just as identical as if they were the same sample. Similarly the R^2-value for comparing pure A2780 cells to 90% pure ones (PS7 versus PS6) was found to be 0.9826. Obviously, this does not mean that the expression levels of each individual gene in these samples are practically the same, but it does suggest that for the vast majority of genes a 90% pure sample could be used as ‘pure sample' in terms of an "overall" GEP.

Comparing the pure CEM sample to the pure A2780 sample (Figure 3.6. B.) produces a scatterplot with an R^2-value of 0.7532. Although we selected the two model cell lines to be very different (CEM is a human, T-lymphoblastoid cell line while A2780 cells are derived from human ovarian carcinoma cells), both are human tumor cells with significant similarities in their GEPs. In our model this R^2-value of 0.7532 is ‘as bad as it gets’. This value probably corresponds to the degree of similarity one sees with very different human cell line types due to the necessary expression of housekeeping and fundamental genes needed for growth and survival in immortalized cells.

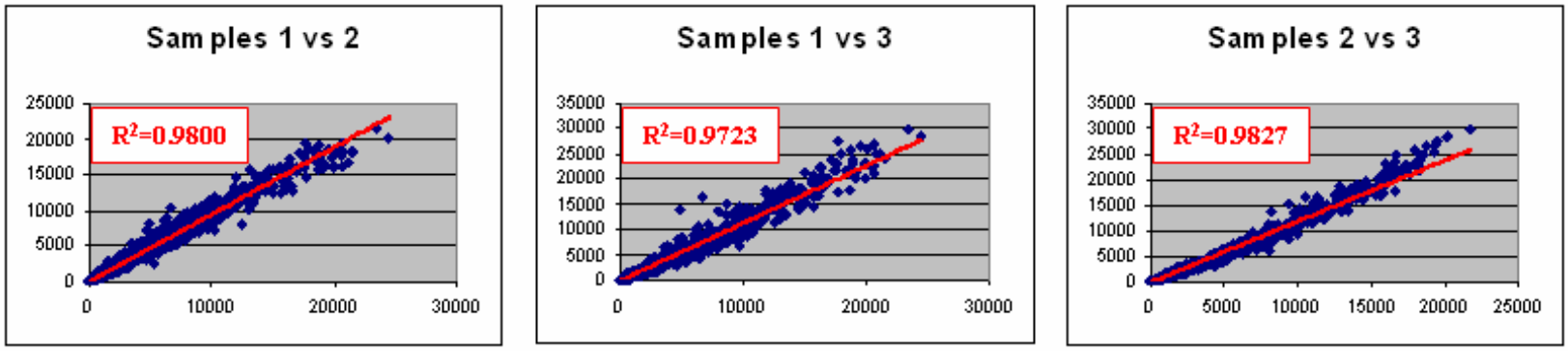

Figure 3.5. Scatterplots of Affymetrix results from triplicate A2780 cell samples

Triplicate samples of A2780 cells were processed and analyzed separately. The expression levels of over 12,000 genes were compared between samples. High R^2 values correspond to high correlation between gene expression levels within the triplicate samples.

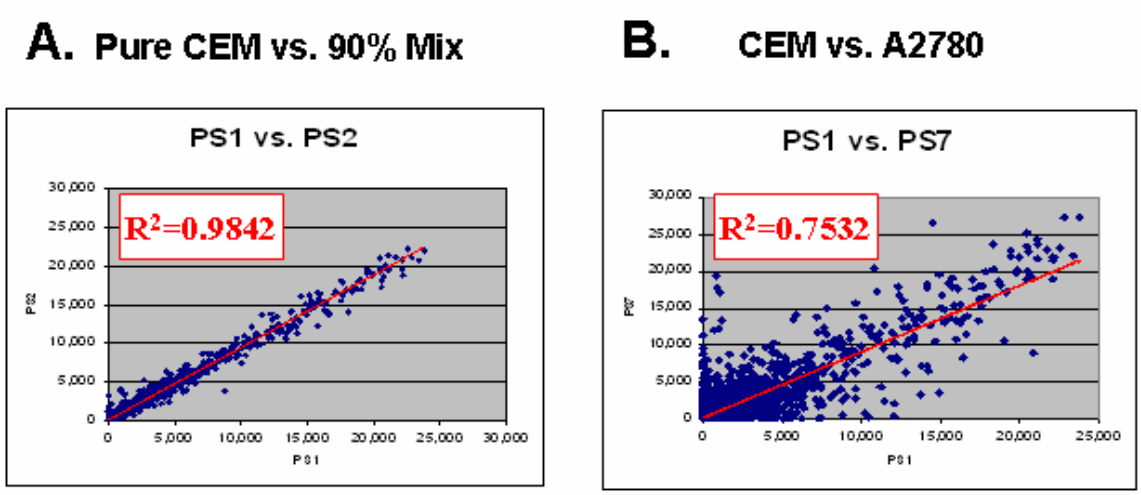

Figure 3.6. Scatterplots of Affymetrix results from different cell lines and cell mixtures

Pure and mixed cell samples were processed and analyzed separately. The expression levels of over 12,000 genes were plotted. Lower R^2 value corresponds to lower correlation between gene expression levels between CEM and A2780 cell samples. Samples: PS1=100% CEM; PS2=90% CEM+10% A2780; PS7=100% A2780

Trellis Plot Analysis of Cell Mixtures

Figure 3.7. displays all possible pair-wise scatterplots of the seven samples in a Trellis plot format. Each scatterplot is positioned in the row and column of the two samples from which it is made (e.g. the second scatterplot in the fourth column is comparing samples PS2 and PS4). By following along a given row or column it is visually easy to see the effect of purity of cell mixtures on GEPs. The overall pattern was strikingly similar to the one observed in Figures 3.3. and 3.4. When calculated the R^2-values of all these scatterplots (data not shown) we found that they also followed the same tendency; the farther two samples were from each other in the Trellis plot the lower R^2-value they produced when compared. Notably, the 75% mix compared to the pure sample (PS3 versus PS1 and PS5 versus PS7) produced R^2-values in the 0.96-0.975 range which was established earlier as the estimated borderline for this method to be able to 'tell the difference" between two samples. However, a very small number of genes were found to fall farther away from the regression line. These outlier genes displayed significantly different expression levels within the otherwise almost identical samples.

All scatterplots where the difference in the cell ratios of the two samples compared were only 25% or less produced R^2-values of 0.96 or higher. Since these R^2 values were all within the sensitivity limits of microarray technology in our hands, based on these sample comparisons we considered the overall GEPs of the compared samples indistinguishable. These samples could be considered identical in terms of overall GEP. We noted that a few genes did display significantly different expression levels (outlier genes).

All scatterplots where the two samples compared were more than 25% different in cell ratios had much lower R^2-values. We considered the overall GEPs of the two samples in such comparisons to be significantly different. These results suggest that a sample needs to be at least 75% pure to truly represent the overall GEP of the given cell population.

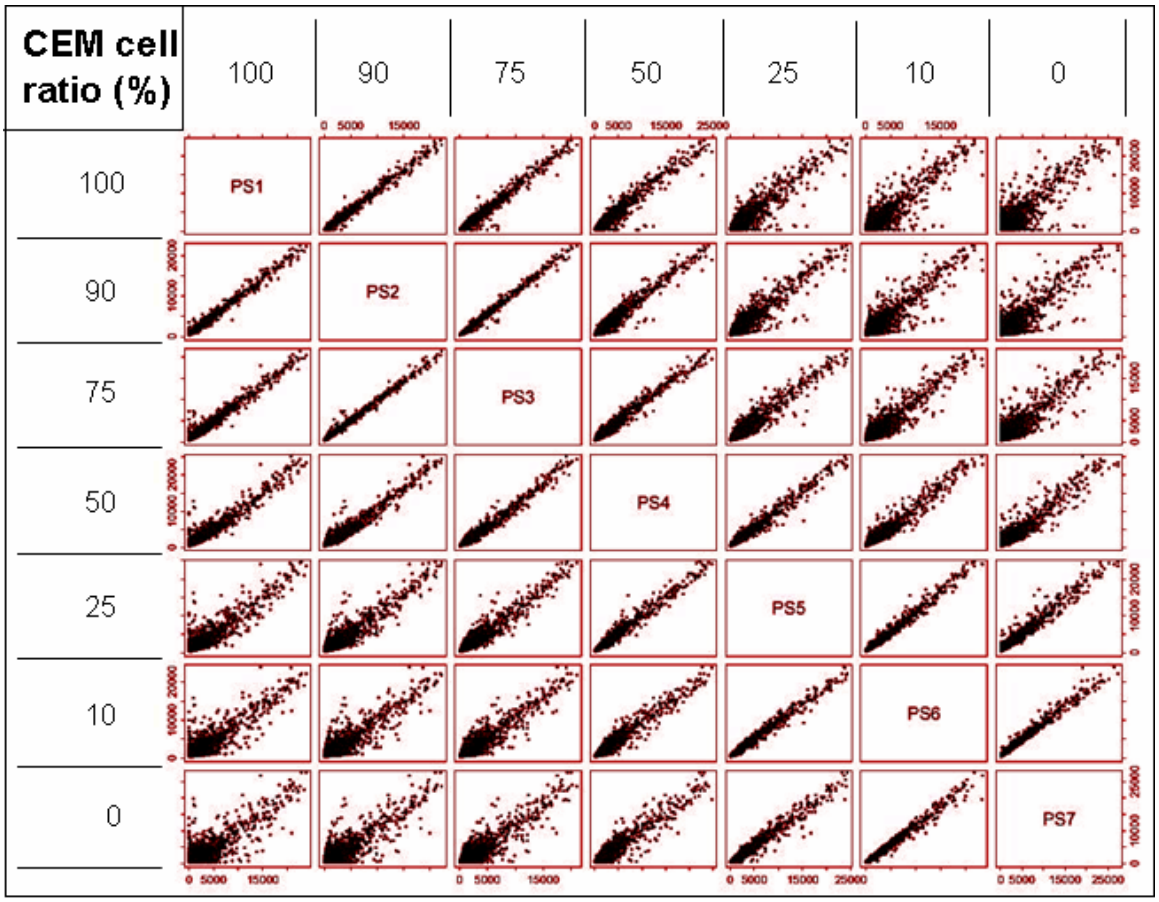

Figure 3.7. Ratio effect of cell mixtures on overall gene expression profiles – Trellis-plot chart

Pure and mixed CEM and A2780 cell samples were processed and analyzed separately. The expression levels of over 12,000 genes were plotted pair wise for all 7 samples. Correlation between cell ratios and overall gene expression profile was visualized using a Trellis-plot chart to emphasize gradually growing differences reflecting changing cell ratios. Each scatterplot is positioned in the row and column of the two samples it is made of.
Samples: PS1=100% CEM; PS2=90% CEM+10% A2780; PS3=75% CEM+25% A2780; PS4=50% CEM+50% A2780; PS5=25% CEM+75% A2780; PS6=10% CEM+90% A2780; PS7=100% A2780

Modeling Genes of 'Real Biological Samples'

After assessing the overall GEPs of cell mixtures we modeled the behavior of individual genes in a possible 'worst case scenario in real biological samples' where the background cells would strongly express the investigated genes. We selected a set of individual genes that were moderately expressed (intensity range: 100-1,000 units) in CEM cells (target cells) and strongly expressed (intensity range: 3,000-12,000 units) in A2780 cells (background cells). As presented in Figure 3.8., the expression level of these genes followed the ratio of A2780 cells in the mixtures. Although CEM cells also expressed these genes, their effect on the sum expression levels was totally washed out by the 'background cells". To monitor changes in the expression levels of these genes in CEM cells even a 90% pure sample was not pure enough.

A possible 'best case scenario' was modeled in Figure 3.9. Again, we selected another set of genes that were moderately expressed (intensity range: 100-1,000 units) in CEM cells, but A2780 cells did not express these genes at all (intensity range: 5-20 units). Now it was possible to monitor the expression levels of these genes in the 'target cells' even in the presence of 9 times more 'background cells'. The charts in Figures 3.8. and 3.9. demonstrate Affymetrix data; very similar charts could be created from spotted array results.

These results suggested that for certain individual genes the "75% rule" we established for the overall GEP does not apply, studying some "worst case" genes might require much higher purity while investigating some "best case genes" might not require cell purification at all.

The remarkable accuracy at which the microarray results for both the overall GEP and individual genes corresponded with the cell ratio of the studied cell mixture throughout the experiments presented so far, confirmed that the gene expression profiles for mixed cell populations are, as expected, the combined expression profiles for each cell subpopulation weighted according to its relative frequency in the cell mixture.

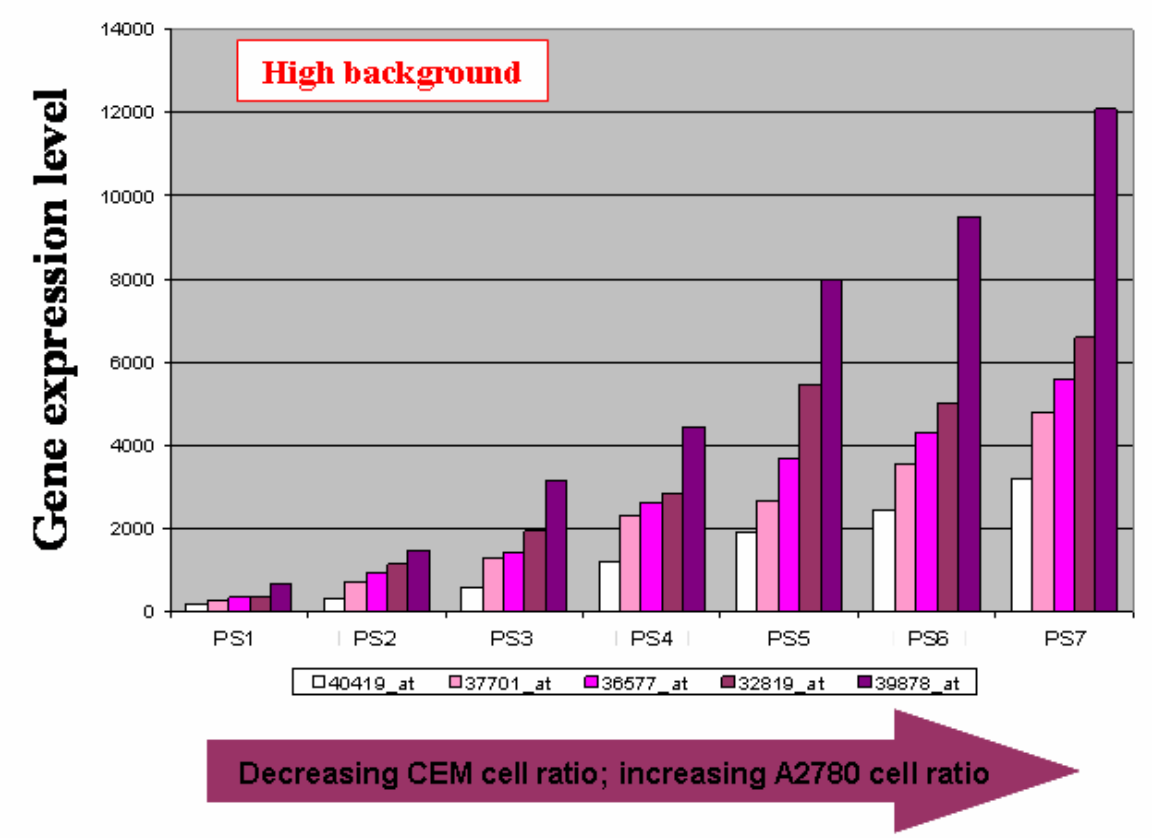

Figure 3.8. Ratio effect of cell mixtures on individual gene expression levels – Moderately expressed genes in the presence of high background

Pure and mixed CEM and A2780 cell samples were processed and analyzed separately using Affymetrix GeneChip® Human Genome U95Av2 array. Five individual genes with moderate expression levels (intensity range: 100-1,000) in CEM cells and high expression levels (intensity range: 3,000-12,000) in A2780 "background" cells were plotted throughout 7 samples with decreasing CEM cell ratios.
Samples: PS1=100% CEM; PS2=90% CEM+10% A2780; PS3=75% CEM+25% A2780; PS4=50% CEM+50% A2780; PS5=25% CEM+75% A2780; PS6=10% CEM+90% A2780; PS7=100% A2780

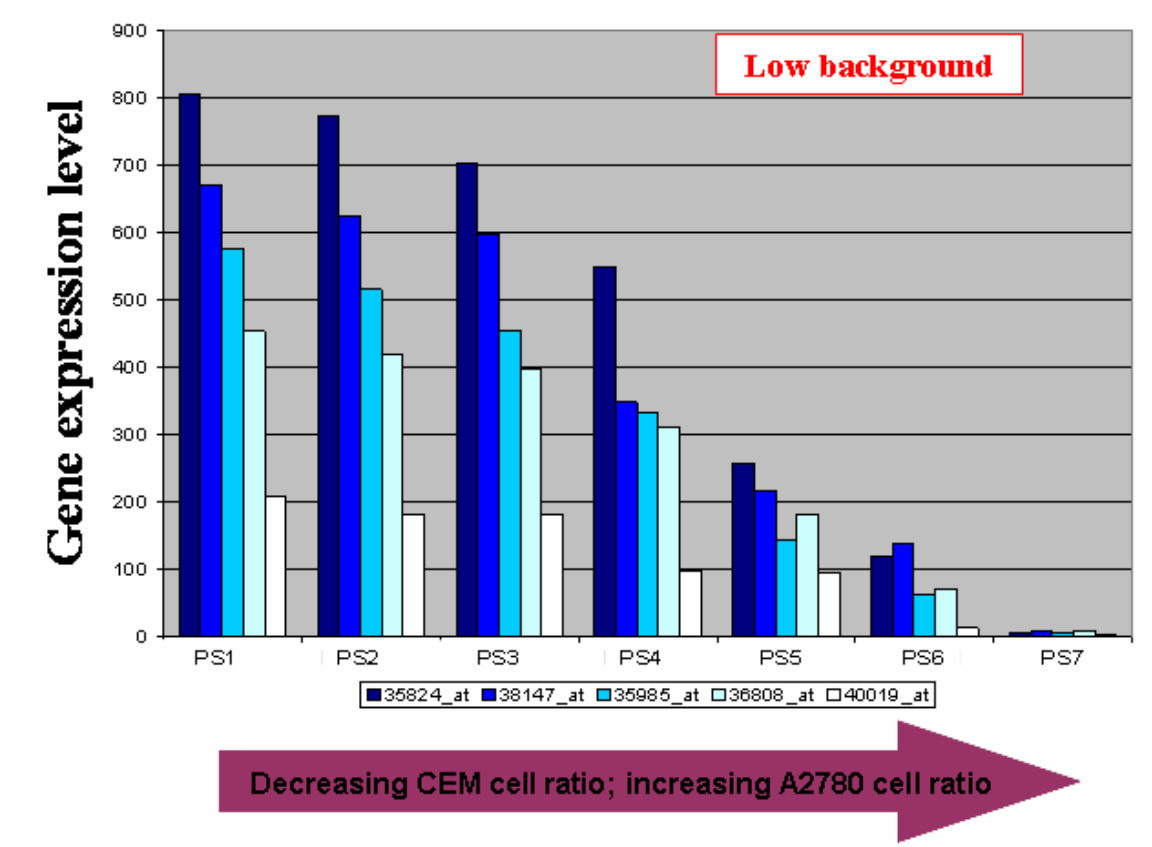

Figure 3.9. Ratio effect of cell mixtures on individual gene expression levels – Moderately expressed genes in the presence of low background

Pure and mixed CEM and A2780 cell samples were processed and analyzed separately using Affymetrix GeneChip® Human Genome U95Av2 array. Five individual genes with moderate expression levels (intensity range: 100-1,000) in CEM cells and very low expression levels (intensity range: 0-20) in A2780 "background" cells were plotted throughout 7 samples with decreasing CEM cell ratios.
Samples: PS1=100% CEM; PS2=90% CEM+10% A2780; PS3=75% CEM+25% A2780; PS4=50% CEM+50% A2780; PS5=25% CEM+75% A2780; PS6=10% CEM+90% A2780; PS7=100% A2780

Effects of Sample Handling

Our results described above confirmed the notion that, in order to investigate the GEP of a cell subset in a mixture, these cells need to be purified first. The next question in our study was, how much distortion the purification process itself (including cell fixation, labeling and sorting) would introduce into the studied GEP. We tested methanol-fixation, since alcohols are known to preserve nucleic acids better than cross-linking agents(60). Figure 3.10. displays the GEP of live, unlabeled CEM cells plotted against the GEP of antibody-labeled, methanol-post-fixed CEM cells. With the R^2-value of 0.982 we considered the overall GEP of the fixed and labeled sample unaltered.

Figure 3.11. presents a Trellis plot of six CEM cell samples all handled somewhat differently prior to microarray analysis. Two steps in the RNA isolation protocol, DNase digestion and purification of the isolated RNA by ethanol-precipitation are generally thought to be important to generate good quality RNA. We omitted these steps in handling two samples to test their importance. The effects of methanol fixation and antibody labeling - separately and combined - were also tested. None of the scatterplots presented here had an R^2-value lower than 0.975 and were practically identical. Again, a very small number of outlier genes displayed significantly different expression levels within the otherwise almost identical samples.

While the GEP might not be altered significantly by methanol fixation, we were concerned that this fixation might alter the cell subset separation process. However, as demonstrated by the flow cytometric results of Figure 3.12., the fluorescent intensity of CEM cells labeled with anti-CD4, R-phycoerythrin-conjugated (PE) antibody after methanol fixation was found to be comparable to that of live, labeled cells. Similar results were obtained for fluorescein isothiocyanate (FITC)-conjugated antibodies. Autofluorescence of methanol fixed cells was actually somewhat lower than that of live cells allowing improved separation of labeled and unlabeled cells after fixation.

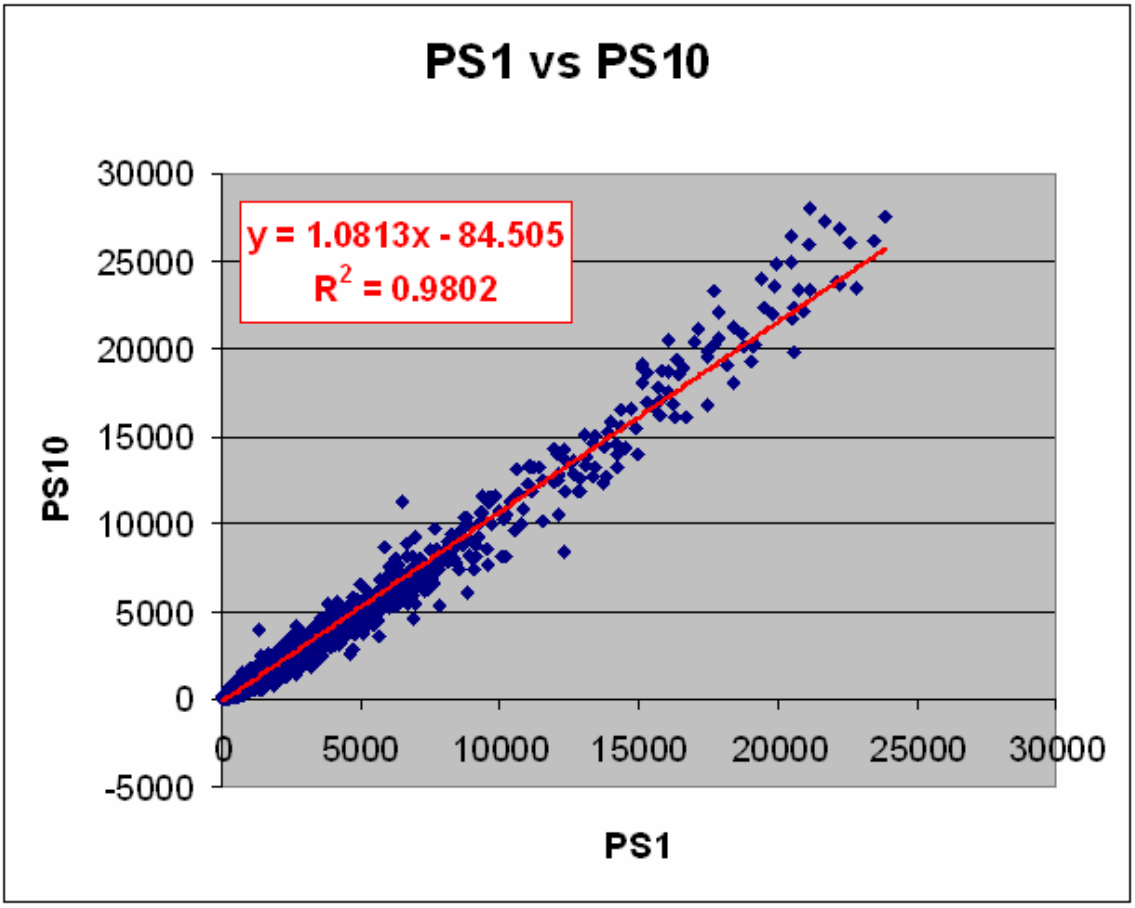

Figure 3.10. Effects of cell fixation and labeling on overall gene expression profiles

Live, unlabeled CEM cells and antibody-labeled, methanol-post-fixed CEM cells were processed and analyzed separately. The expression levels of over 12,000 genes were plotted. High R^2 value corresponds to high correlation between gene expression levels of the two cell samples. The overall GEP of the fixed, labeled sample remained unaltered Samples: PS1= live, unlabeled CEM cells; PS10= antibody-labeled, methanol-post-fixed CEM cells

Figure 3.11. Effects of alternative protocols on overall gene expression profiles

Live, unlabeled CEM cells and antibody-labeled, methanol-post-fixed CEM cells were processed and analyzed separately. The expression levels of over 12,000 genes were plotted pair wise for all 6 samples. Correlation between cell ratios and overall gene expression profile was visualized using a Trellis-plot chart. Each scatterplot is positioned in the row and column of the two samples it is made of.
The samples were all CEM cells: **PS1**= live, unlabeled; **PS8**= methanol fixed, unlabeled; **PS9**= live, antibody-labeled; PS10= antibody-labeled, methanol-post-fixed; **PS11**= live, unlabeled cells with no DNA precipitation in the protocol; **PS12**= live, unlabeled cells with no DNA precipitation and DNase digestion in the protocol.

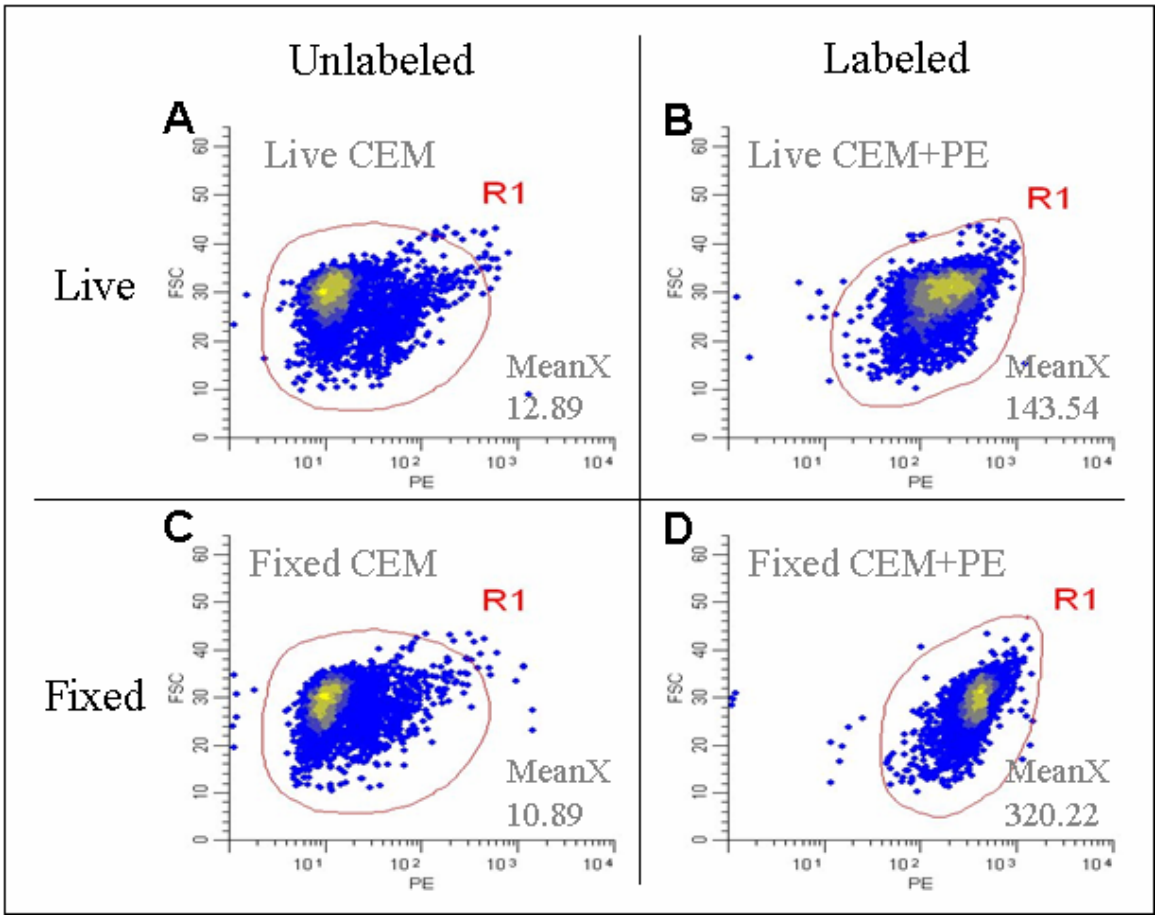

Figure 3.12. Effects of labeling and fixation on cell morphology – Flow cytometry scatterplots

Flow cytometric results of the fluorescent intensity of CEM cells labeled with anti-CD4-PE antibody after methanol fixation (D) are compared to that of live, labeled cells (B) as well as to live, unlabeled cells (A) and methanol fixed, unlabeled cells (C).

Recovering Gene Expression Profiles by Cell Sorting

Next, we studied if it was possible to recover the pure GEP of a cell subset after sorting these cells out of a mixture, and how close the recovered profile would be to the original especially, when the target cells were only a minor cell population in a mixture. Figure 3.13. displays the flow cytometry scattergrams of pure CEM cells, pure A2780 cells, their 1:9 mixture and the CEM cells sorted from this mixture after anti-CD4-PE labeling. We studied the effects of both Miltenyi magnetic bead based cell sorting and flow cytometry-cell sorting on the recovered GEP.

All four samples were analyzed by both spotted microarray technology (Figure 3.14.) and short-oligo arrays. From both types of array images we concluded that the GEP of the cell mix containing 10% CEM cells was very similar to that of A2780 cells, and the characteristics of CEM cells seemed to be lost in it. The CEM cells sorted from this 10% mixture exhibited the lost characteristics again very much resembling the original, pure CEM sample.

We compared the GEP scatterplots of these four samples after both spotted array (Figure 3.15.) and short-oligo array (Figure 3.16.) analysis. The generated scatterplots and their R^2-values were found to be very similar between the two methods. As demonstrated earlier, the GEPs of CEM and A2780 cells were indeed different with R^2-values of 0.808 (Clontech array) and 0.815 (Affymetrix array). The CEM profile was lost in the 10% mix with R^2-values of 0.802 (Clontech array) and 0.841 (Affymetrix array). After sorting the 10% CEM cells from the mixture the- recovered GEP of the sorted cells was found to be virtually identical to the original, pure CEM cells' profile with R^2-values of 0.990 (Clontech array) and 0.984 (Affymetrix array). For the Clontech array analysis CEM cells were purified by magnetic bead sorting (Miltenyi Biotec) and for the Affymetrix array analysis a combination of magnetic bead sorting and flow cytometry-cell sorting was used.

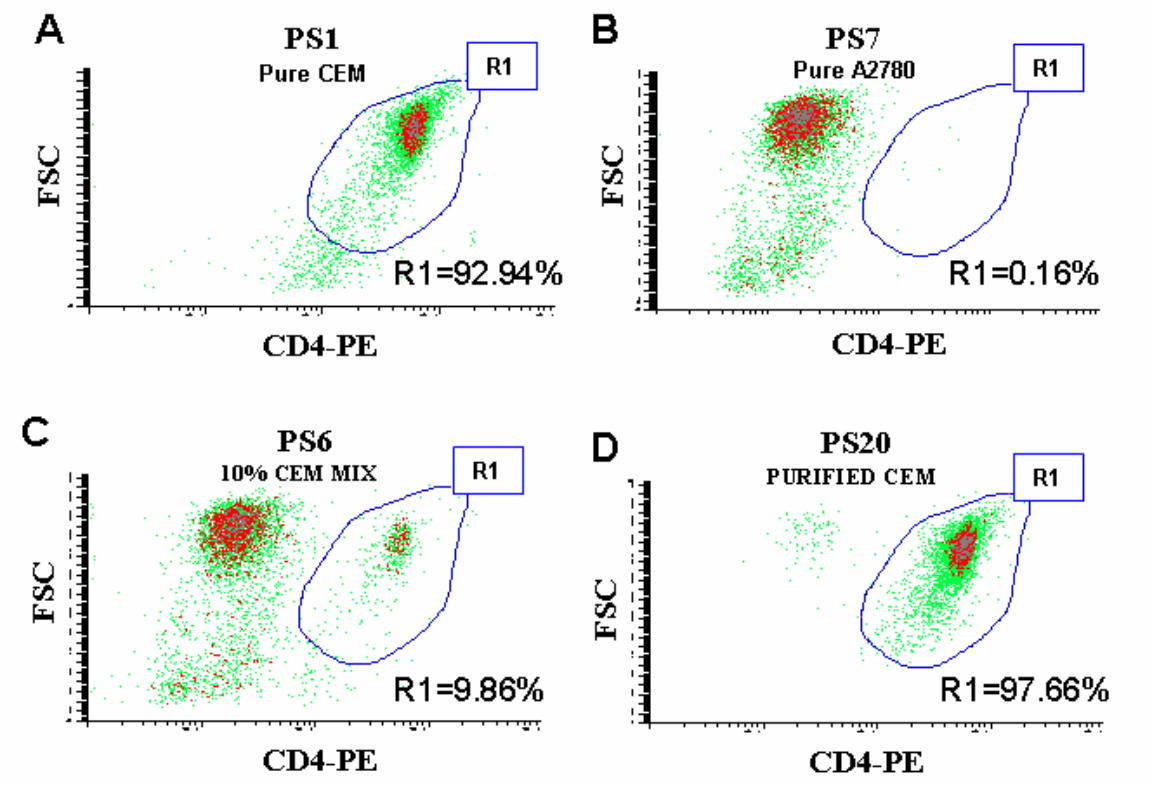

Figure 3.13. Sorting a minor cell population from a cell mixture – Flow cytometry scatterplots

Flow cytometric results of pure CEM cells labeled with anti-CD4-PE antibody (A), unlabeled pure A2780 cells (B), a 1:9 mixture of the two cell types (C), and immuno-magnetically purified CEM cells from the 1:9 CEM/A2780 mixture (D). The purified CEM cell sample is almost 98% pure.

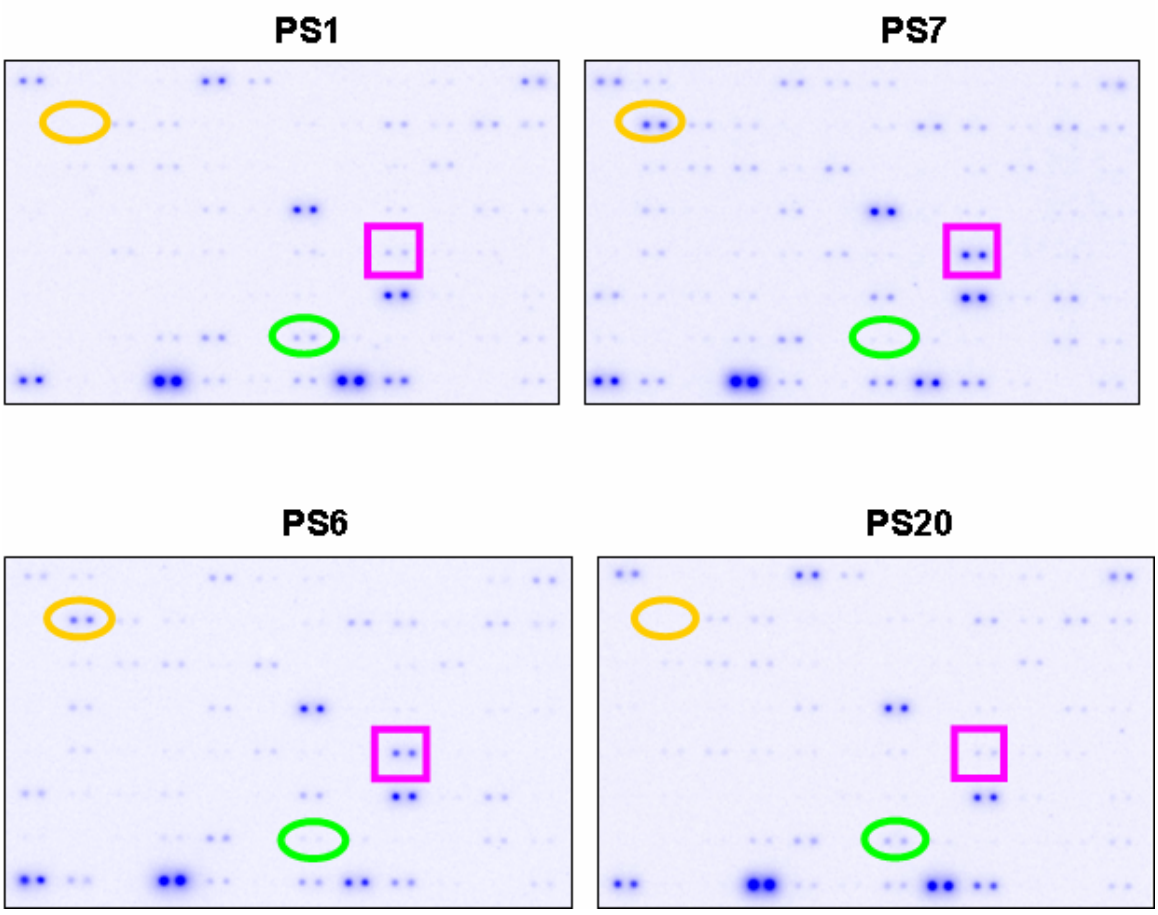

Figure 3.14. Sorting a minor cell population from a cell mixture – Microarray images

Clontech Atlas Trial microarray images of pure CEM cells labeled with anti-CD4-PE antibody (PS1), unlabeled pure A2780 cells (PS7), a 1:9 mixture of the two cell types (PS6), and immuno-magnetically purified CEM cells from the 1:9 CEM/A2780 mixture (PS20). Each pair of spots represents the expression level of an individual gene sequence. Examples for gene sequences expressed only in one of the cell types are marked with ovals 1, and 2. A differentially expressed gene sequence with much higher expression level in A2780 cells than in CEM cells is visible in square 3. The microarray images are presented to help visually verify the recovery of the gene expression profile for the re-purified CEM cells.

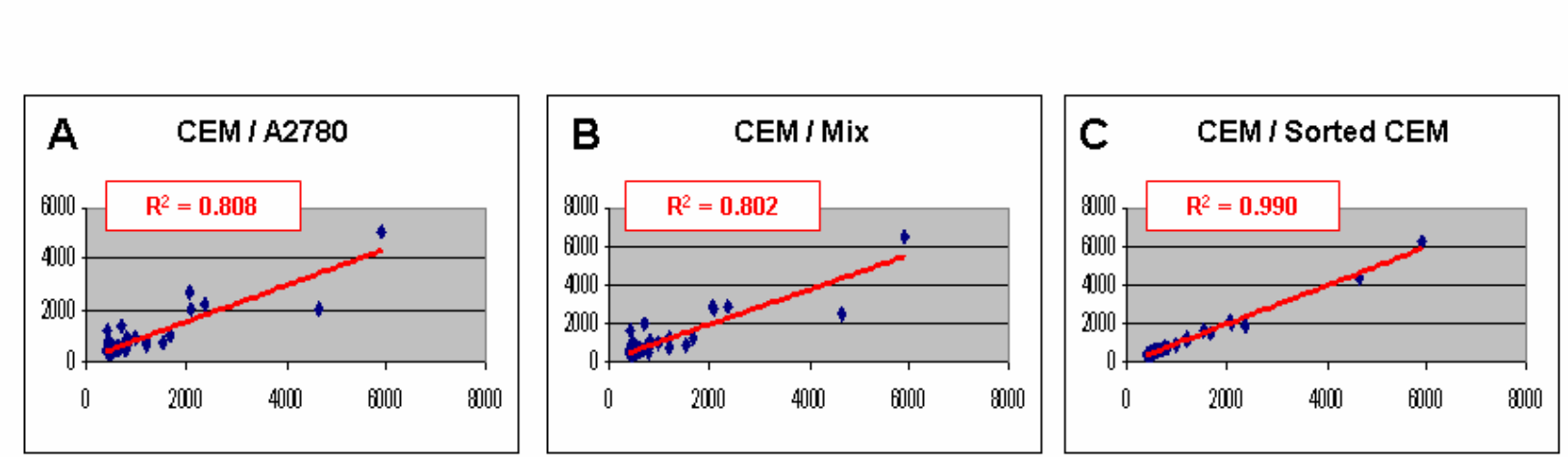

Figure 3.15. Sorting a minor cell population from a cell mixture – Scatterplots of Clontech microarray results

Gene expression levels of pure CEM cells are plotted against pure A2780 cells (A), against a 1:9 mixture of the two cell types (B), and against immuno-magnetically re-purified CEM cells from the 1:9 CEM/A2780 mixture (C). Expression levels of 96 different gene sequences are plotted for each sample. High R^2 values correspond to high correlation between gene expression levels of the original, pure CEM sample and the re-purified CEM sample.

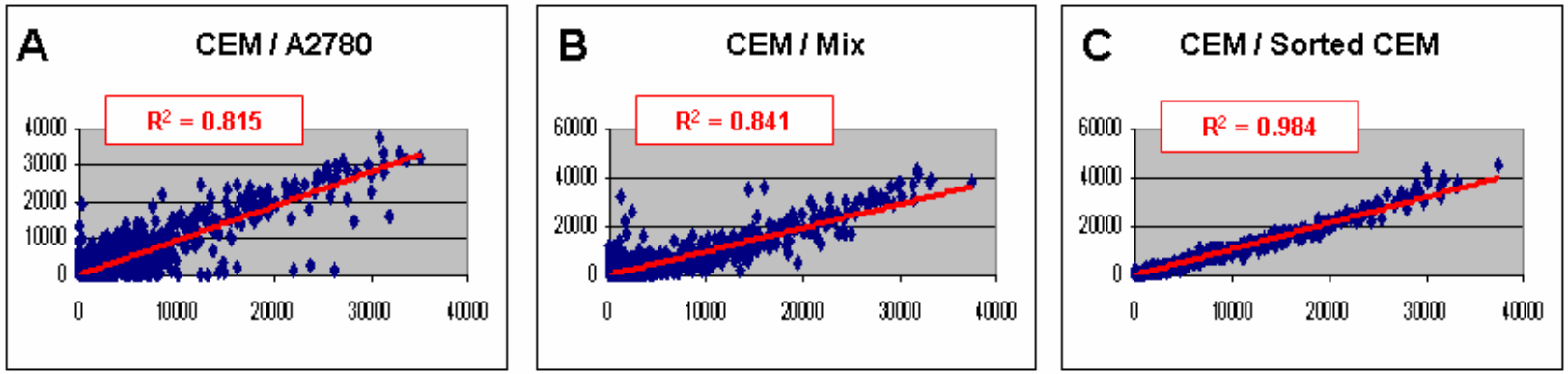

Figure 3.16. Sorting a minor cell population from a cell mixture – Scatterplots of Affymetrix microarray results

Gene expression levels of pure CEM cells are plotted against pure A2780 cells (A), against a 1:9 mixture of the two cell types (B), and against re-purified CEM cells from the 1:9 CEM/A2780 mixture (C). Expression levels of over 12,000 different gene sequences are plotted for each sample. High R^2 values correspond to high correlation between gene expression levels of the original, pure CEM sample and the re-purified CEM sample.

Profiling CD34+ Stem/Progenitor Cells

Purifying CD34+ Stem/Progenitor Cells from Cord Blood

After validating each step of cell purification for microarray analysis we tested the method on a real biological sample, which was one of our original motivating applications for this technology. CD34+ stem/progenitor cells were isolated from pooled cord blood using anti-CD34 antibody coated Miltenyi magnetic beads. KG-1a cells are a 100% CD34+ cell line usually used for modeling stem/progenitor cells. In this experiment they served as a control for monitoring possible GEP distortions caused by cell purification. Figure 3.17. displays the flow cytometric scattergrams of the four samples after anti-CD4-PE labeling. Sample PS301 was the mononuclear cells isolated from cord blood by Ficoll-Paque method. This sample obviously contained a mixture of different mononuclear blood cell types. Sample PS302 contained CD34+ stem/progenitor cells purified from cord blood by Ficoll-Paque method followed by magnetic bead sorting. This sample represents a purified cell subset from PS301. PS303 contained unpurified KG-1a cells and PS304 was purified from PS303 using the same method that was used to sort PS302. This last sample (PS304) served as an internal control for the effects of cell handling, since all KG-1a cells were CD34+. Putting them through the labeling and sorting process would have revealed any possible distortion the GEP of a labeled and magnetic bead purified cell subset might have suffered.

Flow cytometry scatterplots of samples PS303 and PS304 were indeed very similar confirming that all KG-1a cells were CD34+, and they all got sorted. CD34+ stem/progenitor cells constitute less than 1% of all cord blood mononuclear cells (114,115). As expected, in our experiments CD34+ cells from human cord blood samples could be detected only as a very small cell subset on the PS301 flow cytometry scattergram. Flow cytometry data of PS302 confirmed their successful isolation. Fluorescent microscopy also confirmed that the cells in this sample were indeed 100% PE-positive.

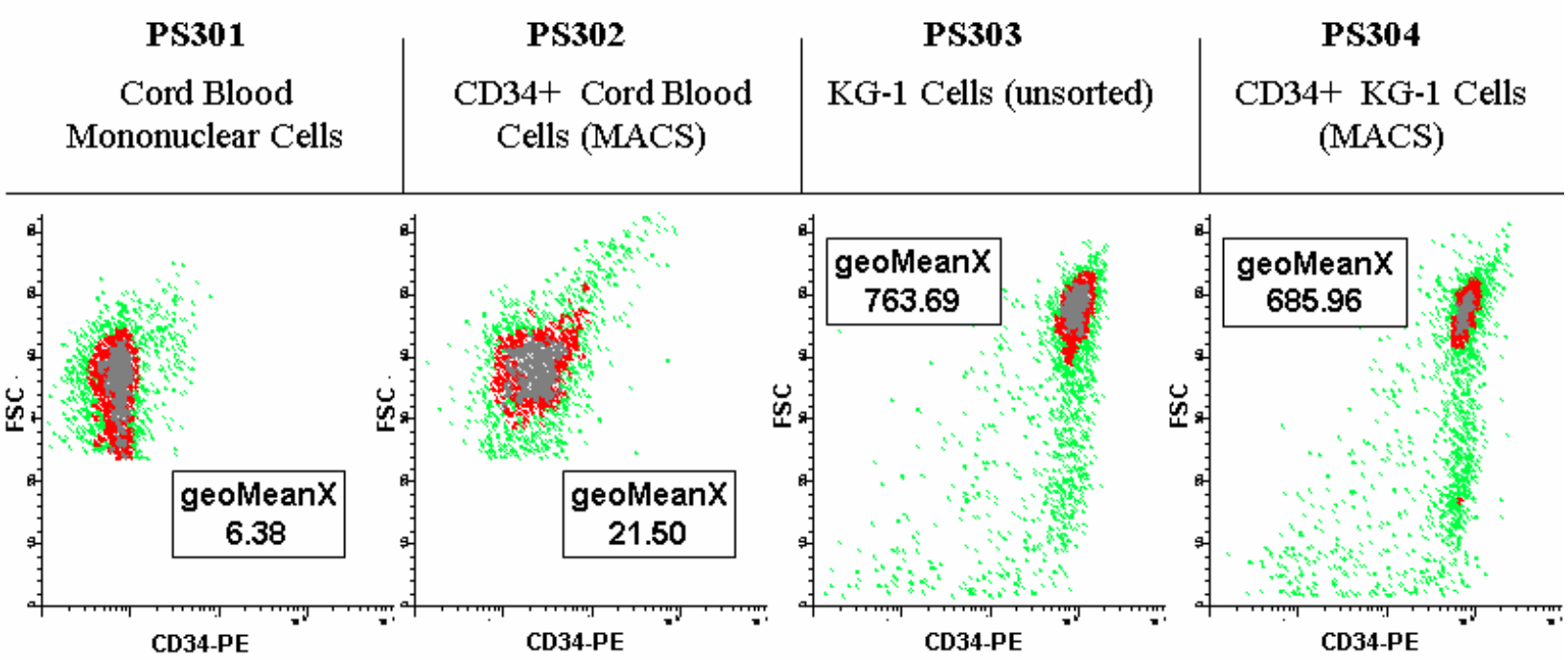

Figure 3.17. Purifying CD34+ Stem/Progenitor Cells from Cord Blood - Flow cytometry scatterplots

Flow cytometric results of mononuclear cells isolated from cord blood (PS301), immuno-magnetically purified CD34+ cord blood cells (PS302), unsorted KG-1 cells (PS303), and immuno-magnetically purified KG-1 cells (PS304), all labeled with anti-CD34-PE antibody.

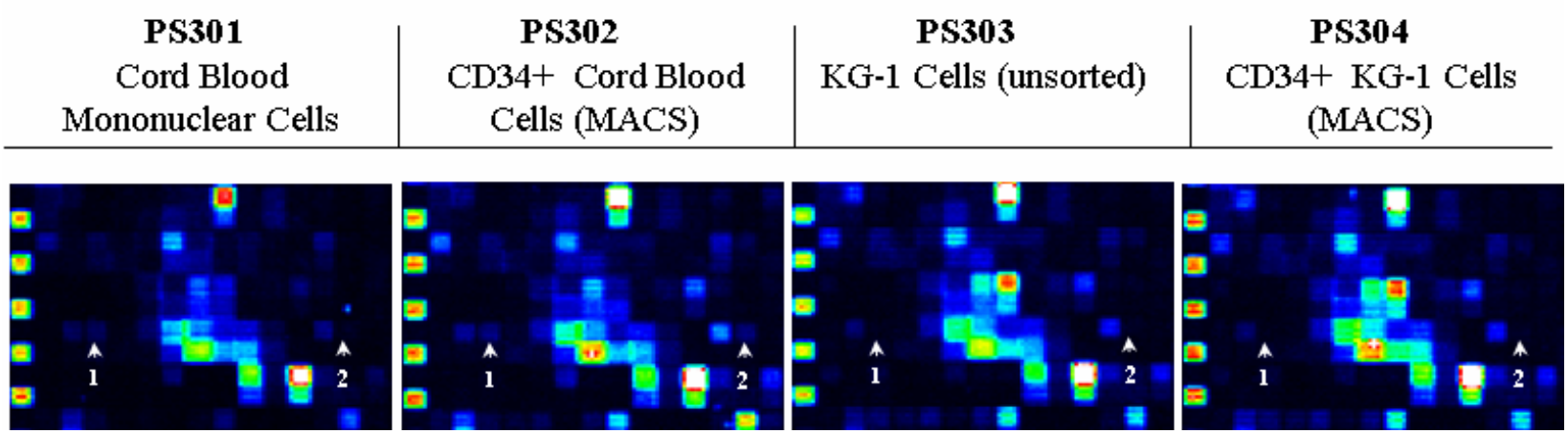

Figure 3.18. Purifying CD34+ Stem/Progenitor Cells from Cord Blood – Affymetrix microarray images

Corresponding segments of Affymetrix images for mononuclear cells isolated from cord blood (PS301), immuno-magnetically purified CD34+ cord blood cells (PS302), unsorted KG-1 cells (PS303), and immuno-magnetically purified KG-1 cells (PS304) are visualized by pseudo-coloring. Each small square represents the expression level of an individual gene sequence. Arrow 1 points at a gene that is expressed in purified CD34+ cord blood cells and CBMCs, but not expressed in KG-1 cells. Arrow 2 points at a gene that is only expressed in purified CD34+ cord blood cells.

Microarray Image Analysis

All four samples were analyzed by Affymetrix microarrays. Figure 3.18. displays the array images of the samples. The same portions of the full images were cropped and magnified and pseudo-colored as in earlier experiments. The images of sorted and unsorted KG-1a cells (PS303 and PS304) were virtually identical suggesting that the sort process did not introduce much distortion into the GEPs (at least into the genes present in this image segment). Comparing samples PS301 and PS302 we found that some genes expressed by purified CD34+ cord blood cells were similarly expressed in cord blood mononuclear cells (CBMCs) and not expressed in KG-1a cells (arrow 1). Expression levels of other genes were different in CD34+ purified cord blood cells than in any of the other three samples (arrow 2). These genes might be characteristic to stem/progenitor blood cells only.

Scatterplot Analysis

The scatterplot presented in Figure 3.19.C. also confirm that the sort process did not introduce any distortion into the overall GEPs, since the sorted and unsorted KG-1a cells (PS303 vs. PS304) were truly identical (R^2=0.99). The overall GEP of purified CD34+ cord blood stem/progenitor cells (PS302) was significantly different from the GEP of unsorted CBMCs (PS301) with an R^2-value of 0.83 when plotted against each other (Figure 3.19.A.). Surprisingly, the CD34+ stem/progenitor cells isolated from human cord blood were also very different from KG-1a cells (R^2=0.81), despite the fact that these cells are generally used as a model stem/progenitor cell line in experiments (Figure 3.19.B.). While KG-1a cells - originally derived from a bone marrow tumor - also express CD34 protein, the rest of their gene expression profile is much more different than expected from the GEP of normal CD34+ stem-progenitor cells isolated from normal human cord blood.

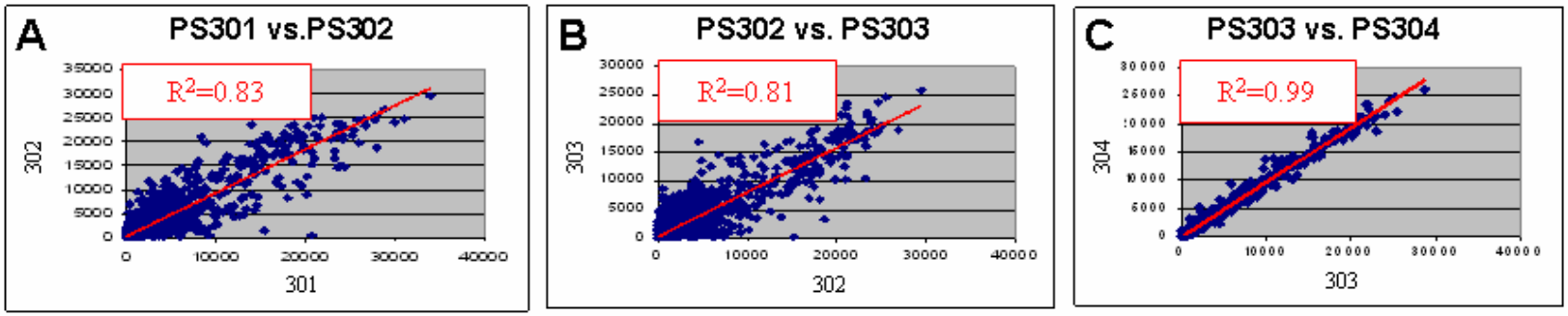

Figure 3.19. Purifying CD34+ Stem/Progenitor Cells from Cord Blood – Affymetrix microarray scatterplots

Gene expression levels of immuno-magnetically purified CD34+ cord blood cells (PS302) are plotted against mononuclear cells isolated from cord blood (PS301) on panel A, and against unsorted KG-1 cells (PS303) on panel B. To visualize the effects of immuno-magnetic labeling and cell sorting, unsorted KG-1 cells (PS303) and immuno-magnetically purified KG-1 cells (PS304) are plotted against each other on panel C. Expression levels of over 12,000 different gene sequences are plotted for each sample. R^2 values correspond to degree of correlation between gene expression levels.

Hierarchical Clustering

Figure 3.20. is a software (Spotfire, Inc. version 7.2) generated heat map comparing the expression levels of all of the genes judged to be "valid" (approximately 4800 genes) for the four samples. Samples and genes are ordered by Spotfire hierarchical clustering analysis based on normalized expression levels. Each sample is represented by a column of red and green stripes. Each gene is represented by a colored stripe where light green represents very low, light red represents very high expression levels. Each gene is placed on the same horizontal line in all four samples. The genes are arranged into clusters (both by gene expression levels in the vertical axis) and by cell sample (in the horizontal axis) by hierarchical clustering to visualize characteristic groups of genes as patterns.

Purified and unpurified KG-1a cells (PS304 and PS303) are represented by two almost identical columns underlining our previous conclusion that our purification protocol did not distort the purified cells' overall GEP. CD34+ stem/progenitor cells (PS302) truly displayed a different overall pattern from both KG-1a cells and from CBMCs (PS301), although some groups of genes were expressed similarly. Gene cluster 'a' is an example of genes that were expressed similarly in CD34+ stem/progenitor cells and CBMCs but not in KG-1a cells. Cluster 'b' was expressed similarly in CD34+ stem/progenitor cells and KG-1a cells but not in CBMCs. Genes in cluster 'c' were expressed differentially in CD34+ stem/progenitor cells from both other cell types. The expression levels of these genes were characteristic to CD34+ stem/progenitor cells only. This type of hierarchical clustering analysis found that the overall GEP of CD34+ stem/progenitor cells isolated from cord blood was still slightly closer to the GEP of the KG-1a cells than to the overall GEP of the mature blood cells in their original cord blood mixture. Prior purification of these CD34+ cells from cord blood was necessary to uncover their characteristic profile. Without cell purification it would not be possible to study changes in the GEP of these rare cells.

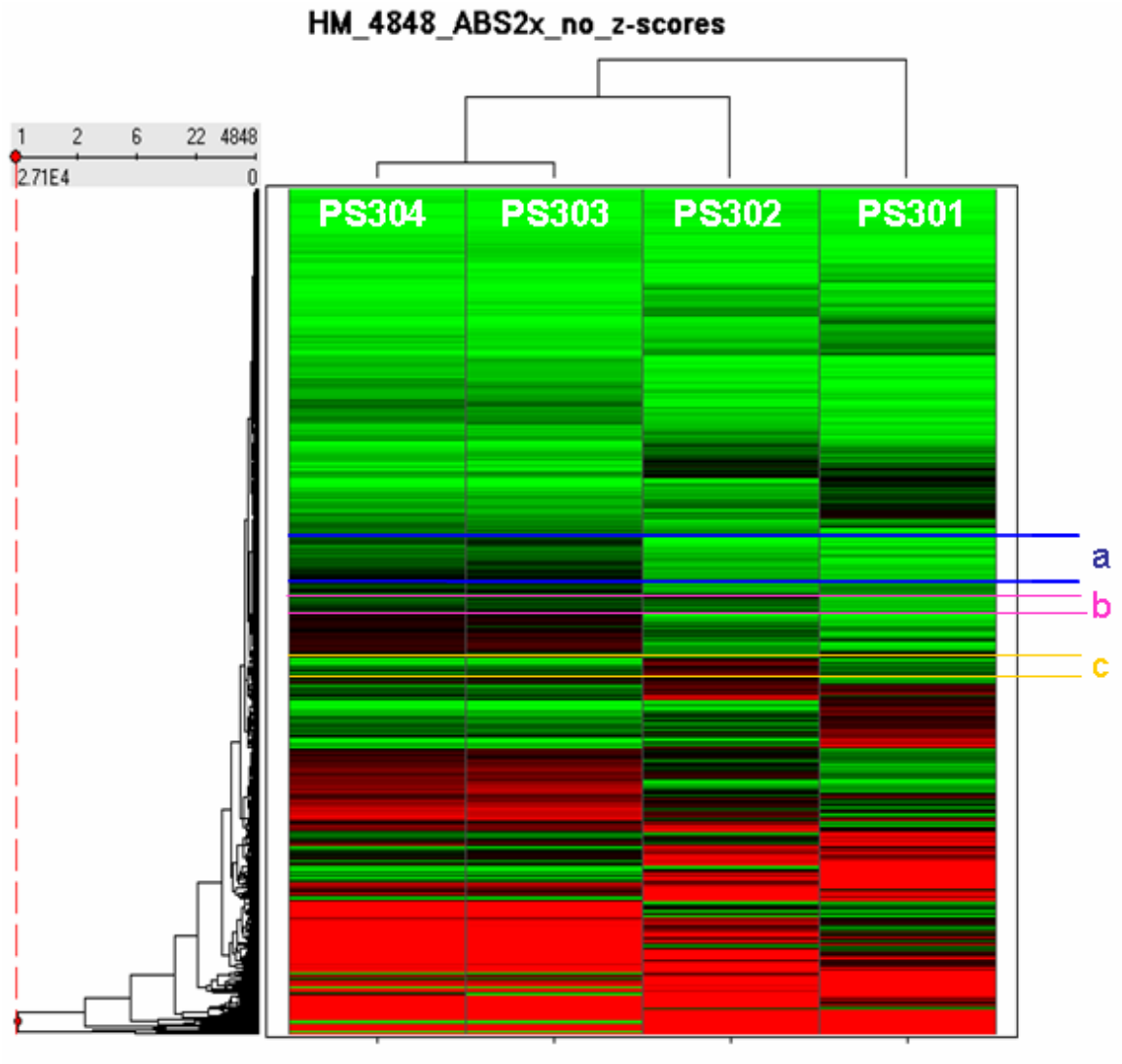

Figure 3.20. Purifying CD34+ Stem/Progenitor Cells from Cord Blood - Hierarchical clustering analysis

Computer software generated heat map comparing the gene expression levels in CBMCs (PS301), purified CD34+ cord blood cells (PS302), unpurified, and purified KG-1a cells (PS303 and PS304). Each gene is represented by a colored stripe where light green represents very low, light red represents very high expression levels. (a: A group of genes that are expressed in PS302 similarly to PS301; b: Genes that are expressed in PS302 similarly to PS303 and PS304; c: Genes that are uniquely expressed in PS302)

DISCUSSION

We have examined both the capabilities and limitations of the microarray approach in studying gene expression in defined, model cell mixtures and real biological samples. To determine if meaningful data could be obtained by this method even in the unfortunate scenario when the investigated cell type was a small minority in the sample, we tested the effects of cell–subset ratios and sample processing methods on the overall GEP as well as on individual gene expression levels. We modeled biological samples by cell mixtures of two cell types with different ratios and analyzed their GEPs by both spotted and short-oligo microarrays.

First, we confirmed that both platforms – when the corresponding protocols were carried out accurately – generated highly reproducible results with a 3-to 4 log dynamic range as reported in recent literature (24,37,39).

We determined that without applying any cell separation the cell type in majority dominated the overall GEP while the GEPs of minor cell subsets got washed out. Examining the overall GEP when investigating a minor cell subset of a cell mixture is like 'only seeing the tip of the iceberg'. The differences between the GEPs of the gradually changing cell mixtures convincingly mirrored the changes in cell ratios. In summary of our model cell mixture experiments, we concluded that the gene expression profiles for mixed cell populations are, as expected, the combined expression profiles for each cell subpopulation weighted according to its relative frequency in the cell mixture.

We noted that while "housekeeping genes" are often used to standardize samples in Northern blots and other technologies analyzing gene expression, we found these genes differentially expressed between cell samples. Therefore, it seems more prudent to use overall gene expression levels or at least a set of "housekeeping genes" rather than any single one of these genes as standard in these experiments.

We determined the sensitivity of the microarray approach in analyzing cell mixtures. In our model, the overall GEP of a more than 75% pure sample was virtually indistinguishable from a 100% pure sample. In fact, any two samples that had a

difference of less than 25% in their mixture ratios could be considered identical based on their overall GEP, since when compared, they produced very high (greater than 0.975) R^2-values. Based on these results we established that the sensitivity threshold for overall GEP comparisons for these techniques in our hand was 25% cell mixture ratio difference.

The number of outlier genes within these samples as seen in the Trellis plots was very small (fewer than 10 out of more than 12000 genes) considering that the raw data was not pre-processed prior to scatterplot analysis. However, these outlier genes might indicate that purity requirements can be very different for monitoring individual genes, depending on whether or not those same genes are expressed at high or low levels in the contaminating cell types.Our results indicated that, in the case of minor cell subsets, to be able to see 'more than just the tip of the iceberg' cell purification is necessary. To address the question whether cell labeling, fixation, and sorting alters the cells' GEP, and the reasonable concern that the more a sample is processed the more distorted its GEP might get, we also tested handling effects on overall sample GEP. We demonstrated that after antibody labeling and methanol fixation the overall GEP remained unaltered and even omitting steps traditionally used to improve RNA quality did not have any significant effect on the overall GEP. Again, the presence of a few outlier genes indicated that individual genes might be much more affected by certain processing steps, e.g. antibody labeling of a surface receptor on a live cell obviously, might trigger certain pathways altering the expression levels of the genes involved. Nevertheless, we were able to conclude that the "overall GEP" of a sample (representing the vast majority of all genes) is more robust and resistant to sample processing than it has been generally appreciated.

Methanol/PBS fixation of the antibody labeled cells prior to cell purification and microarray analysis turned out to be a rather fortunate choice. It did not unfavorably alter the detection/selection process by either immuno-magnetic cell separation or flow cytometric cell sorting. Both PE- and FITC labeled antibodies used in these experiments maintained good separation characteristics after methanol post-fixation. In summary, the methanol/PBS fixation method not only preserved the GEP of labeled cells, but also allowed fluorescence-based labeling for cell sorting.

To address the question of how much purity we need in a sample, we concluded that generally it is not necessary to achieve 100% purity. In our model for the overall GEP a sample that was at least 75% pure was found to be indistinguishable from the pure sample. Starting with a 10-50% cell mix his level of purity can be achieved by two rounds of magnetic bead sorting. One round typically results in about 70% purity, two rounds raises the purity to approximately 90%, while after 3 rounds it is generally above 95%. One round of magnetic bead sorting followed by one round of flow cytometry/cell sorting results in about 90-98% purity as well. Studying individual genes however, might require much higher or lower sample purity depending on the gene's relative expression levels in the cell subsets. With good cell biomarkers and technique, multiparameter flow cytometry/ cell sorting can be used to obtain purities of more than 99 percent. This degree of purity may be needed for correct GEP analysis of low-expressing genes where even 90 percent purity may be insufficient to obtain accurate GEP results for those specific genes.

To test just how much of the 'iceberg' could be revealed, we purified the minor cell subset of a 10% cell mix where the GEP of these cells was almost completely covered by the background cells. Using both magnetic bead cell purification and flow cytometry cell sorting we managed to recover the 'hidden' GEP virtually perfectly, also proving that the sort process itself did not distort the profile. The almost perfectly recovered profiles suggest that both magnetic bead and conventional flow cytometry/cell sorting purification methods are capable of purifying cells without significantly distorting their GEP. Results from the control KG-1a cells from the cord blood experiment confirmed this finding since the cells that went through the sort process matched the unsorted cells with an R^2-value of 0.99. We concluded that for meaningful gene expression microarray profiling a minor cell subset of a cell mixture, purification of these cells is not only necessary but also very much achievable, recovering the 'pure profile' without any significant distortion despite the concerns expressed previously in the literature (31,59,60). In our hands, following the procedures described in this study, the effects of sample handling on the GEP were minimal and not significant.

As a proof-of-principle experiment, we measured the GEP of purified, CD34+ cord blood stem/progenitor cells. Since these cells are present in cord blood in less than 1% minority of all mononuclear cells (114,115), their GEP had been heavily masked by the overwhelming presence of mature contaminating cell types. The true stem/progenitor cell-GEP was 'invisible' without purification. We demonstrated that the recovered GEP of these cells was characteristically different from both CBMCs and KG-1a cells. This result seriously questions the use of KG-1a cells as a model cell line for stem/progenitor cells in gene expression studies, even though that cell line was originally established from a bone marrow tumor.

For the cord blood experiment we needed to pool several samples. Individual cord blood samples could not be directly analyzed simply because they did not provide enough purified CD34+ cells necessary for one microarray analysis. This problem would be even more serious if we wanted to further purify this cell subset based on the cells' other surface antigen properties. Unfortunately, many biological samples do not provide enough purified cells of a certain cell type for direct gene expression profiling. For these samples, RNA-amplification is necessary prior to microarray analysis.

In summary, we found that both Clontech and Affymetrix arrays performed at a very high level of reproducibility, the generated profiles followed the cell mixing ratios very accurately, these profiles proved to be surprisingly robust, and hidden GEPs could be accurately recovered from cell mixtures by cell separation techniques. Based on these results, we concluded that MAT combined with sample purification would allow us to study alterations in the GEP of our cells of interest.

CHAPTER 4.

CELL PURIFICATION FOR MICROARRAY ANALYSIS – MULTISTAGE MAGNETIC SORTING

INTRODUCTION

The objective of the experiments described in this chapter is to test whether using multistage magnetic sorting it is possible to purify mixed cell samples without significantly altering the cells' GEP even with small, biohazardous samples. The two cell separation methods we used successfully in the experiments described in Chapter 3. (flow cytometric cell sorting and column based immuno-magnetic cell separation) performed very well in our hands, but for some special applications, especially in purifying cell subsets out of a small, biohazardous cell sample, these methods have serious disadvantages. Recently, a new cell separation technology, called multistage magnetic sorting, has been developed by SHOT, Inc. (66,116,117), for the purpose of purifying homogeneous cell subsets from small cell mixtures and potentially biohazardous biological samples in extraterrestrial, and other unconventional biological research applications. We studied the applicability of this new technology, for sorting our target cell populations prior to microarray analysis.

Multistage magnetic sorting has the ability to sort particles and cells according to their degree of magnetization and to collect up to 15 fractions sorted (65-68). This novel cell sorting technology utilizes a Multistage Magnetic Sorter (Magsort™) instrument that is described in detail in Chapter 2. Briefly, the Magsort instrument (Figure 4.1.) contains a pair of horizontal circular plates housing the sample insertion chamber and the sort chambers that collect the separated fractions. An onboard computer controls the movement of these plates as well as the movement of a magnet changer via stepper motors. The magnet changer holds up to 6 earth magnets of different strength

Figure 4.1. The Multistage Magnetic Sorter

The Magsort beta prototype used in our laboratory. As an interesting feature it utilized a Supernintendo graphical user interface. The Magsort equipment is described in detail in Chapter 2.

each of which can be placed above the sample well for a period of time as determined by the user. The onboard computer is programmable by a graphical user interface (GUI). Separations are accomplished by exposing magnetically labeled cells (or magnetic microparticles) to a magnetic field that pulls them out of the sample insertion chamber into the sort chambers (65-68).

By multistage magnetic sorting one can divide a magnetically labeled cell mixture into individual fractions on the basis of magnetophoretic mobility. Cells are separated based on the amount of magnetic material attached to the cell surface, which is thought to be directly proportional to the surface antigen content of immuno-magnetically labeled cells (64). Receptor or cell-specific antibodies – attached to the surface of magnetic beads – can therefore be used not only to separate the cells of interest from unwanted cells, but to separate those selected cells into specific fractions by applying increasingly stronger magnetic fields to the immuno-magnetically labeled cells in a stepwise fashion (65-68).

One objective of the Magsort experiments in our laboratory was the separation of stem/progenitor blood cell subsets from cord blood for gene expression profile (GEP) microarray analysis. In model cell mixture experiments we used anti-CD34 antibodies for the immuno-magnetic cell labeling since CD34 antigen expression is one of the best described characteristics of the human stem/progenitor blood cell compartment (118-120). Human CD34+ cord blood cells are heterogeneous both in function and in CD34 antigen expression level where different levels of CD34 antigen expression correlate with different function and differentiation stage (56,119,120). The KG-1a cell line was established from bone marrow tumor and these cells all express CD34 antigen. (ATCC CCL-246.1) To test if multistage magnetic sorting was capable of separating different cell subsets based on different levels of CD34 antigen expression, we sorted immuno-magnetically labeled CD34+ KG-1a cell subsets out of model cell mixtures. We used concomitant anti-CD34 immuno-fluorescent labeling to investigate the efficacy of multistage magnetic sorting by analyzing the sorted cell fractions using fluorescent microscopy and flow cytometry.

Another objective of our microgenomics project was to study the changes in the GEPs of human immunodeficiency virus (HIV) infected cells and HIV-challenged human stem/progenitor blood cell subsets. The CD4 surface antigen is the main receptor used by HIV (121,122) and the virus is known to alter the CD4 antigen expression level of infected cells (121). We designed a set of experiments to separate cells based on the expression level of CD4 antigen on the sample cells' surface. This capability could be important in the study of HIV infection, especially in separating different fractions of live, HIV-infected cell samples. High sorting accuracy combined with the ability to separate biohazard samples in a closed system could be the key to meaningful microarray experiments in the fields of stem cell and HIV research.

With the unique capability of sorting cells into several fractions (as opposed to traditional "all-or nothing" immuno-magnetic cell sorting) in a closed system (as opposed to flow cytometric cell sorting) multistage magnetic sorting was a promising candidate for a highly accurate cell purification method capable of safely sorting cell subsets out of small, biohazardous samples. To develop a protocol that enhances the accuracy and sensitivity of this approach we tested the Magsort instrument using several different cell types, antibodies, magnetic bead types, loading methods and sort programs.

RESULTS

Magsort Separations of Commercial Magnetic Microparticles

In initial experiments performed at SHOT, Inc., well-characterized, commercially manufactured magnetic microparticles were used to test the capabilities of multistage magnetic sorting. Mixtures of two different microparticles mixed in approximately equal proportions were suspended in cell culture medium and loaded into the sample chamber. The samples were exposed to the full array of six magnets and the resulting fractions were examined on a hemacytometer to evaluate whether the two different particles separated into different fractions, how high purity could be achieved and whether further subsets within a certain type of microparticle could be sorted out from the mix. Figure

4.2. displays the results of an experiment in which two Estapor particle types with different diameters (2.39 μm and 0.7 μm), both containing magnetite, were sorted into 10 fractions. After the sort a particle count of each fraction was taken. The Estapor 0.7 μm particles were found to be 80% pure in fraction 6 (377 mT magnet), and the Estapor 2.39 μm particles were 98% pure in fraction 8 (518 mT magnet). A small portion of the Estapor 0.7 μm particles were found in the first fraction that was not exposed to any of the magnets. These particles might have been highly buoyant and never settled at the bottom of the sample chamber. The rest of the fractions either contained very few particles or an unsorted mix of them. No further subsets got sorted.

Figure 4.3. displays the results of another two-particle experiment where approximately 3.4 x 10^7 particles/ml of Polysciences particles with 1-2 μm diameter and approximately 4.4 x 10^7 particles/ml of Magsphere particles with 5.56 μm diameter were mixed and suspended in DMEM cell culture medium. Again, this mixed sample was exposed to the full array of six magnets, and the resulting 10 fractions were examined on a hemacytometer. The Magsphere (5.56 μm) particles were collected mostly in the leftover sample (which was 96% pure 5.56 μm particles) indicating that they were not magnetic enough to be attracted by any of the magnets, but some appeared in the 68 mT and 172 mT-magnet fractions (fractions #2 and #3) indicating that about 1/3 of the particles were strongly magnetic. The purity of these Magsphere particle subsets found in the early fractions was approximately 80%. The Polysciences (1-2 μm) particles were collected by the 445 mT and 495 mT magnets and were separated from the mixture to about 75-80% purity. These smaller particles were found to be more homogeneous than the Magsphere particles although a small subset of them was found in the first fraction at approximately 70% purity. These experiments demonstrated the potential of multistage magnetic sorting in separating 1-6 μm sized magnetic particles into purified or enriched fractions based on the particles' magnetophoretic mobility.

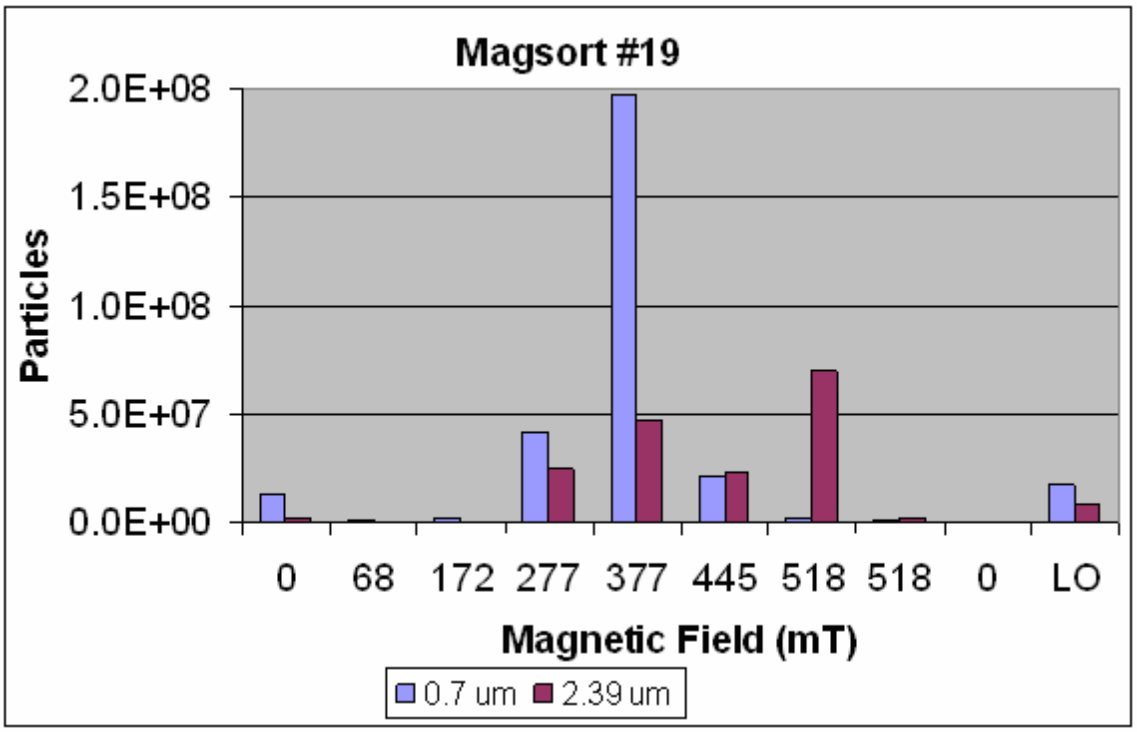

Figure 4.2. Magsort separation of Estapor (0.7 µm) and Estapor (2.39 µm) microparticles

Distributions of the two particle types separated from an approximately 50% mixture. Ordinate: particles per fraction. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction on the left and the ninth fraction were not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets.

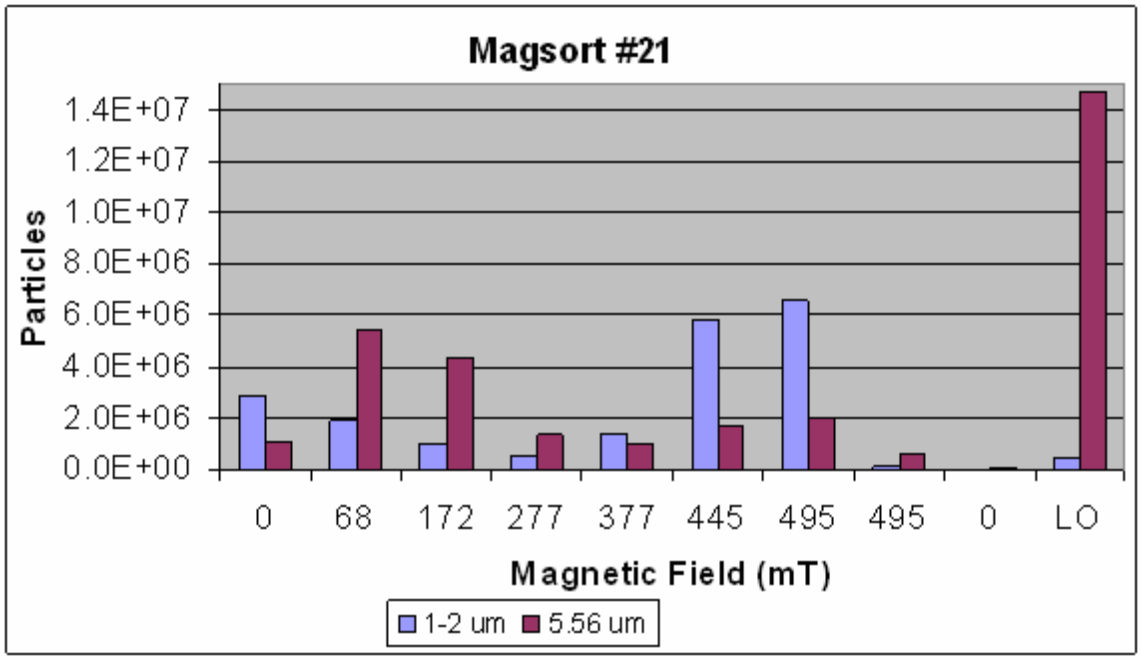

Figure 4.3. Magsort separation of Polysciences (1-2 µm) and Magsphere (5.56 µm) microparticles

Distributions of the two particle types separated from a 45% / 55% mixture. Ordinate: particles per fraction. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction on the left and the ninth fraction were not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets

Test Separation of Model Cell Mixtures Using the Original Protocol

To develop a protocol designed for multistage magnetic sorting of homogeneous cell subsets from mixed cell samples for subsequent microarray analysis we evaluated the feasibility of four magnetic bead types: two different Dynal beads with 4.5 μm and 1.0 μm diameter (Figure 4.4.A and B), Bangs beads with and average diameter of 0.86 μm (not shown), and Miltenyi beads with 50 nm diameter (Figure 4.4.C and D). Using greatly different-sized magnetic beads (Figure 4.5.) enabled us to determine the optimal range of bead size for this approach. As illustrated in Figures 4.4. and 4.5., while the other 3 bead types are spherical and fairly uniform Bangs beads have irregular shape and their size varies in the range of 0.3-1.2 μm. Miltenyi beads were coated with either anti-CD34 or anti-CD4 antibody; the other 3 bead types were all streptavidin coated.

Figure 4.6. presents confocal images of cells labeled with Dynal 4.5 μm beads (A), Dynal 1.0 μm beads (B), and Bangs 0.86 μm beads (C). Miltenyi beads are two small to be visible using conventional confocal microscopy. (not shown) As figure 4.6. demonstrates, the average number of labeling beads per cell depends mainly on the bead size and type. Since Dynal 4.5 μm beads are comparable to cells in size, usually 1-6 beads attached to a positive cell. It was common to find the same Dynal 4.5 μm beads attached to more than one cell as demonstrated in Figure 4.6.A. This often resulted in the forming of multiple-cell/bead aggregates. The range of labeling Dynal 1.0 μm beads per cell was found to be 1-12 in our experiments (Figure 4.6.B) while for Bangs beads it was approximately 5-20 (Figure 4.6.C). According to the manufacturer, the same ratio for Miltenyi beads can be up to several hundreds of beads per cell (62). The cells presented in Figure 4.6. were fluorescently labeled as well to enable us to evaluate the purity of the resulting sort fractions by fluorescence microscopy. Figure 4.6.A displays KG-1a cells labeled with R-phycoerythrin (PE)-conjugated anti-CD34 antibodies and Dynal 4.5 μm beads. Figure 4.6.B displays CEM cells labeled with Alexa-488-conjugated anti-CD4 antibodies and Dynal 1.0 μm beads. Figure 4.6.C displays KG-1a cells labeled with PE-conjugated anti-CD34 antibodies and Bangs 0.86 μm beads.

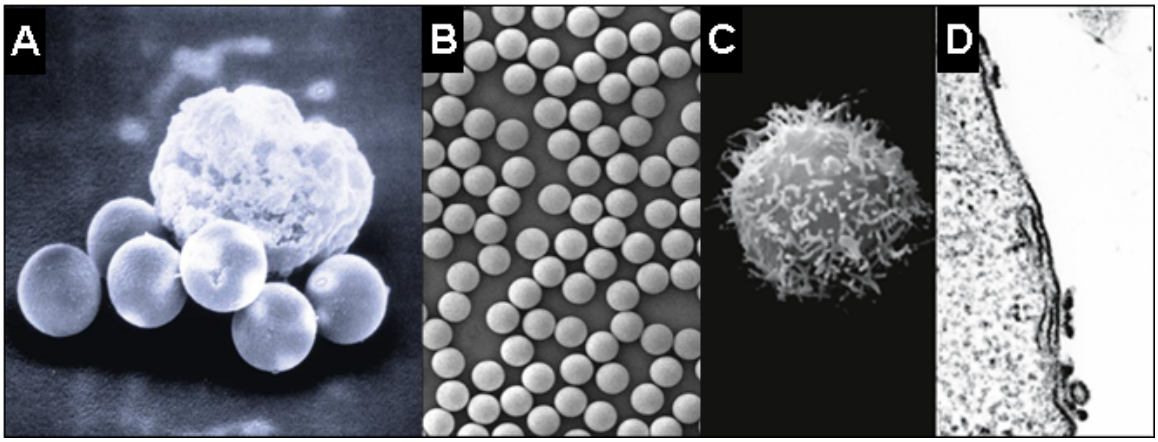

Figure 4.4. Commercially available immuno-magnetic microparticles

A: Scanning electromicrograph of a Dynabead (4.5 μm) labeled H9 cell (source: www.invitrogen.com-Prof. R.H. Dennin) B: Scanning electromicrograph of a MyOne Dynal beads (source: www.invitrogen.com-Mikal Heldal) C and D: Scanning and transmission electromicrograph of a Miltenyi bead (50 nm) labeled CD8+ T-cell (source: www.miltenyibiotec.com-Prof. Groscurth) All three types of beads are spherical and uniform.

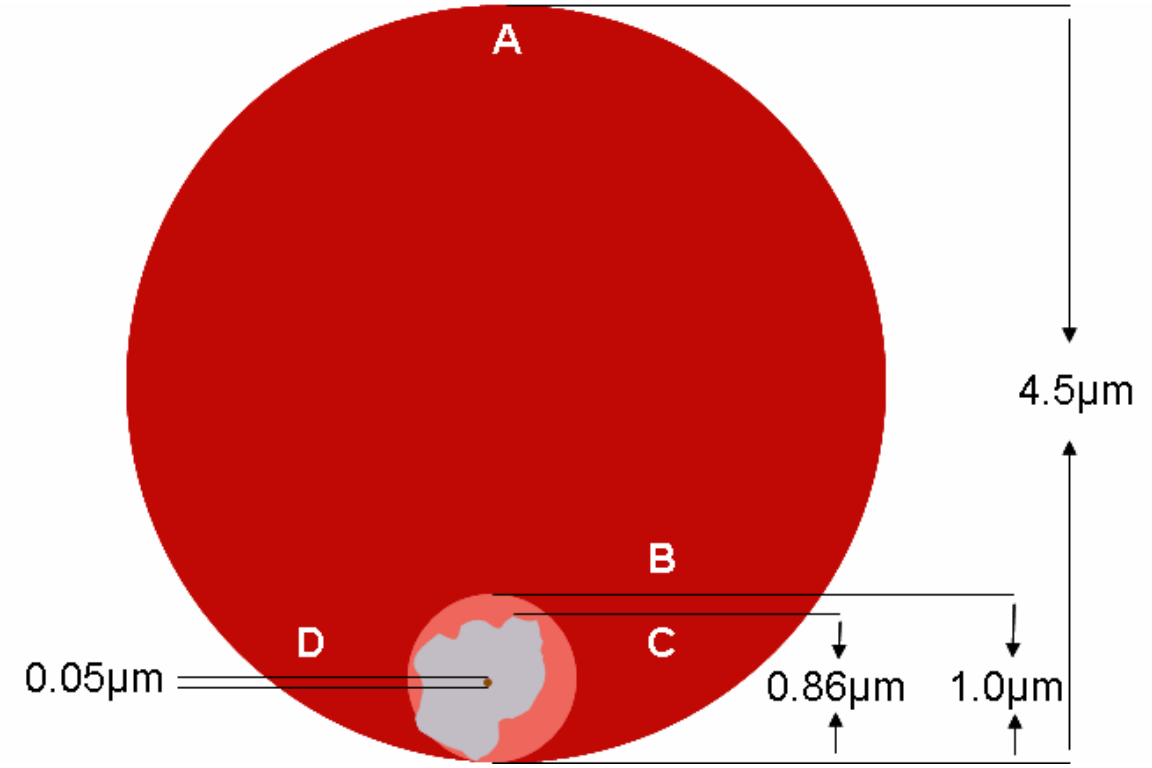

Figure 4.5. Scaled drawing of 4 types of immuno-magnetic microparticles

Dynal-4.5 µm (A), Dynal-1.0 µm (B) Bangs-0.86 µm (C), and Miltenyi-50 nm (D) microparticles were drawn to scale for comparison. The drawing illustrates that while the other three bead types are spherical Bangs beads are irregular-shaped.

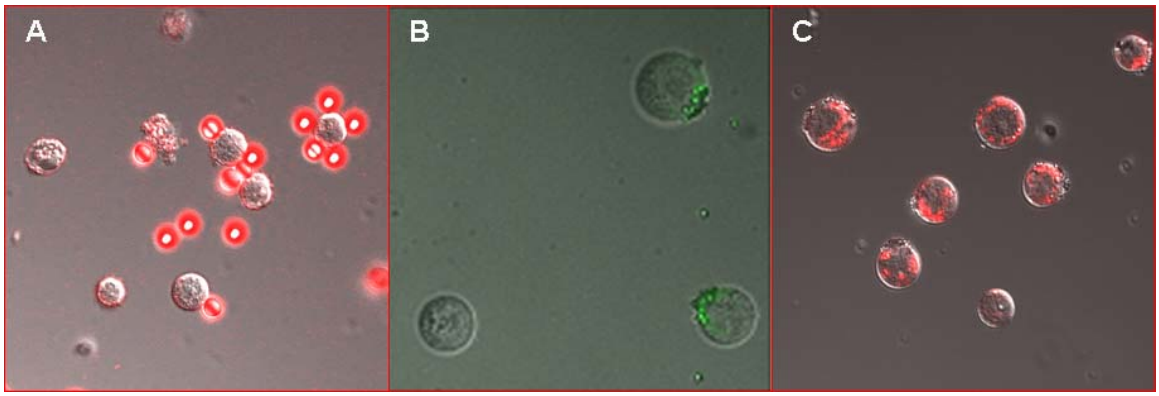

Figure 4.6. Immuno-magnetically and fluorescently labeled cells

A: KG-1a cells labeled with PE-conjugated anti-CD34 antibodies and Dynal 4.5 μm beads. The beads also display strong fluorescence at 560/590 nm. One of the cells in the view is labeled by five beads, while another cell is labeled by one bead only. Another two cells form an aggregate with several beads. B: CEM cells labeled with Alexa-488-conjugated anti-CD4 antibodies and Dynal 1.0 μm beads. The cells display capping with both the fluorescent antibody and 6-10 beads. C: KG-1a cells labeled with PE-conjugated anti-CD34 antibodies and Bangs 0.86 μm beads. Capping is less prominent with Bangs beads than with Dynal beads. Bangs beads display uneven shape, size, and distribution.

We tested all four types of immuno-magnetic beads separately for multistage magnetic sorting. We loaded samples of magnetically and fluorescently labeled cells mixed with unlabeled cells at approximately 50% mixture ratios into the Magsort instrument and analyzed the sorted fractions by cell counting and fluorescence microscopy. As demonstrated in figures 4.7.-4.10., none of the four magnetic bead types was able to selectively pull magnetically labeled cells from the bottom of the sample chamber into the sort chamber. The few cells that did end up in the sort chambers were not delivered there selectively based on their immuno-magnetic labeling, because the cell mixture ratio of these fractions was found unaltered compared to the initial sample. The vast majority of all recovered cells (98% with Miltenyi beads – Figure 4.7., 86% with Dynal 4.5 μm beads – Figure 4.8., 93% with Dynal 1.0 μm beads– Figure 4.9., and 92% with Bangs beads– Figure 4.10.) remained in the sample insertion chamber. Unbound magnetic particles did get sorted similarly to the microparticle mixture experiments described earlier. The cell mixture ratio in the leftover sample was found to be virtually identical to the initial cell mixture loaded into the system. This was invariably the case with all cell types (KG-1a, CEM, PBMC), and magnets (68 mT, 172 mT, 277 mT, 377 mT, 445 mT, and 495 mT) tested using any user definable sort program.

We tested the effects of extending or shortening the time allowed for initial cell settling and for the individual sort periods, but even when the settling time was omitted and the sort time was extended to several hours using the strongest magnet (495 mT) the recovered cell subsets were not specifically purified or enriched. Shortening or omitting the initial cell settling period did result in slightly increased numbers of cells in the first two sort fractions but the cell subset ratio in these fractions remained unchanged. The optimal settling time before applying the first magnet was found to be 30-60 minutes. Increasing or decreasing the number of cells in the initial cell mixture loaded into the system did not have any effect on the purity of the sort fractions but starting with too many or too few cells did result in increased cell loss. The optimal sample size was found to be 1-6x10^6 cells with 10-25% cell loss measured by the loaded/recovered cell ratio.

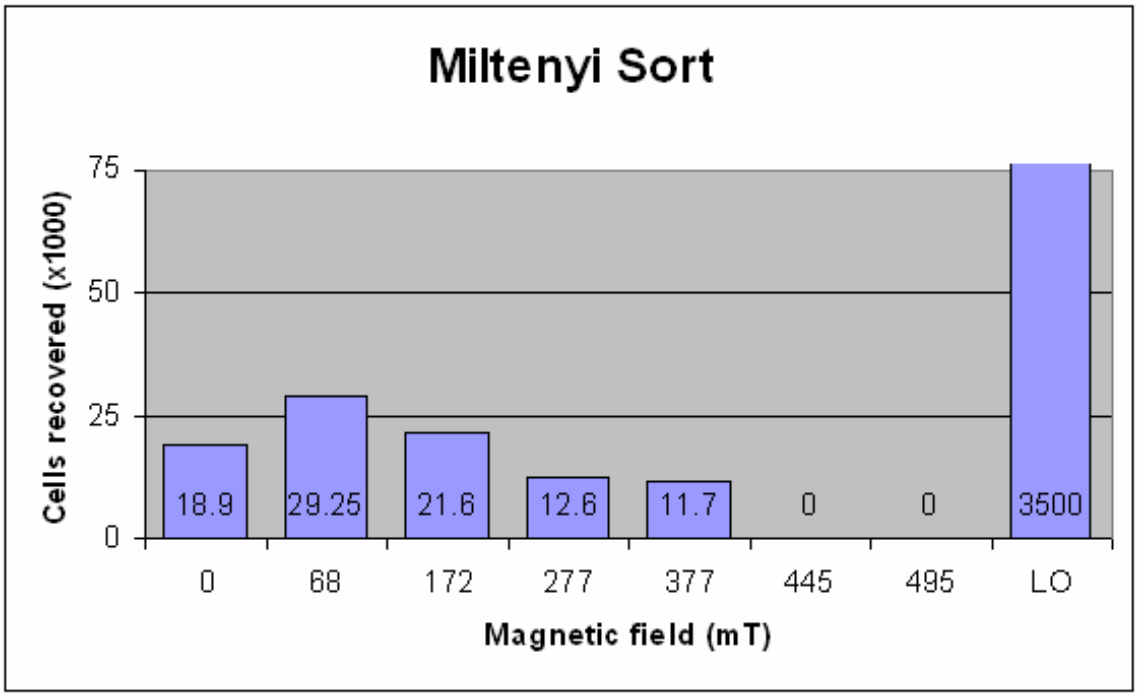

Figure 4.7. Magsort results with Miltenyi bead labeled cells

Distribution of cells after a model cell mixture sorting experiment. Magsort was loaded with a 50% mixture of KG-1a/CEM cells, 6 x 10^6 cells total. KG-1a cells were labeled with anti-human CD34-coated Miltenyi beads and anti-human CD34 PE-conjugated antibodies. A regular incremental program was used. Chambers 1-5 had a very small number of cells after the sort, but they did not display increased fluorescence ratio. The last two magnets did not sort any cells. 98% of all recovered cells were found in the unsorted fraction. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction on the left and the ninth fraction were not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets.

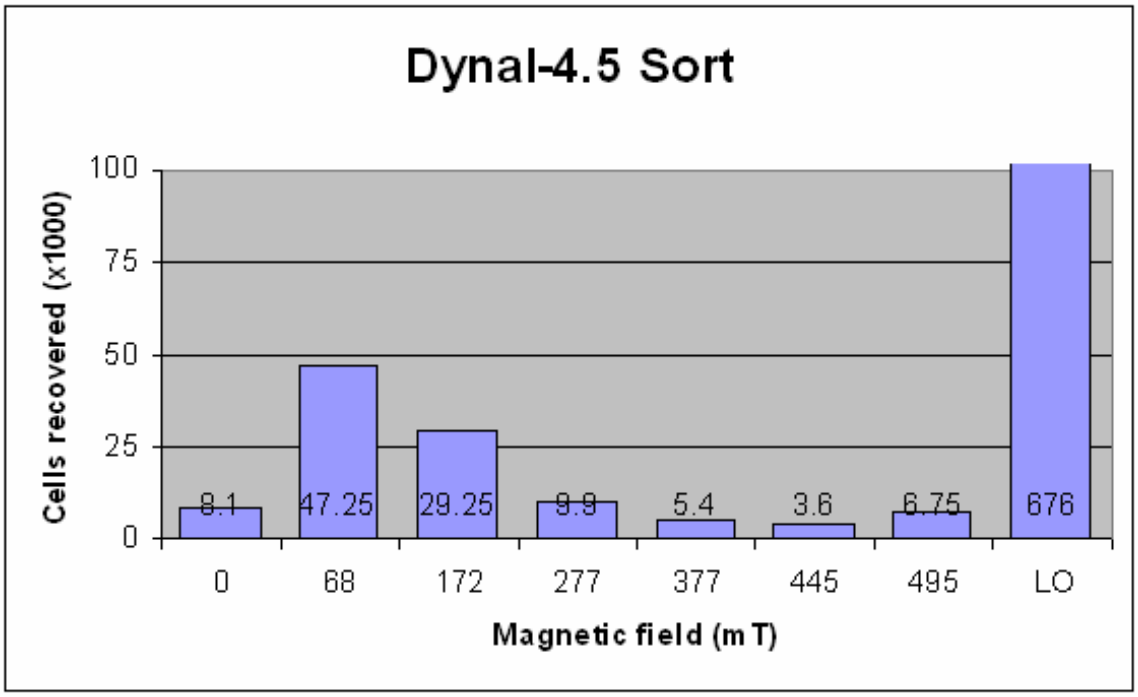

Figure 4.8. Magsort results with Dynal 4.5 µm bead labeled cells

Distribution of cells after a model cell mixture sorting experiment. Magsort was loaded with a 50% mixture of KG-1a/PBMC cells, 2.6 x 10^6 cells total. KG-1a cells were labeled with anti-human CD34-biotin conjugated antibodies, streptavidin coated Dynal beads, and anti-human CD34 PE-conjugated antibodies. A regular incremental program was used. All sort chambers had a small number of cells after the sort, but none of the fractions displayed increased fluorescence ratio. 86% of all recovered cells were found in the unsorted fraction. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction on the left and the ninth fraction were not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets.

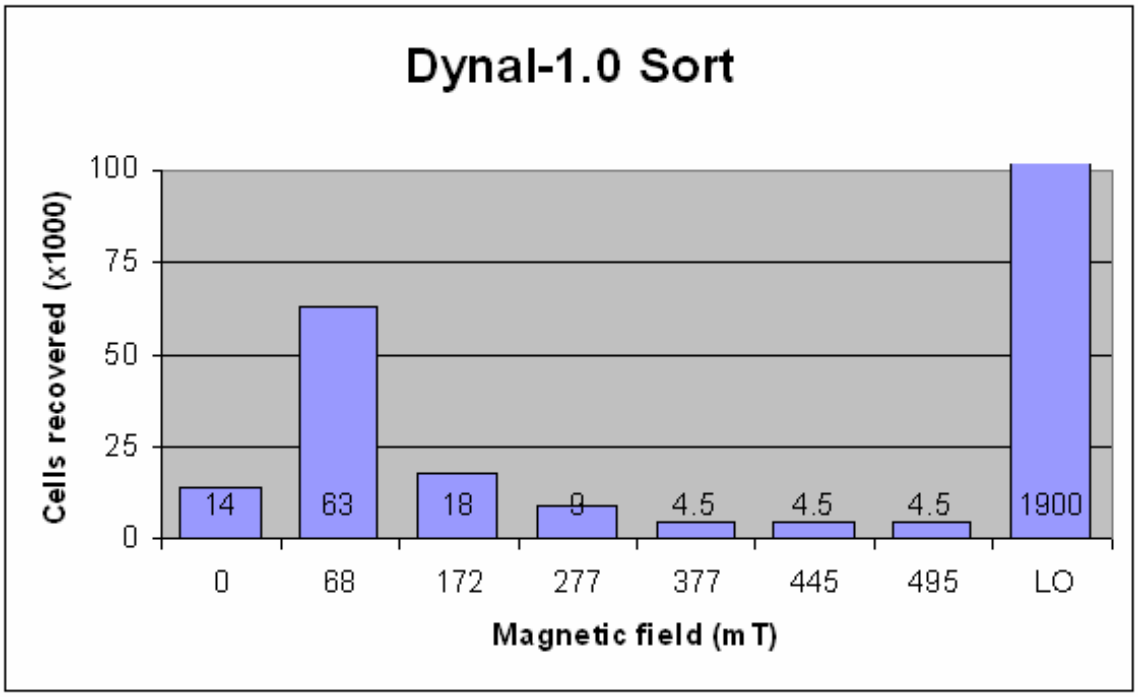

Figure 4.9. Magsort results with Dynal 1.0 µm bead labeled cells

Distribution of cells after a model cell mixture sorting experiment. Magsort was loaded with a 50% mixture of KG-1a/CEM cells, 4x 10^6 cells total. KG-1a cells were labeled with anti-human CD34-biotin conjugated antibodies, streptavidin coated Dynal beads, and anti-human CD34 PE-conjugated antibodies. A regular incremental program was used. All sort chambers had a small number of cells after the sort, but none of the fractions displayed increased fluorescence ratio. 93% of all recovered cells were found in the unsorted fraction. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction on the left and the ninth fraction were not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets.

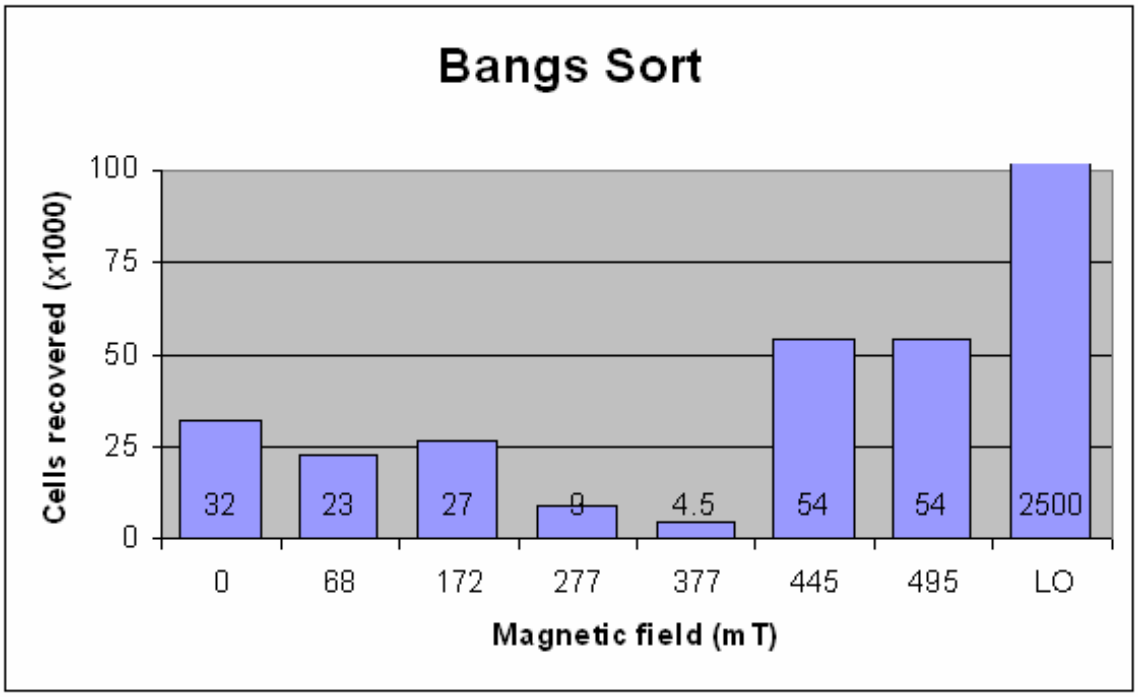

Figure 4.10. Magsort results with Bangs bead labeled cells

Distribution of cells after a model cell mixture sorting experiment. Magsort was loaded with a 50% mixture of labeled and unlabeled KG-1a cells, 4.6 x 10^6 cells total. Half of the cells were labeled with anti-human CD34-biotin conjugated antibodies, streptavidin coated Bangs beads, and anti-human CD34 PE-conjugated antibodies. A regular incremental program was used. All sort chambers had a small number of cells after the sort, but none of the fractions displayed increased fluorescence ratio. 92% of all recovered cells were found in the unsorted fraction. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction on the left and the ninth fraction were not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets.

Test Separation of Model Cell Mixtures Using the Optimized Protocol with Percoll Layering

Besides the greater than manageable distance between the cells and the magnets another cause for the lack of antigen specific sorting could have been that once the cells settled on the plastic surface in the sample chamber it was too hard to break them off and get them to start moving upward. To shorten the distance between the cells and the sorting magnet and to reduce the necessary force the cells needed to start moving we layered Percoll underneath the sample. Percoll is a dense, non-toxic liquid often used for density gradient cell centrifugation. In our experiments the cells sedimented in the cell medium of the load chamber onto the Percoll surface. This brought the cells closer to the magnets and created a surface from where the cells could be pulled off easier than from the plastic surface of the bottom of the sample chamber. Figure 4.11. illustrates the modified loading and sorting conditions. We tested all four bead types using the earlier described model cell mixtures following the optimized protocol we developed with Percoll layering.

Figure 4.12. demonstrates that anti CD34+ antibody coated Miltenyi beads were still unable to selectively deliver the labeled KG-1a cells into the sort chambers from a 50% KG-1a/CEM cell mixture. Although a few cells were recovered from each sort chamber these fractions were not purified or enriched for CD34+ cells. Besides the immuno-magnetic Miltenyi bead labeling CD34+ cells were also labeled with anti-human CD34 PE-conjugated antibodies and the CD34+ cell ratio of each fraction was verified by fluorescent microscopy. Figure 4.13 reveals that there was no significant difference between the initial sample and the sorted fraction #6 (magnet strength 445 mT) in terms of fluorescent cell ratio. Just like in the initial cell sorting experiments, the cells found in the sort chambers were delivered there by a non-specific process. It is possible that the mass movement of unbound magnetic beads stirred the cell medium in the sample chamber enough for a few buoyant cells to float up into the sort chambers.

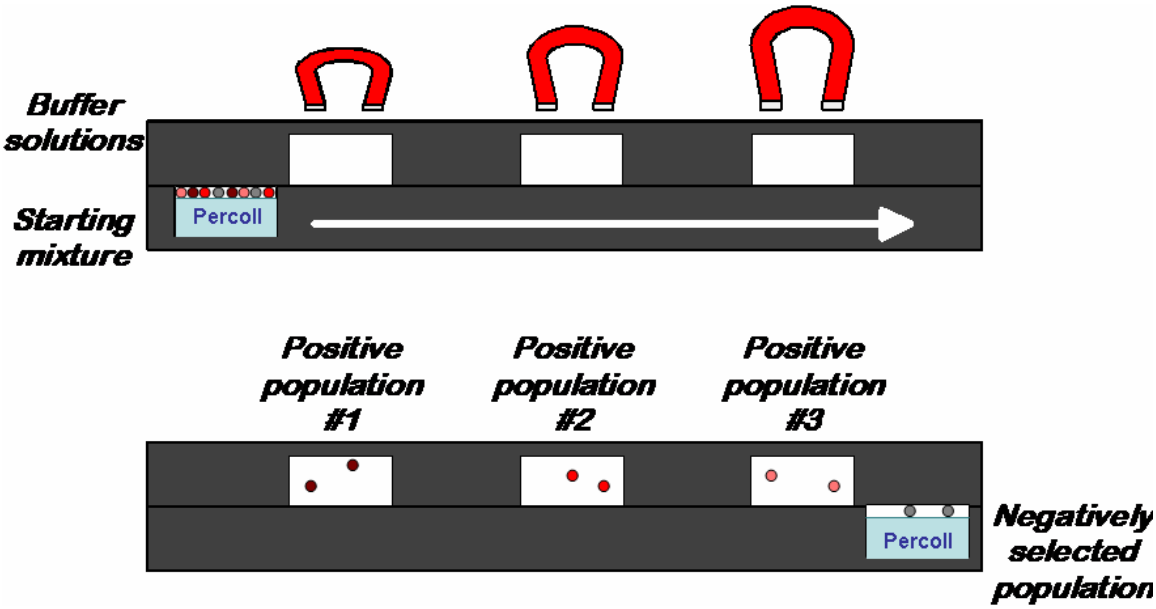

Figure 4.11. General principle of Percoll mediated Magsort purification of cells

Percoll is layered into the sample insertion chamber underneath the cell medium containing the cell sample. The cells settle on top of the Percoll layer that brings them closer to the magnets and creates a surface from where they can be pulled off easily.

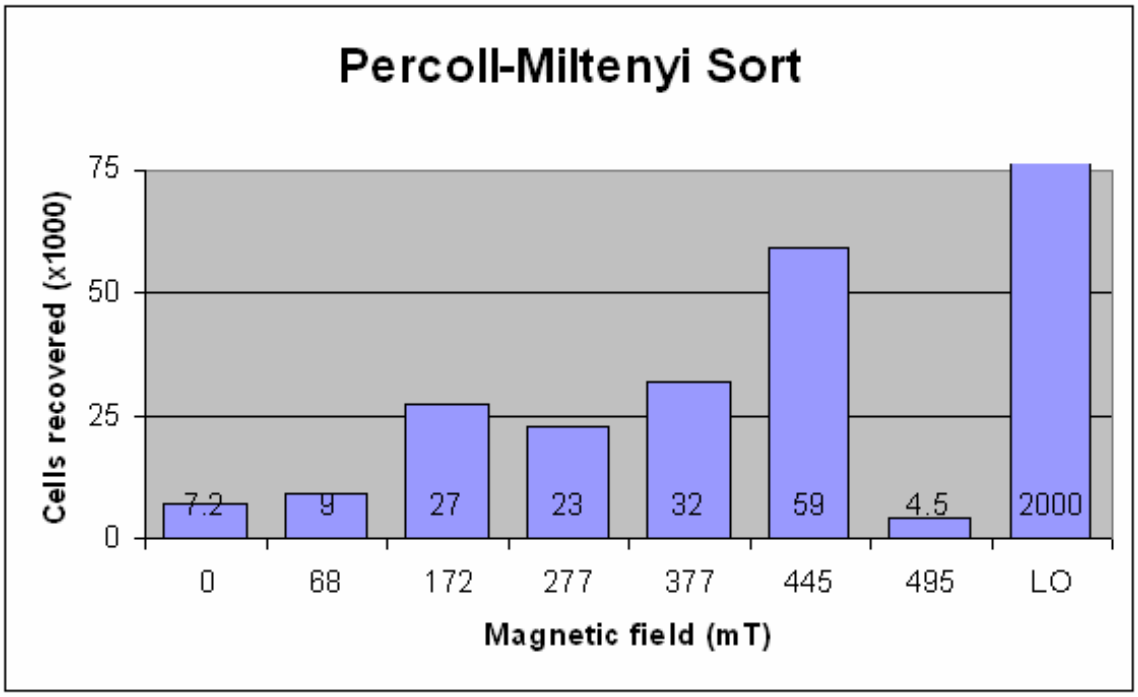

Figure 4.12. Magsort results with Miltenyi bead labeled cells after Percoll mediated cell sorting

Distribution of cells after a Percoll-mediated model cell mixture sorting experiment. Magsort was loaded with a 50% mixture of KG-1a/CEM cells, 4.9 x 10^6 cells total. KG-1a cells were labeled with anti-human CD34-coated Miltenyi beads and anti-human CD34 PE-conjugated antibodies. An optimized incremental program was used. All chambers had a small number of cells after the sort, but none of the fractions displayed increased fluorescent cell ratio. 93% of all recovered cells were found in the unsorted fraction. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction on the left and the ninth fraction were not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets.

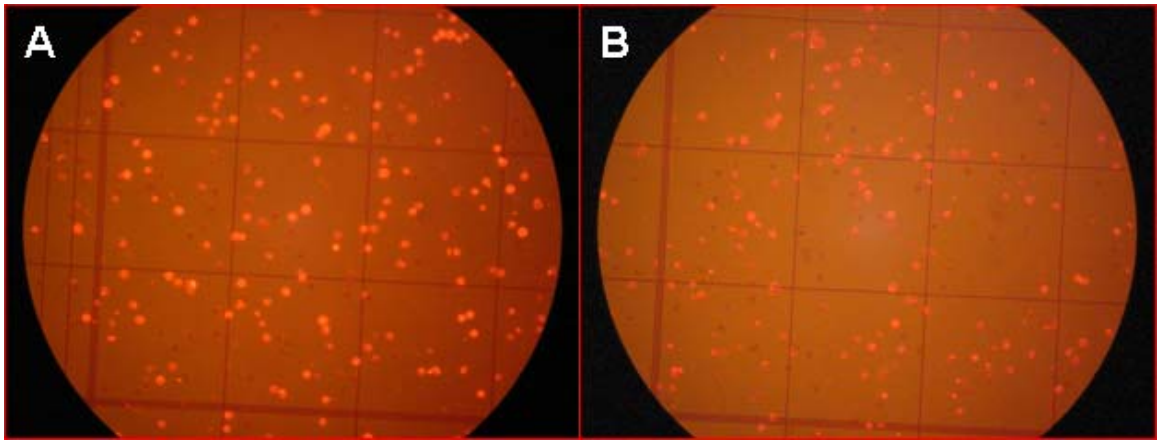

Figure 4.13. Comparison of fluorescent cell ratios before and after Percoll mediated Magsort purification of KG-1a cells labeled with Miltenyi beads

Fluorescence microscopy images of the initial sample and the sorted fraction recovered from Chamber #6. Magsort was loaded with a 50% mixture of KG-1a/CEM cells, 4.9 x 10^6 cells total. KG-1a cells were labeled with anti-human CD34-coated Miltenyi beads and anti-human CD34 PE-conjugated antibodies. An optimized incremental program was used and the recovered fractions were examined by fluorescence microscopy. There was no significant difference in fluorescent cell ratio between the sample from before (A) and after the sort (B).

The calculated amount of iron in a Miltenyi bead with 50 nm diameter is about 8000-fold less than in a single Dynal bead with 1.0 μm diameter or 5000 times less than in a single Bangs bead with 0.86 μm diameter and the difference is even greater when compared to a larger Dynal bead. Even the reported labeling ratio of several hundreds of Miltenyi beads per labeled cell (62) amounts to a magnetic force that is much smaller than with just one of the other magnetic beads.

In Magsort experiments the larger (4.5 μm) Dynal beads led to aggregation of beads and cells when exposed to magnetic fields. Some of these aggregates were attracted to the smallest magnet used (68 mT at the poleface) others remained in the unsorted fraction. When the cells in the sorted fraction were examined by fluorescence microscopy the fluorescence ratio was found to be virtually unaltered. These cells might have been delivered into the sort chamber by either an non-specific process (floating cells) or a specific process delivering many non-specific cells to the sort chamber (unlabeled cells stuck in a mesh of both beads and labeled cells) In addition to aggregation, this type of bead was not suitable for quantitative labeling experiments because of the small number of beads that can be bound to one cell.

Figure 4.14. presents the results of a Percoll mediated model cell mixture sorting experiment utilizing Bangs bead labeling. The Magsort instrument was loaded with an approximately 50% mixture of labeled and unlabeled KG-1a cells, 3.8×10^6 cells total. The labeled cells were labeled with anti human CD34, biotin conjugated antibodies, streptavidin coated Bangs beads, and anti-human CD34 PE-conjugated antibodies. An optimized incremental program with one hour cell settling time was used exposing the cells to the complete array of all six magnets. All sorted fractions were collected and the recovered cells were analyzed by microscopy and flow cytometry. There were no signs of toxic effects on the cells caused by the labeling and sorting procedures. The Magsort related cell loss was 24%. The number of recovered cells and the ratio of bead labeled cells in each fraction were determined by microscopy.

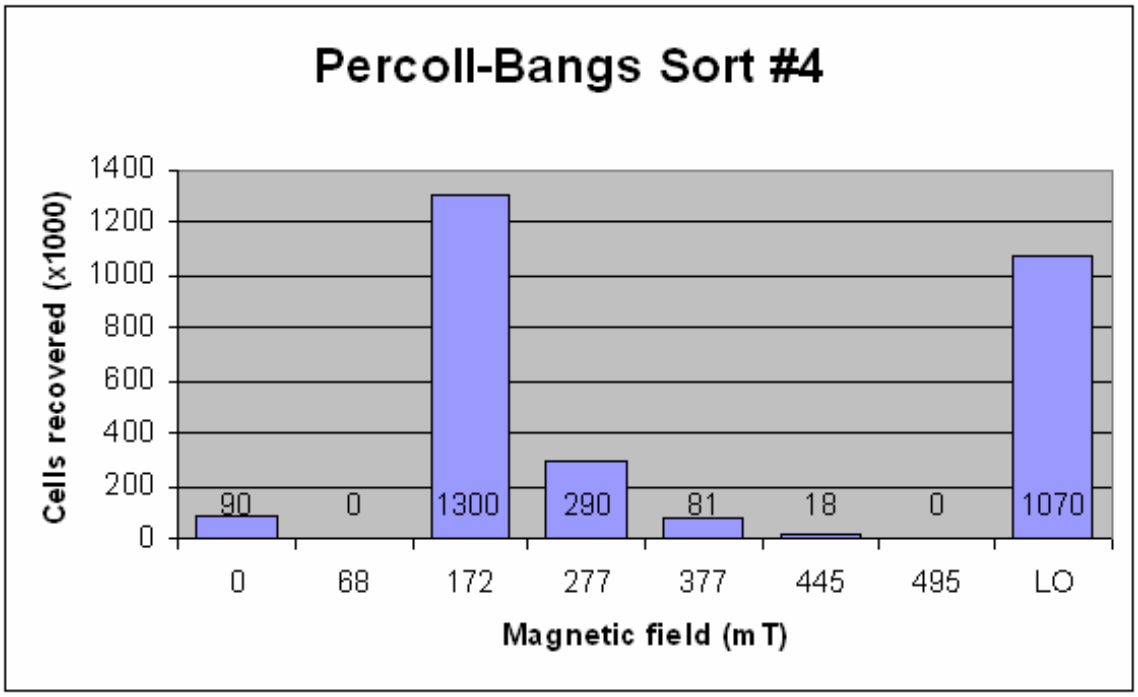

Figure 4.14. Percoll mediated Magsort purification of Bangs bead labeled KG-1a cells

Distribution of KG-1a cells after a Percoll-mediated model cell mixture sorting experiment. Magsort was loaded with an approximately 50% mixture of labeled and unlabeled KG-1a cells, 3.8 x 10^6 cells total. The labeled cells were labeled with anti human CD34-biotin conjugated antibodies, streptavidin-coated Bangs beads, and anti-human CD34-PE-conjugated antibodies. An optimized incremental program was used. Fraction #3 corresponding to Magnet #2 (172 mT) contained 46% of all recovered cells. An additional 10% were found in the next fraction. The unsorted fraction contained 38% of all recovered cells. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction was not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets

Fraction #3 corresponding to Magnet #2 (172 mT) contained 46% of all recovered cells and was 87% bead-positive. Fraction #4 had an additional 10% of all cells and 80% of this fraction was found to be bead-labeled. The number of sorted cells and the bead-positive/bead-negative ratio in consecutive fractions and the remaining sample displayed a decreasing pattern, with the complete depletion of the bead-positive cells from the leftover fraction (Fraction #8) by microscopy. There was also a strong correlation between the average number of bound beads/cell and the magnet strength that sorted the fraction, with labeled cells sorted by weaker magnets binding a higher average number of beads than labeled cells sorted by stronger magnets. Table 4.1. summarizes the results of this experiment and presents a good correlation between the microscopic and flow cytometric data.

Fraction #	Magnet #	Magnetic Field (mT)	Percent of all Cells	Percent Bead-Positive Cells	Percent with Many Beads	Percent Flow Positive Cells	Mean PE by Flow
3	2	172	45.63	87.1	62.9	78.65	297.53
4	3	277	10.17	80.0	26.8	61.98	244.46
5	4	377	2.84	73.3	16.7	40.63	225.24
8 (LO)	-	-	37.56	-	--	5.34	-

Table 4.1. Summary of results for KG-1a cells after Percoll mediated Magsort purification

Fraction #8 was the unsorted (leftover) fraction after passage under all magnets. The "Percent of all cells" column displays what portion of all recovered cells was found in the given fraction. The ratio of bead-positive cells was determined by microscopy. "Percent with many beads" represents the portion of cells in a given fraction with more than 4 beads attached to them. "Percent flow positive cells" were the ratio of PE-positive cells determined by flow cytometry. "Mean PE by flow" was the mean PE value for the positive cells (in the R1 region on the flow cytometry scatterplots) in a given fraction.

Flow cytometry was performed on all fractions that had sufficient cell numbers. In the flow cytometry scatterplots displayed in Figures 4.15. and 4.16. instrument settings and R1 boundaries remained unchanged throughout the analyses.

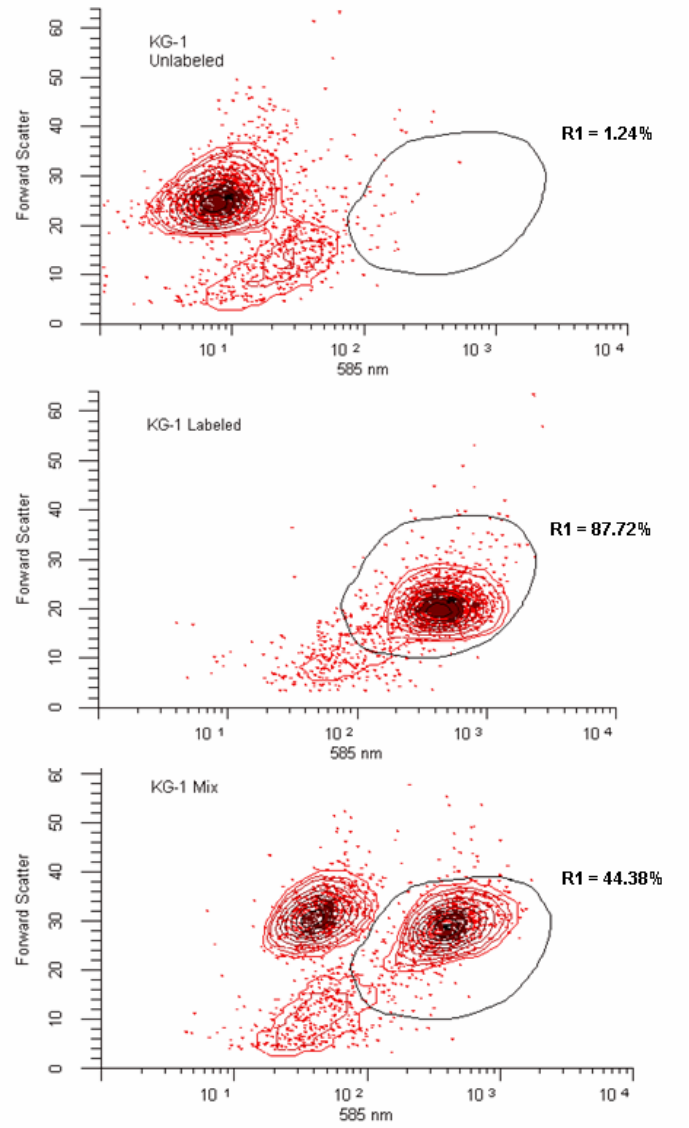

Figure 4.15. Flow cytometry scatterplots of KG-1a cell samples before Percoll mediated Magsort cell purification

The scatterplots display unlabeled KG-1a cells, labeled KG-1a cells, and a mixture of the two samples. The labeled cells were labeled with anti human CD34-biotin conjugated antibodies, streptavidin-coated Bangs beads, and anti-human CD34-PE-conjugated antibodies. PE-fluorescence was acquired at 585nm wavelength. The R1 zone represents the fluorescently labeled CD34+ cells. Instrument settings and R1 boundaries remained unchanged throughout the experiment. About 44% of all cells were in the R1 region of the mixture, but since about 10% of all cells were cell debris, the PE-positive portion of the cell mixture is approximately 50%.

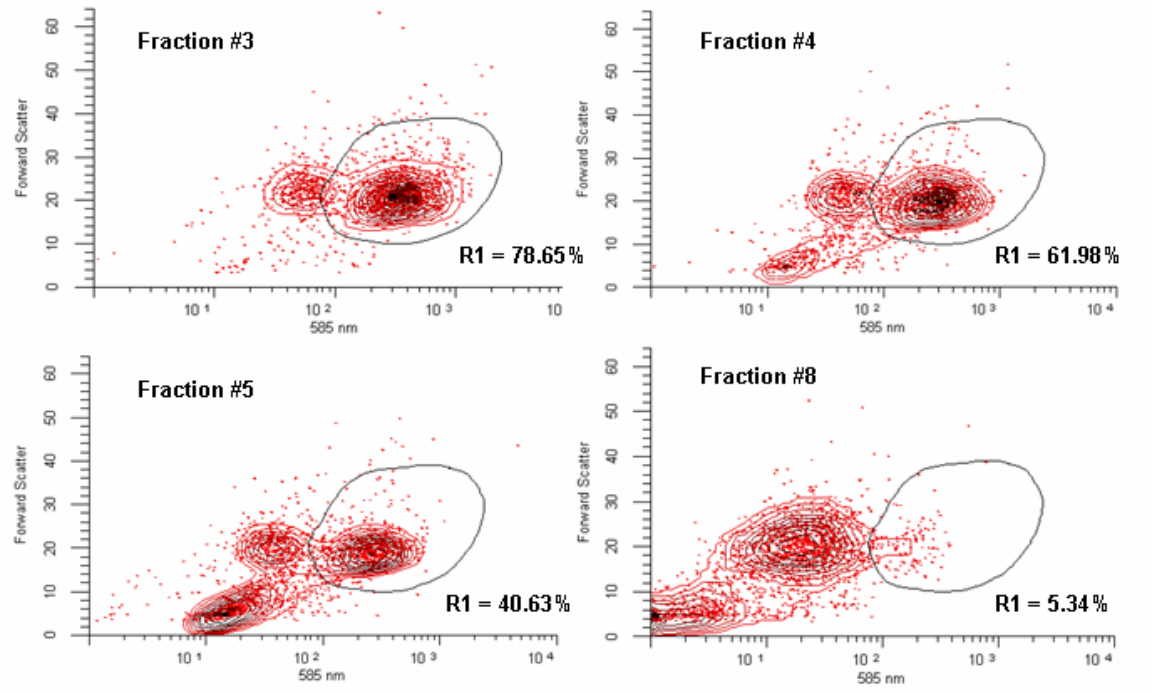

Figure 4.16. Flow cytometry scatterplots of KG-1a cell sample fractions after Percoll mediated Magsort cell purification

The scatterplots display four sort fractions of KG-1a cells sorted from an approximately 50% cell mixture of labeled and unlabeled cells based on their magnetophoretic mobility. The labeled cells were labeled with anti human CD34-biotin conjugated antibodies, streptavidin-coated Bangs beads, and anti-human CD34-PE-conjugated antibodies. PE-fluorescence was acquired at 585nm wavelength. The R1 zone represents the fluorescently labeled CD34+ cells. Instrument settings and R1 boundaries remained unchanged throughout the experiment. Consecutive fractions (fractions #3, #4, and #5) contained gradually decreasing percentage of PE labeled cells. The leftover fraction only contained 5.34% lightly PE-positive cells.

The R1 zone represents the fluorescently labeled CD34+ cells. Figure 4.15. displays the flow cytometry scatterplots of unlabeled KG-1a cells, labeled KG-1a cells, and a mixture of the two samples. The cell mixture loaded into the system was approximately 50% PE-positive (excluding the 10% cell debris from the calculations). Figure 4.16. demonstrates that Fraction #3 corresponding to Magnet #2 (172 mT) contained 78.65% PE-positive cells. The next two fractions contained 61.98% and 40.63% labeled cells displaying a gradually decreasing ratio of labeled cells in consecutive fractions. The last fraction contained cells that passed under all magnets and remained unmoved. This leftover fraction was depleted of PE-positive cells to 5.34%. These results confirmed that the newly developed method specifically sorted CD34 antigen labeled cells based on their magnetophoretic mobility.

To test whether Percoll mediated multistage magnetic sorting also worked with other antigens and cell types we purified CEM cells from an approximately 50% mixture of labeled and unlabeled cells. Figure 4.17. displays the results of this sorting experiment utilizing Bangs bead labeling. The Magsort instrument was loaded with an approximately 50% mixture of labeled and unlabeled CEM cells, 3.3×10^6 cells total. The labeled cells were labeled with anti human CD4, biotin conjugated antibodies, streptavidin coated Bangs beads, anti-human CD4 PE-conjugated antibodies, and anti-human CD4 Alexa-488 conjugated antibodies. An optimized incremental program with one hour cell settling time was used exposing the cells to the complete array of all six magnets. All sorted fractions were collected and the recovered cells were analyzed by microscopy and flow cytometry. There were no signs of toxic effects on the cells caused by the labeling and sorting procedures. The Magsort related cell loss was 30%. The number of recovered cells and the ratio of bead labeled cells in each fraction were determined by microscopy. Fraction #3 corresponding to Magnet #2 (172 mT) contained 49% of all recovered cells and was 85% bead-positive. Fraction #4 contained an additional 17% of all cells and 86% of this fraction was found to be bead-labeled. The bead-positive/bead-negative ratio of the other fractions and the remaining sample displayed a decreasing pattern, with the

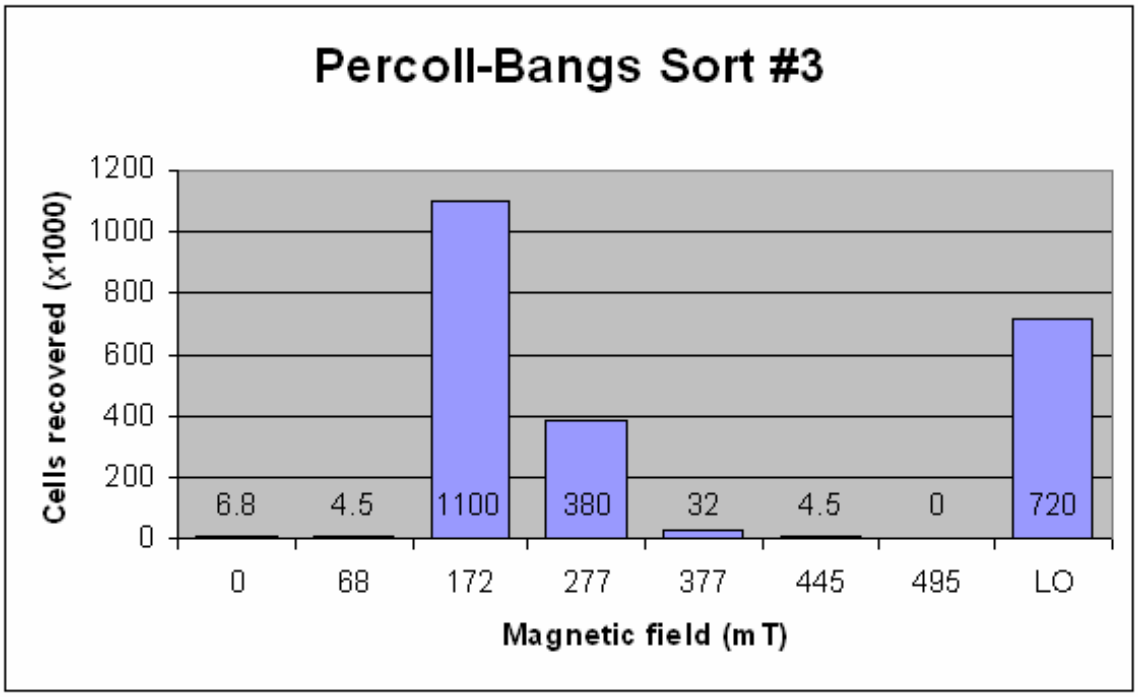

Figure 4.17. Percoll mediated Magsort purification of Bangs bead labeled CEM cells

Distribution of CEM cells after a Percoll-mediated model cell mixture sorting experiment. Magsort was loaded with a 50% mixture of labeled and unlabeled CEM cells, 3.3 x 10^6 cells total. The labeled cells were labeled with anti human CD4, biotin conjugated antibodies, streptavidin coated Bangs beads, and anti-human CD4 Alexa 488-conjugated antibodies. An optimized incremental program was used. Fraction #3 corresponding to Magnet #2 (172 mT) contained 49% of all recovered cells. An additional 17% were found in the next fraction. The unsorted fraction contained 32% of all recovered cells. Each fraction is represented by the magnetic field at the poleface of permanent magnets used to collect it. The first fraction was not exposed to any of the magnets. LO: residual particles left in sample chamber after passage under all magnets

depletion of the positive cells from the leftover fraction to 4.5% by microscopy. There was also a strong correlation between the average number of bound beads/cell and the magnet strength that sorted the fraction, with labeled cells sorted by weaker magnets binding a higher average number of beads than labeled cells sorted by stronger magnets. Table 4.2. summarizes the results of this experiment.

Fraction #	Magnet #	Magnetic Field (mT)	Percent of all Cells	Percent Bead-Positive Cells	Percent with Many Beads	Percent Flow Positive Cells	Mean Alexa-488 by Flow
3	2	172	48.93	94.9	58.1	66.83	260.00
4	3	277	16.90	85.5	51.3	56.89	241.94
5	4	377	1.42	76.0	20.0	37.76	200.36
9 (LO)	-	-	32.03	4.5	--	1.90	-

Table 4.2. Summary of results for CEM cells after Percoll mediated Magsort purification

Fraction #9 was the unsorted (leftover) fraction after passage under all magnets. The "Percent of all cells" column displays what portion of all recovered cells was found in the given fraction. The ratio of bead-positive cells was determined by microscopy. "Percent with many beads" represents the portion of cells in a given fraction with more than 4 beads attached to them. "Percent flow positive cells" were the ratio of Alexa-488 - positive cells determined by flow cytometry. "Mean Alexa-488 by flow" was the mean Alexa-488 value for the positive cells (in the R1 region on the flow cytometry scatterplots) in a given fraction.

Flow cytometry was performed on all fractions that had sufficient cell numbers. In the flow cytometry scatterplots displayed in Figures 4.18. and 4.19. instrument settings and R1 boundaries remained unchanged throughout the analyses. The R1 zone represents the fluorescently labeled CD4+ cells. Figure 4.18. displays the flow cytometry scatterplots of unlabeled CEM cells, labeled CEM cells, and a mixture of the two samples. The cell mixture loaded into the system was approximately 50% PE-positive. Figure 4.19 demonstrates that fraction #3 corresponding to Magnet #2 (172 mT) contained 78.65% PE positive cells. The next two fractions contained 61.98% and 40.63% labeled cells displaying a gradually decreasing ratio of labeled cells in

consecutive fractions. The last fraction contained cells that passed under all magnets and remained unmoved. This leftover fraction was depleted of PE-positive cells to 5.34%. These results confirm that Percoll mediated Magsort purification, utilizing Bangs beads with 0.86 μm diameter can be applied to different cell types with different labeled surface antigens. Similar results could be generated with Dynal 1.0 μm magnetic beads, but since these microparticles were originally designed for molecular biology applications and not for cell purification these beads were found to be moderately toxic to live cells. For this reason we concluded that Percoll mediated Magsort purification worked best using Bangs beads with 0.86 μm diameter.

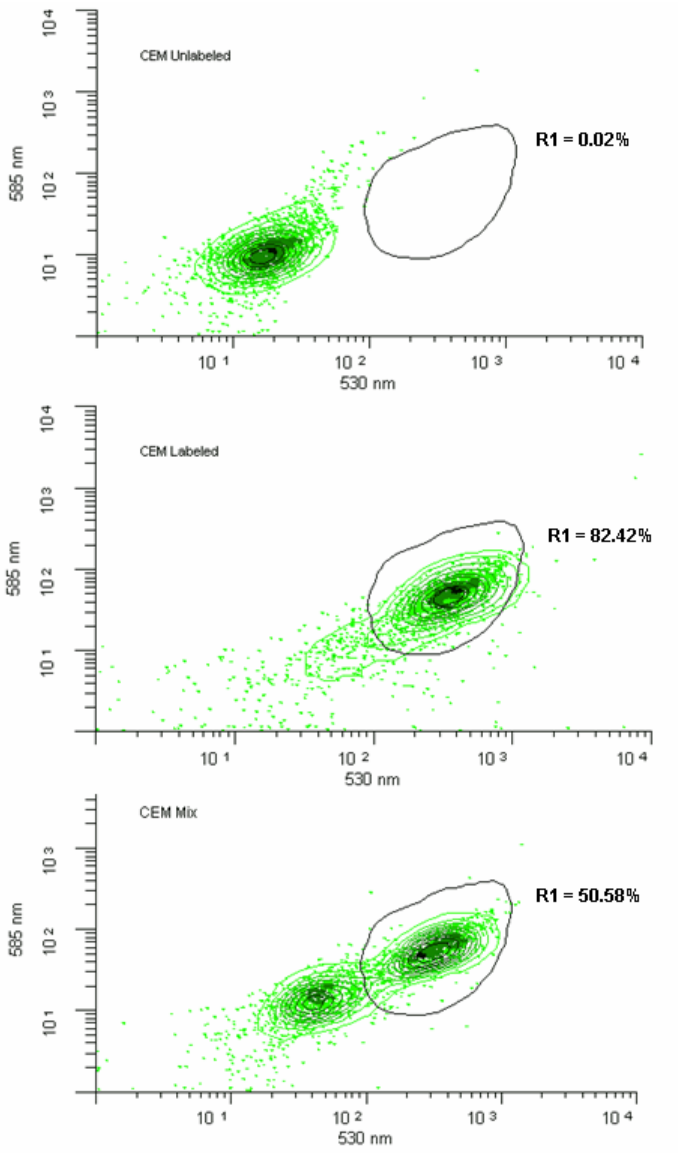

Figure 4.18. Flow cytometry scatterplots of CEM cell samples before Percoll mediated Magsort cell purification

The scatterplots display unlabeled CEM cells, labeled CEM cells, and a mixture of the two samples. The labeled cells were labeled with anti human CD4-biotin conjugated antibodies, streptavidin-coated Bangs beads, and anti-human CD4-Alexa 488-conjugated antibodies. Alexa 488-fluorescence was acquired at 530 nm wavelength. The R1 zone represents the fluorescently labeled CD4+ cells. Instrument settings and R1 boundaries remained unchanged throughout the experiment. About 50% of all cells were in the R1 region of the mixture with very few cell debris.

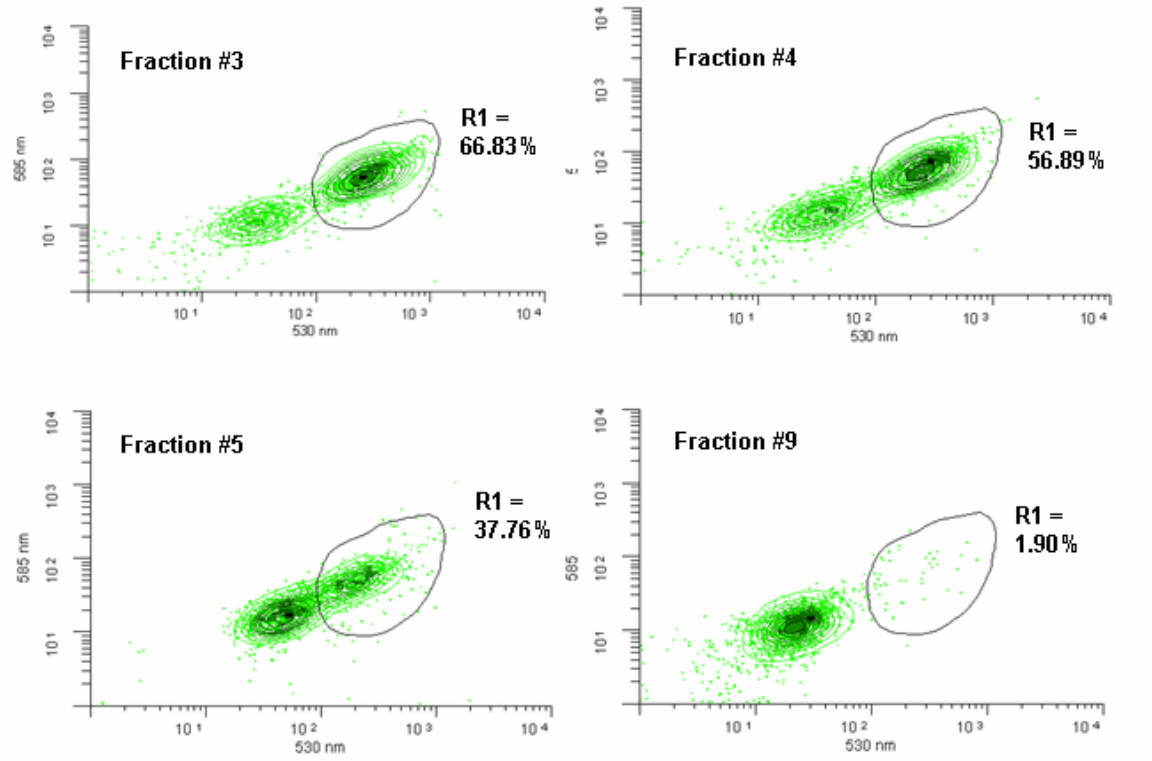

Figure 4.19. Flow cytometry scatterplots of CEM cell sample fractions after Percoll mediated Magsort cell purification

The scatterplots display four sort fractions of KG-1a cells sorted from an approximately 50% cell mixture of labeled and unlabeled cells based on their magnetophoretic mobility. The labeled cells were labeled with anti human CD4-biotin conjugated antibodies, streptavidin-coated Bangs beads, and anti-human CD4-Alexa 488-conjugated antibodies. Green fluorescence was acquired at 530 nm wavelength. The R1 zone represents the fluorescently labeled CD34+ cells. Instrument settings and R1 boundaries remained unchanged throughout the experiment. Consecutive fractions (fractions #3, #4, and #5) contained gradually decreasing percentage of PE labeled cells. The leftover fraction only contained 1.9% lightly PE-positive cells.

DISCUSSION

We have tested the applicability of multistage magnetic sorting, a new cell sorting technology to purify selected cell subsets from live, biohazardous cell samples in a closed system for consecutive microarray analysis. Multistage magnetic sorting is able to separate magnetic particles and magnetically labeled cells based on their magnetophoretic mobility by exposing them to gradually increasing magnetic fields. Several fractions can be collected separately based on the strength of the magnetic field needed to move the cells or particles into the collection chamber. Whereas most published work in the area of magnetic cell sorting has made use of high magnetic field gradients (which inevitably results in binary sorting), Magsort uses low magnetic field gradients but rather high permanent-magnet fields (62,65-68). In any magnetically labeled cell population the main source of heterogeneity is considered to be the distribution of receptor sites among cells. The more of the labeled receptor a cell expresses, the more magnetically labeled it gets. The level of magnetization of the labeling particles is considered to be fairly even; therefore it does not contribute significantly to the differences in magnetization of the labeled cells. Based on these assumptions magnetically labeled particles and cells were sorted into fractions according to their degree of magnetization utilizing Magsort technology.

In the first set of experiments Magsort was tested using mixtures of different magnetic microparticles. It proved to be capable of separating each type of microparticle to 75-98% purity from approximately even mixtures of two different particle types. To some extent it also sorted several subsets of the main particle types into separate fractions. However these subset separations resulted in low purities. We concluded that the magnetic particle experiments demonstrated the potential of this method for successful cell sorting.

We tested Magsort using several different cell types, antibodies, magnetic bead types, loading methods and sort programs. Out of these variables the magnetic bead type and the loading method used proved to be the most important factors determining the

outcome of the sort experiments. We used model cell mixtures to evaluate the capabilities and limitations of this technology on well-defined samples. We loaded samples of magnetically and fluorescently labeled cells mixed with unlabeled cells at approximately 50% mixture ratios into the system and analyzed the sorted fractions by fluorescence microscopy and flow cytometry.

We tested four different magnetic bead types: Dynal beads with 4.5 μm diameter, Dynal beads with 1.0 μm diameter, Bangs beads with 0.86 μm average diameter, and Miltenyi beads with 50 nm diameter. We demonstrated that while Magsort was capable of sorting each magnetic bead type alone using the original protocols it was not able to selectively pull magnetically labeled cells from the bottom of the sample chamber into the sort chamber regardless to which magnetic bead type was used. The vast majority (86-98%) of all recovered cells remained in the sample insertion chamber. The few cells that did end up in the sort chambers were not delivered there selectively based on their immuno-magnetic labeling, since the cell mixture ratio of these fractions remained unaltered compared to the initial sample. This was invariably the case with all cell types, and magnets tested using any user definable sort program.

We tested the effects of extending or shortening the time allowed for initial cell settling and for the individual sort periods, but even when the settling time was omitted and the sort time was extended to several hours using the strongest magnet (495 mT) the Magsort instrument did not produce specifically purified or enriched cell subsets. Shortening or omitting the initial cell settling period did result in slightly increased numbers of cells in the first two sort fractions but the labeled and unlabeled cell ratio in these fractions was very similar to that of the initial, unsorted cell mixture. The optimal initial settling time before applying the first magnet was found to be 30-60 minutes. Increasing or decreasing the number of cells in the initial cell mixture loaded into the system did not have any effect on the purity of the sort fractions but starting with too many or too few cells did result in increased cell loss. The optimal sample size was found to be 1-6x10^6 cells. Loading this size of initial samples into the Magsort instrument resulted in 10-25% cell loss measured by the loaded/recovered cell ratio. We concluded

that the distance between the bottom of the sample chamber and the magnets were too great for the magnets to pull up the labeled cells from the sample chamber into the sort chambers. Another problem with the original protocol appeared to be that once the cells settled on the plastic surface in the sample chamber it was too hard to break them off and get them to start moving upward.

To solve these problems we needed to shorten the distance between the cells and the sorting magnet and also reduce the necessary force the cells needed to start moving. We layered Percoll underneath the sample to bring the cells closer to the magnets and to create a surface from where the cells could be pulled off easier than from the plastic surface of the bottom of the sample chamber. The overall effect of this approach depended on the magnetic bead type used for labeling. Cells labeled with Dynal beads (1.0 μm diameter) or with Bangs beads (0.8-1 μm diameter) selectively moved up into the sort chambers in the increasing magnetic field. Miltenyi beads with their 50 nm diameter were too small to lift the cells up even under the modified conditions. Dynal beads with 4.5 μm diameter formed large, multicell-multibead aggregates during sorting that prevented individual cell sorting.

We demonstrated that when using Bangs beads and Percoll with optimized sample loading and running protocols, Magsort was capable of sorting up to 4 distinguishable cell fractions with different labeled/unlabeled cell ratios of immuno-magnetically labeled cells. The maximum purity we could achieve with this method was 75-80%. We had similar results with PE-conjugated, anti-CD34 antibody labeled KG-1 cells and Alexa-488-conjugated anti-CD4 antibody labeled CEM cells suggesting that the most important factors in Multistage Magnetic Sorting experiments are the sample loading method (with Percoll or other dense, non-toxic fluid) and the type of magnetic beads used. Different cell types (CEM, KG-1a), once labeled, behaved similarly and labeling different antigens (CD34, CD4) did not affect the sort results noticeably. In our experiments Bangs beads with 0.86 μm average diameter were the most effective in sorting. Dynal beads with 1.0 μm diameter were somewhat effective in cell sorting but since these beads were originally designed for molecular biology applications and not for

cell labeling, they were highly toxic to live cells. The other two bead types we tested were not feasible for multistage magnetic sorting applications for the reasons described above. The type of cells to be sorted, the antigen and antibody used for labeling, the fluorochrome used for labeling verification, and the sort program applied (other than the initial cell settling time) did not significantly affect the outcome of the sort experiments.

Although the achieved 70-80% purity barely reaches the minimum requirements we established in the first section of this study for meaningful microarray experiments and the technology is inferior to traditional flow cytometric cell sorting and column-based magnetic bead mediated cell sorting in accuracy, reproducibility, and ease of handling, we concluded that multistage magnetic sorting may be useful in processing special samples where the traditional methods are not feasible, for example in purifying small, live, potentially biohazardous cell samples.

CHAPTER 5.

LASER MEDIATED CELL SORTING AND MANIPULATION FOR MICROARRAY ANALYSIS – LEAP

INTRODUCTION

The objective of the experiments described in this chapter was to test whether using Laser Enabled Analysis and Processing it was possible to purify mixed cell samples and manipulate selected cells prior to microarray analysis without significantly altering the target cells' GEP. Multistage magnetic cell sorting (described and discussed in chapter 4.) proved to be capable of sorting small, mixed cell samples into up to four distinguishable fractions based on the cell subsets' surface antigen expression and reaching 70-80% purity in some of these fractions. We tested another new technology, Laser Enabled Analysis and Processing (88-91), for highly accurate cell purification of homogeneous cell subsets from small, biohazardous samples. We also tested the unique capability of laser mediated micromanipulation of selected cells offered by this new technology. Most of the results presented here have been published (72).

Laser Enabled Analysis and Processing (LEAP) is part of the new generation of scanning cytometry technologies making use of the wide variety of fluorescent molecular probes now available. These new technologies have many of the features and power of multiparameter flow cytometry as well as the advantages of an imaging technology. Recently developed instruments combine some of the capabilities of scanning cytometry with the ability to manipulate cells. LEAP provides many of the advantages of Laser Capture Microdissection and Laser Microdissection and Pressure Catapulting technologies in manipulating single, live cells and offers a method of live cell sorting using laser ablation or laser catapulting (86). In addition, unlike other imaging technologies, the LEAP instrument also provides a convenient method for high-speed microinjection of macromolecules into living cells using a pulsed laser ("laser optoinjection") set at sub-lethal energies. This optoinjection is very fast (hundreds of

cells/sec) and, unlike electroporation, has a very low rate of injury to cells which can be individually selected on the basis of multiple fluorescent probes in an automated molecular imaging process. With these capabilities LEAP offers a new method of sorting or optoinjecting live, potentially biohazardous cells under sterile conditions in a closed system.

The LEAP instrument (Figure 5.1.) is a complex system combining brightfield and fluorescent microscopy, high speed, broad-field optical imaging, high precision laser targeting, laser mediated cell-ablation, -catapulting, and -manipulation, a stepping motor system with microstepping capability, and multithreading computer technology. The instrument and its operating principals are described in detail in Chapter 2. Briefly, the system includes three interacting computers running several custom software packages, three monitors, an optical imaging and targeting path with a broad band illuminating lamp, several filter wheels, lenses, mirrors and prisms, two CCD cameras, a high energy, pulsed laser providing high peaks of power output in the UV, visible, and IR ranges of the spectrum, two high-speed galvanometer mirrors alternatively targeting the laser beam to the sample and the reflected light from the sample to the cameras, 16 stepping motors accurately moving and positioning the sample holding platform, the filter wheels, and the galvanometer mirrors. Figure 5.2. demonstrates the general principal of LEAP in a simplified model. The model demonstrates that the sample is placed on a flat platform and illuminated first by a broad-band light source to select the target cells (Figure 5.2.A). Selected cells are then targeted one at a time with the laser beam (Figure 5.2.B). Highly accurate targeting is achieved with the fast moving galvanometer mirrors. Targeted cells can be eliminated by ablation or by laser catapulting. Using lower laser energy settings the target cells can be induced to go into apoptosis instead of being ablated (90). Shooting with even lower laser energy can result in optoinjection of the targeted cells by macromolecules added to the medium prior to LEAP (72,88,89). These unique capabilities LEAP offers could be essential for consecutive microarray experiments.

Figure 5.1. The LEAP instrument in our laboratory

A: The LEAP instrument was placed on an air-table to reduce the effects of the vibrations of the laboratory. B: With the side panels off the intricate maze of CCD cameras, filter wheels, lenses, mirrors and prisms are visible. The laser is in the bottom left corner in the picture. C: The laser in action.

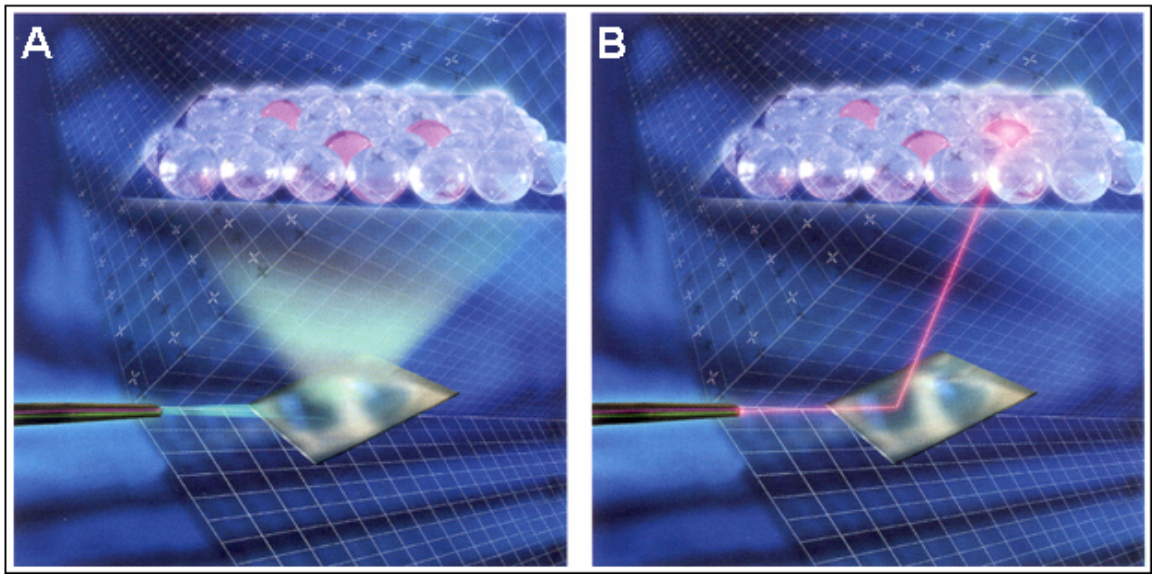

Figure 5.2. The general principal of LEAP

A: the sample is placed on a flat platform and illuminated first by a broad-band light source to select the target cells. B: Selected cells are targeted one at a time with the laser beam. (Source: www.cyntellect.com)

One objective of the LEAP experiments in our laboratory was to test the cell purification capabilities of this new approach in purifying adherent cell types. A major advantage of LEAP is the potential ability to purify cells from a cell monolayer by eliminating unwanted cells without dissociating the purified cell subset from the surface of the tissue culture vessel it was cultured in. This could be very important in studying sensitive cell types like isolated primary hepatocytes that do not tolerate other cell purification methods well. We modeled adherent cell purification experiments using HeLa cells first then using the experience we gained with these model experiments we purified mouse primary hepatocytes by LEAP.

Another objective was to test the applicability of LEAP for separation of stem/progenitor blood cell subsets from cord blood for gene expression profile (GEP) microarray analysis. In model cell mixture experiments we used KG-1a cells labeled with fluorescent anti-CD34 antibodies to mimic CD34 antigen expression based cell sorting of human stem/progenitor blood cells (118-120). The KG-1a cell line was established from bone marrow tumor and these cells all express CD34 antigen at a constitutively high level (ATCC CCL-246.1).

A third objective was to test CD4 surface antigen based cell purification of human immunodeficiency virus (HIV) susceptible cells and HIV infected cells to study the changes in the GEPs of these cells after HIV infection. Since the CD4 surface antigen is the main cell surface receptor used by HIV (121,122) we used CEM cells in our model cell purification experiments. CEM is a lymphoblastoid T-cell line that constitutively expresses CD4 antigen. We designed a set of experiments to separate CEM cells based on their expression level of CD4 antigen on the sample cells' surface. The capability to sort live, HIV-infected cell samples in a closed system could enable us to use microarray analysis to study gene expression profile alterations caused by HIV infection in the LEAP-purified cell subsets.

Finally, we tested and characterized the optoinjection capabilities of LEAP in both adherent and suspension cell cultures using different size fluorescent dextran molecules. Selectively delivering macromolecules into targeted cells like primary

hepatocytes, stem/progenitor blood cells, or HIV infected cells without significantly damaging them would offer new possibilities in studying and manipulating them.

RESULTS

Ablation and Purification of Live Adherent Cells

Figures 5.3 and 5.4 demonstrate the results of a set of experiments designed to explore the power and accuracy of the LEAP platform for cell purification from a monolayer of adherent cells. A model cell culture of HeLa cells was plated in a semi-confluent monolayer and processed live by LEAP. Defined, easily recognizable geometrical regions of the cell monolayer were LEAP-ablated using the grid-pattern shooting method. Figure 5.3. displays the successful ablation of a square region. To assess LEAP-related effects, targeted regions, their immediate surrounding areas, and the rest of the culture were monitored for ablation/cell damage and compared to control cultures. After optimizing the shooting conditions, virtually 100% of the targeted cells were completely ablated with only a very few unattached, dead/damaged cells found in the targeted areas. The targeted regions could be very accurately ablated with sharp edges and angles circumscribing the ablated area. Indeed, targeted regions could also be ablated at any desired configuration. Figure 5.4. displays three microscopic views of the same slide chamber where the letters M, C, and U (abbreviation of Molecular Cytometry Unit, our laboratory) were ablated. A few damaged cells were observed within a 1-2 cell diameter range (5-20μm) from the targeted area (Figure 5.3. Region 1). The rest of the cells throughout the culture did not display any signs of damage. Most of the cells bordering the ablated region (Figure 5.3. Region 2) and the cells plated towards the periphery of the slide chamber (Panel 1, Region 3) appeared to be unaffected. For this ablation application, high laser power (50-100%) with very few (1-3) repeats was found to be the most efficient.

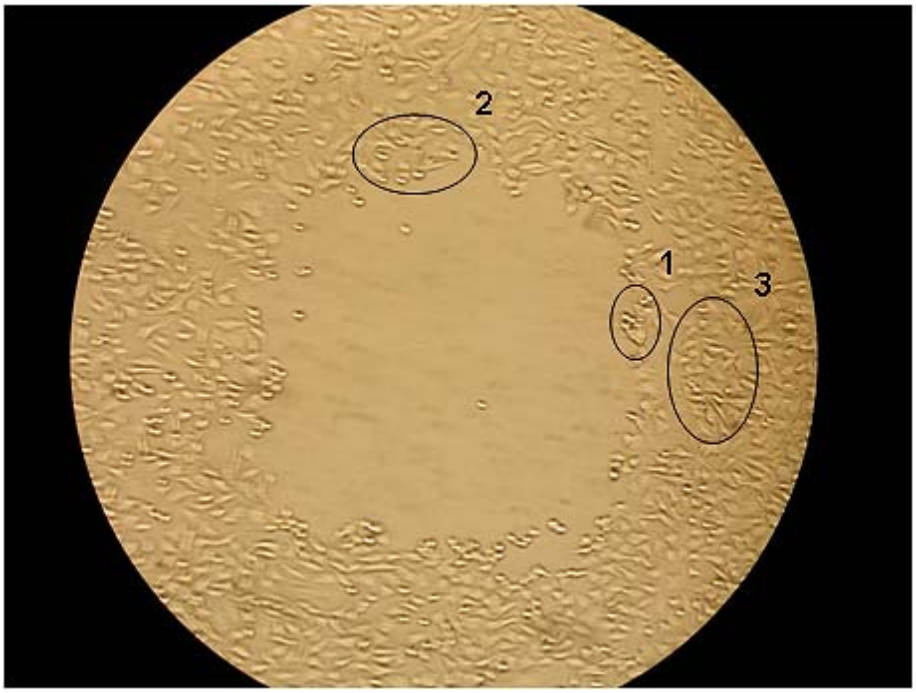

Figure 5.3. LEAP-mediated ablation of a square region from a cell monolayer

HeLa cells were plated in a semi-confluent monolayer and processed live by LEAP. A square region of the cell monolayer was LEAP-ablated using the grid-pattern shooting method with high laser power (100%) and 2 repeats. Brightfield view of the targeted area after the LEAP-ablation is presented.
Region 1: A few damaged cells observed within a 1-2 cell diameter range (5-20μm) from the targeted area. Region 2: unaffected cells bordering the ablated region. Region 3: Unaffected cells towards the periphery of the slide chamber.

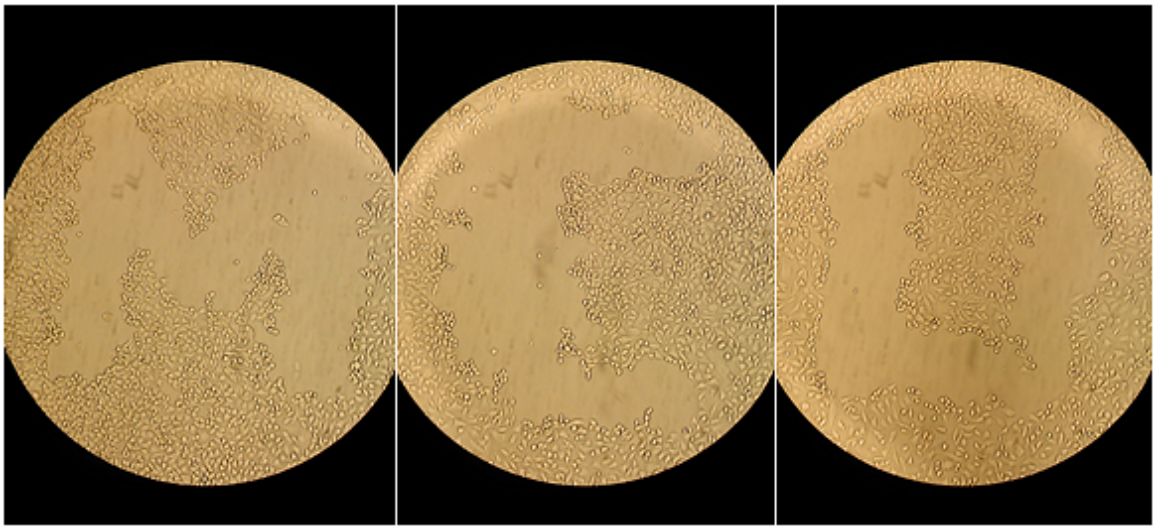

Figure 5.4. LEAP-mediated ablation of complex regions from a cell monolayer

HeLa cells were plated in a semi-confluent monolayer and processed live by LEAP. Different regions of the cell monolayer in the shape of the letters M, C, and U were LEAP-ablated to characterize the accuracy of the method and the cell damage it caused. Grid-pattern shooting method with high laser power (100%) and 2 repeats was used. Brightfield views of the targeted areas after the LEAP-ablation.

After establishing the basic characteristics of LEAP-mediated cell ablation, we modeled the more lifelike problem of purifying a cell culture from contaminating individual cells. Purification of a confluent/semi-confluent monolayer of unlabeled HeLa cells from a small subset of contaminating labeled HeLa cells (below 5%) could be achieved by targeting these unwanted cells individually, using the shooting conditions described. The ablating/damaging LEAP-effect on the non-targeted neighboring cells caused a 10-20% loss in the purified unlabeled HeLa cell population. This approach resulted in virtually 100% purity with above 80% yield. To purify a cell mixture with above 5% unwanted cell ratio, lower cell culture density was needed to avoid significant cell loss by LEAP-purification. The optimal plating density had to be determined empirically based on the contaminating cell ratio and the fragility/sensitivity of the given cell type to LEAP-irradiation. Figure 5.5. demonstrates an example of a highly LEAP-sensitive, adherent cell type with 50% contaminating cell ratio. In this experiment, freshly isolated hepatocytes were infected with an Adenovirus expressing the Green Fluorescent Protein (GFP) reporter gene at approximately 50% infection efficiency. The GFP expressing cells were subsequently LEAP-purified for further experiments. These conditions required low-density cell plating of single cells with the cells being 2-5 cell diameters (30-100μm) apart. The culture was briefly treated with CellTracker Orange (CTO) to label all cells for targeting. (90) For visualization in the instrument, which produces grey-level images on each of its two cameras, the fluorescence is displayed in pseudo-colors generated by the LEAP imaging system (GFP=green and CTO=orange).

Figure 5.5.A presents a simple two-color fluorescent image of hepatocytes taken by LEAP before laser processing. GFP-expressing cells appear green to yellow depending on the relative strength of their green and orange fluorescence, while GFP-negative cells appear orange. We targeted these negative cells using the laser and another two-color image was created of the same field-of-view immediately after the shooting was performed (Figure 5.5.B). Hepatocytes have the tendency to form multicellular conglomerates (Figure 5.5. A and B, Regions 1) that may contain positive and negative cells at the same time. These cell clusters were not targeted because LEAP-processing

was found to have an all-or-nothing effect on them where targeting any one of these cells detached the entire cluster. The rest of the GFP negative cells were all ablated (Figure 5.5. A and B, Regions 2, 3) or seriously damaged (Region 4). For this sensitive cell type, maintaining attachment to the plate surface is crucial. Detached hepatocytes are much more likely to die by apoptosis than to re-attach. To monitor the exact position of each individual cell before and after LEAP-shooting we superimposed the before and after images of each view. Figure 5.6.A was created by overlaying the green fluorescent images of the same view before (green) and after (black) the shooting. Only GFP-positive cells are visible on this image since the negative cells do not emit light at this wavelength (525nm). Black spots on top of green spots represent GFP-positive cells that have not been moved by LEAP processing (Figure 5.6.A, Region 1). An exposed green spot represents a GFP positive cell that has moved away from that position during the shooting while a black spot alone represents the new position of a moved GFP-positive cell (Figure 5.6.A, Region 2). The vast majority of GFP-positive cells were found to be unaffected. Similarly, Figure 5.6.B was created by overlaying the orange fluorescent images of the same view before (orange) and after (black) LEAP processing. Since all cells are CTO-positive, both GFP-positive and -negative cells are visible in this image. GFP-negative cells are visible in this image but not in the previous one (Figure 5.6.B, Regions 2, 3, 4). Orange-only spots represent removed cells (Regions 2, 3), while small gray spots appeared on top of larger orange spots when the shooting resulted in seriously damaged cells or cell debris (Region 4). Taken together, Figure 5.6.A and B revealed that most of the GFP-negative cells had been removed by LEAP purification while most GFP-positive cells remained intact. This purification method resulted in above 90% cell recovery with approximately 90% purity. The remaining GFP-positive cells suffered no apparent damage and could be cultured further. For this application, medium-level laser power (25-75%) with several (5-15) repeats proved to be the most efficient.

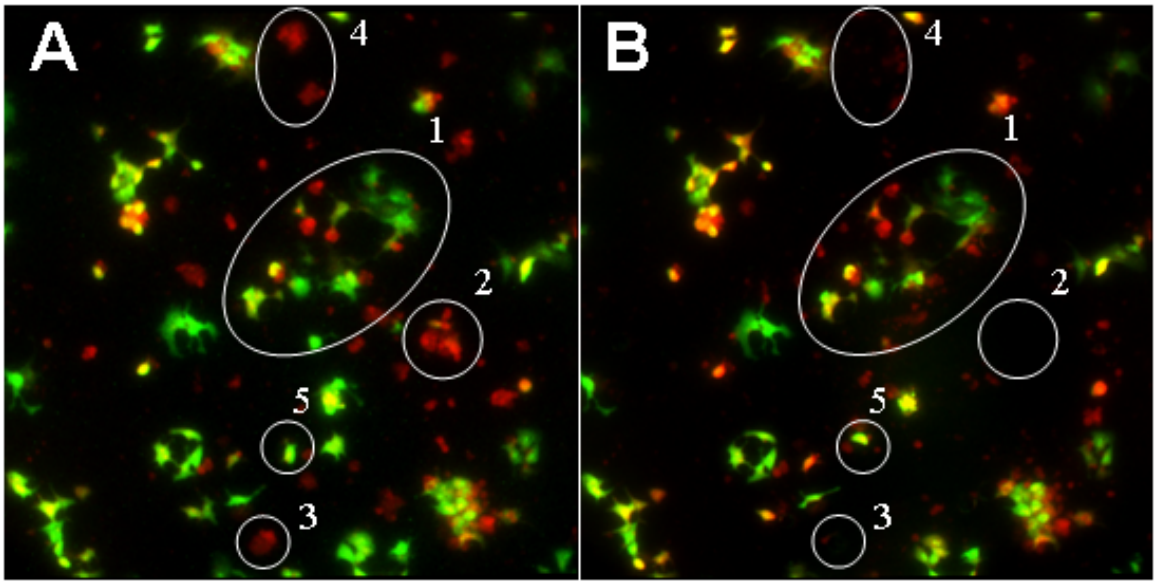

Figure 5.5. LEAP-mediated purification of individual GFP-expressing hepatocytes - Two-color fluorescent images

Green Fluorescent Protein expressing (GFP-positive) and GFP-negative hepatocytes were plated at low density. The culture was briefly treated with CellTracker Orange (CTO) to label all cells for targeting. GFP-negative cells were targeted by LEAP using 50% laser power and 10 repeats at individual shooting mode. GFP and CTO images of the same view were overlaid to visualize GFP-negative cells (orange) and GFP-expressing cells with GFP-positive (green-yellow) and negative (orange) portions before (A) and after (B) LEAP purification.
Region 1: Unaffected multicellular conglomerate with both positive and negative cells. Regions 2, 3, and 4: Ablated and / or seriously damaged GFP-negative cells. Region 5: Detached GFP-positive cell.

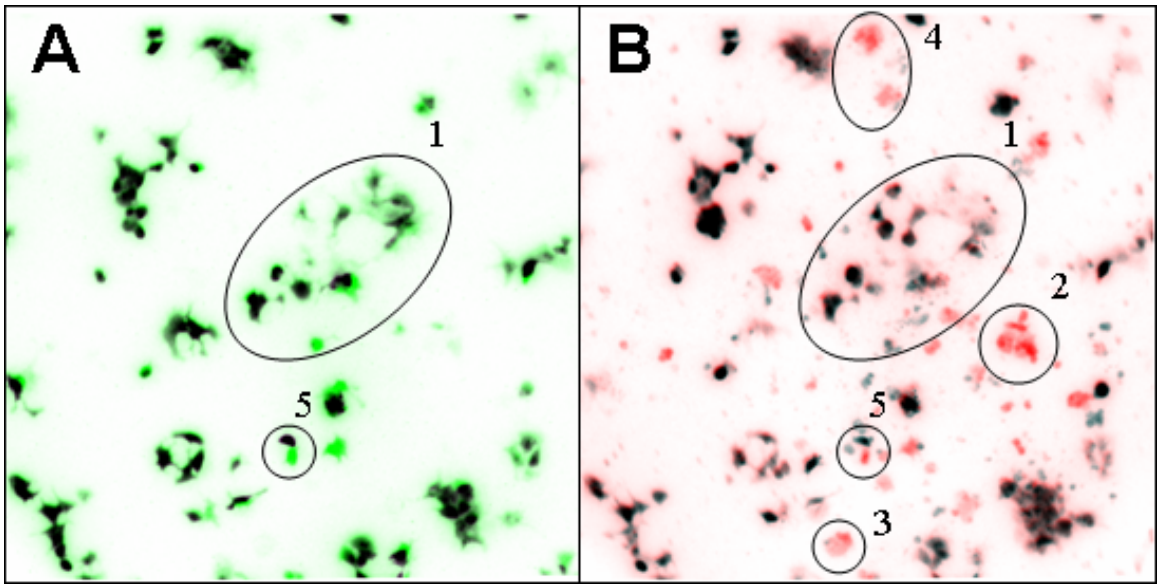

Figure 5.6. LEAP-mediated purification of individual GFP-expressing hepatocytes - Single-wavelength fluorescent images

Green Fluorescent Protein expressing (GFP-positive) and GFP-negative hepatocytes were plated at low density. The culture was briefly treated with CellTracker Orange (CTO) to label all cells for targeting. GFP-negative cells were targeted by LEAP using 50% laser power and 10 repeats at individual shooting mode. Panel A: To visualize detached GFP-positive cells green fluorescent images of the same view before (green) and after (black) LEAP purification were overlaid. Exposed green spots represent detached GFP-positive cells while black spots on top of green spots represent unaffected GFP-positive cells. Panel B: To visualize all detached cells red fluorescent images of the same view before (orange) and after (black) purification were overlaid.
Region 1: Unaffected multicellular conglomerate with both positive and negative cells. Regions 2, 3, and 4: Ablated and / or seriously damaged GFP-negative cells. Region 5: Detached GFP-positive cell.

Purification of Live Suspension Cells

As a model cell mixture for suspension cells, a 50% mix of differentially labeled KG-1a cells (CD34-FITC label) and CEM cells (CD4-PE label) was prepared. The mixture was plated at different cell densities in 8-chambered slides and the slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. Figure 5.7. illustrates purification results after low-density plating with the cells being 5-10 cell diameters (30-100μm) apart. Figure 5.7.A is a two-color fluorescent image of the cells taken by the LEAP instrument prior to laser processing. KG-1a cells appear green; CEM cells are orange. We targeted orange (PE-positive) cells with the laser and a subsequent two-color image was created of the same view after the shooting (Figure 5.6.B). Most CEM cells disappeared (i.e. were ablated or detached and floated away) from the field-of-view (Figure 5.6.A and B, Regions 1 of both images). The rest of the CEM cells were moved from their original location (Regions 2, 3); many of these detached cells appeared obviously damaged. Figure 5.8.A was created by overlaying the green fluorescent images of the same view before (green) and after (black) the shooting of the cells. This image exhibits the KG-1a cells only, and confirms that the majority of them remained attached (Region 4) with only a few of them moved (Region 5). Figure 5.8.B was created by overlaying the orange fluorescent images of the same view before (orange) and after (black) LEAP processing. Only CEM cells were visible here confirming that virtually none of the targeted CEM cells remained attached (Figure 5.8.B, Regions 1, 2, 3). This purification method resulted in greater than 80% cell recovery with 95-100% purity. The recovered, purified KG-1a cells suffered no apparent damage and could be cultured further. For this application, medium-level laser power (25-75%) using a few (1-5) repeats was found to be optimal. Although LEAP purification at this plating density resulted in very high levels of purity it was only feasible with very small samples or when only a few purified cells are needed. To purify more cells at a reasonable speed, higher plating density was required.

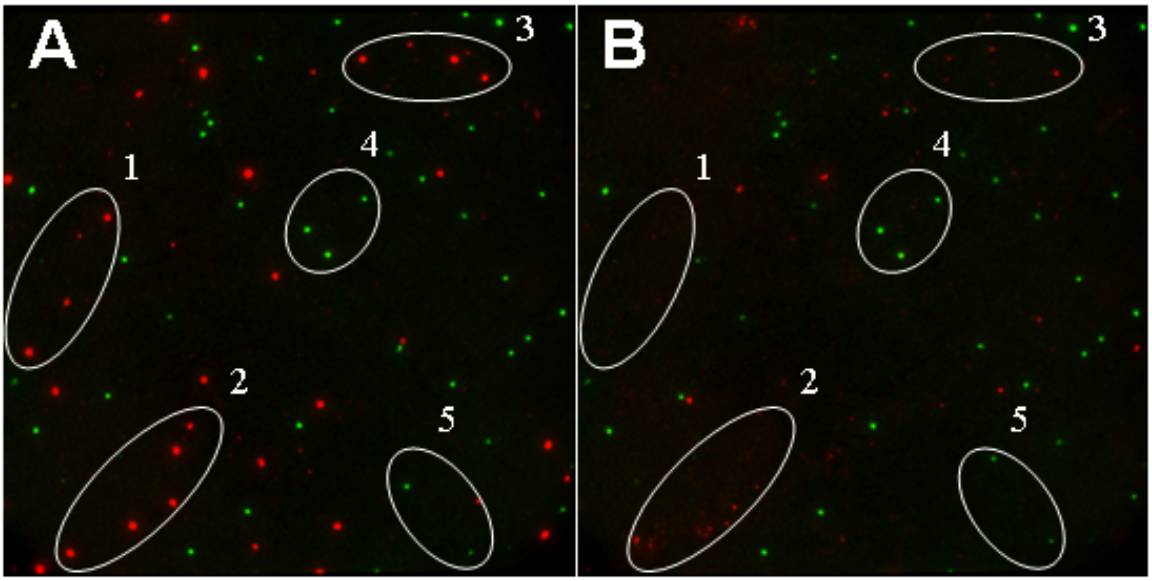

Figure 5.7. LEAP-mediated purification of suspension cells at low density - Two-color fluorescent images

A 50% mix of differentially labeled KG-1a cells (CD34-FITC label) and CEM cells (CD4-PE label) was plated in 8-chambered slides at low density. The slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. KG-1a cells (green) were purified by LEAP ablation/detachment of the CEM cells (orange) using 50% laser power and 3 repeats at individual shooting mode. FITC (green) and PE (orange) images of the same view were overlaid to visualize both cell types before (A) and after (B) LEAP purification.
Region 1: Ablated CEM cells. Regions 2 and 3: Detached and / or seriously damaged CEM cells. Region 4: Unaffected KG-1a cells. Region 5: Detached KG-1a cells.

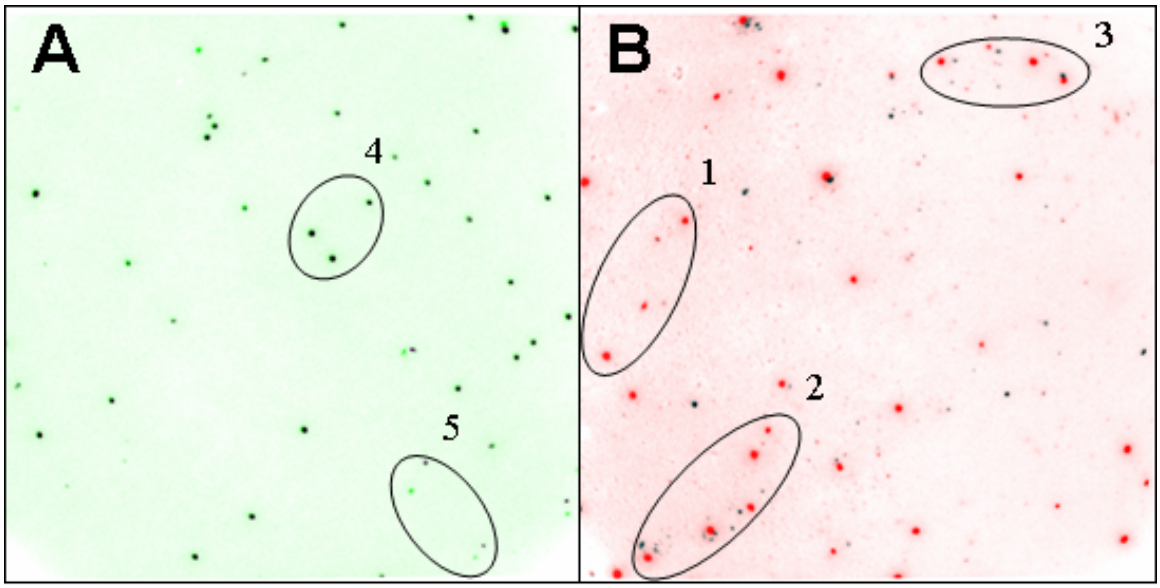

Figure 5.8. LEAP-mediated purification of suspension cells at low density - Single-wavelength fluorescent images

A 50% mix of differentially labeled KG-1a cells (CD34-FITC label) and CEM cells (CD4-PE label) was plated in 8-chambered slides at low density. The slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. KG-1a cells (green) were purified by LEAP ablation/detachment of the CEM cells (orange) using 50% laser power and 3 repeats at individual shooting mode. A: To visualize detached FITC-positive KG-1a cells green fluorescent images of the same view before (green) and after (black) LEAP purification were overlaid. Exposed green spots represent detached FITC-positive cells while black spots on top of green spots represent unaffected FITC-positive cells. B: To visualize detached PE-positive CEM cells red fluorescent images of the same view before (orange) and after (black) purification were overlaid. Region 1: Ablated CEM cells. Regions 2 and 3: Detached and / or seriously damaged CEM cells. Region 4: Unaffected KG-1a cells. Region 5: Detached KG-1a cells.

Figure 5.9. exhibits purification results from the same 50% KG-1a/CEM cell mixture after higher-density cell seeding when the cells were 1-5 cell diameters (10-50μm) apart. Figure 5.9.A is a two-color fluorescent image of the cells taken by the LEAP instrument before laser processing. In the mixture of KG-1a (green) and CEM cells (orange), the LEAP software was used to target and eliminate CEM cells. Another two-color image was created of the same field-of-view after purification (Figure 5.9.B). Since this plating density only allowed lower power LEAP shooting in order to preserve non-targeted KG-1a cells, significantly fewer targeted cells had completely disappeared from the view after shooting than with low density plating. Some targeted CEM cells could be observed to be floating above the focal plane of the instrument's CCD camera (Figure 5.9.B, Region 1). Figure 5.10.A was created by the same overlay technique used in the earlier images, where green areas and black spots represent FITC-labeled cells before and after laser shooting, respectively. Only KG-1a cells were visible and there was visual confirmation that a good portion of them remained attached (Figure 5.10.A, Region 2). It should be noted that a significant number of KG-1a cells were found to have moved away from their origins by the indirect effects of the LEAP shooting (Region 3). Figure 5.10.B was created by overlaying the orange fluorescent images of the same view before (orange) and after (black) LEAP processing. This combined image displays CEM cells only and confirms that very few of the targeted CEM cells remained attached (Figure 5.10.B, Region 4). The majority of the targeted cells had detached, but still remained visible (Region 5). This plating density resulted in greater than 50% cell recovery with more than 90% purity. While the purity remained fairly high, the cell recovery rate fell significantly in trade for the ability to process more cells in less time. The recovered, purified KG-1a cells did not suffer apparent damage and could be cultured further. For this application low-level laser power (25-50%) was applied in several (3-10) repeats. The results of the different types of LEAP-purification experiments and the laser power applied are summarized in Table 1.

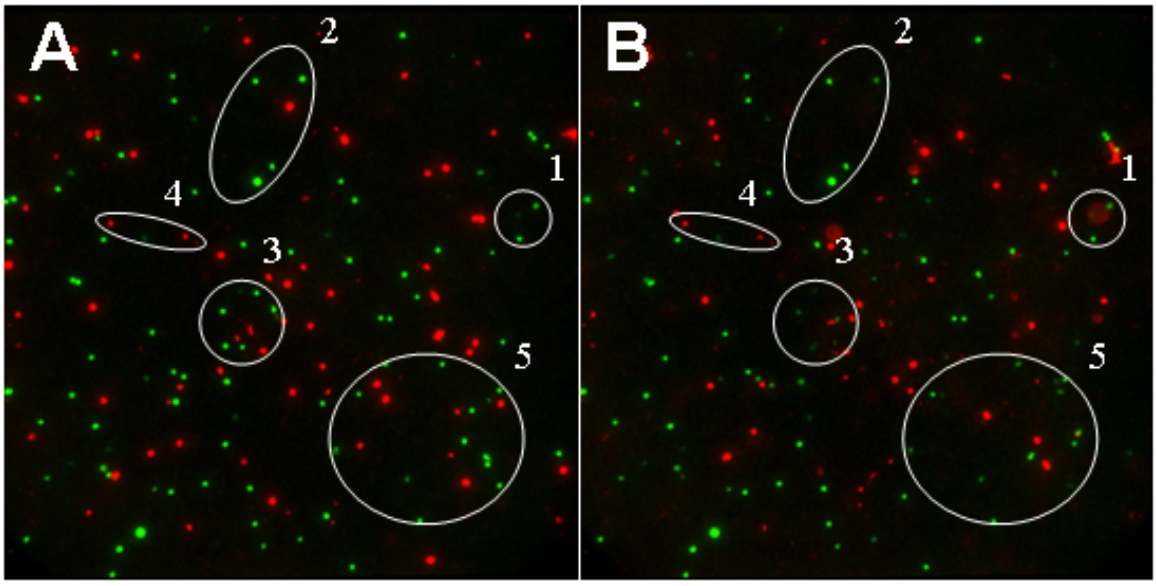

Figure 5.9. LEAP-mediated purification of suspension cells at high density - Two-color fluorescent images

A 50% mix of differentially labeled KG-1a cells (CD34-FITC label) and CEM cells (CD4-PE label) was plated in 8-chambered slides at high density. The slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. KG-1a cells (green) were purified by LEAP ablation/detachment of the CEM cells (orange) using 30% laser power and 5 repeats at individual shooting mode. FITC (green) and PE (orange) images of the same view were overlaid to visualize both cell types before (A) and after (B) LEAP purification.
Regions 1 and 2: Unaffected KG-1a cells with a floating CEM cell in region 1 after the purification. Regions 2 and 3: Detached and / or seriously damaged CEM cells. Region 4: Unaffected CEM cells. Region 5: Detached CEM and KG-1a cells.

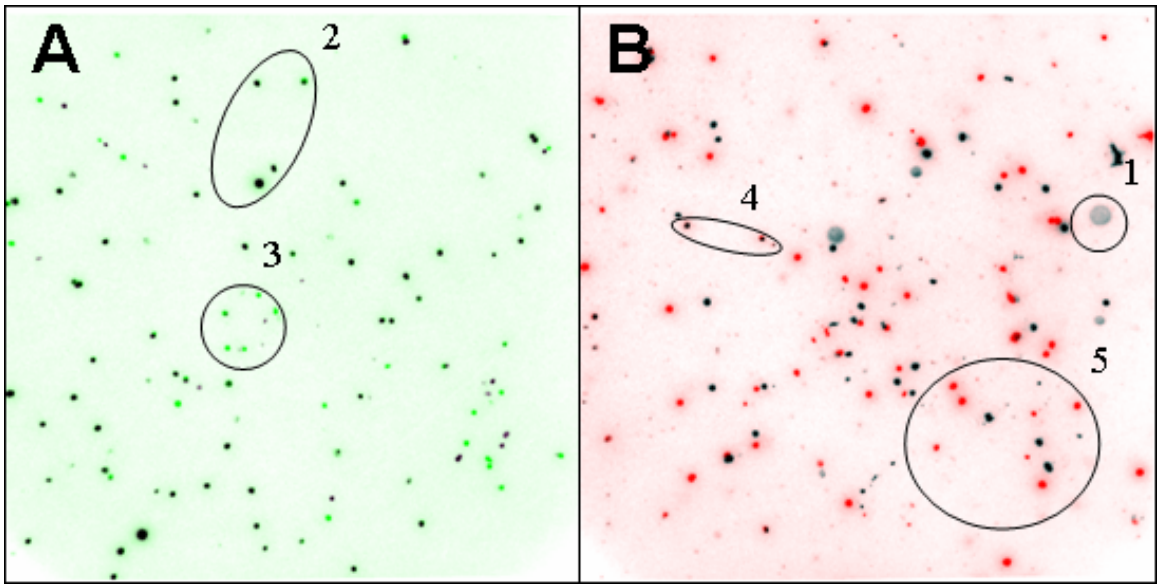

Figure 5.10. LEAP-mediated purification of suspension cells at high density - Single-wavelength fluorescent images

A 50% mix of differentially labeled KG-1a cells (CD34-FITC label) and CEM cells (CD4-PE label) was plated in 8-chambered slides at high density. The slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. KG-1a cells (green) were purified by LEAP ablation/detachment of the CEM cells (orange) using 30% laser power and 5 repeats at individual shooting mode. A: To visualize detached FITC-positive KG-1a cells green fluorescent images of the same view before (green) and after (black) LEAP purification were overlaid. Exposed green spots represent detached FITC-positive cells while black spots on top of green spots represent unaffected FITC-positive cells. B: To visualize detached PE-positive CEM cells red fluorescent images of the same view before (orange) and after (black) purification were overlaid.
Region 1: A floating CEM cell in after the purification. Region 2: Unaffected KG-1a cells. Region 3: Detached KG-1a cells. Region 4: Unaffected CEM cells. Region 5: Detached CEM cells.

Cells	Purity	Yield	Damage	Laser Power
Adherent Cells - Confluent (Region) *	100%	90%	Some	50-100%
Adherent Cells - Confluent (Individual up to 5%) **	100%	80%	Some	50-100%
Adherent Cells - Low Density (Individual) **	90%	90%	Some	25-75%
Suspension Cells - High Density (Individual) **	90-95%	50-75%	None	25-50%
Suspension Cells - Low Density (Individual) **	95-100%	80%	None	25-75%

Table 5.1. LEAP-mediated purification of adherent and suspension cells

Purity, yield, and cell damage of the purified cell subpopulation as well as the laser power used for the purification are compared for each of the cell types and plating conditions tested.

* Ablation of defined regions from a confluent monolayer of cells.

**Purification of cell samples from individual contaminating cells.

Optoinjection of Live Adherent Cells

Optoinjection (delivering macromolecules into cells by laser irradiation) effects after shooting at cells by LEAP have been reported by Clark et al (88). When we added fluorescent dextran (TMR-conjugated dextran; MW=10kD) into the medium prior to the ablation experiments described above, we observed such optoinjection effects (Figures 5.11. and 5.12.) on non-targeted cells. The cells bordering the ablated regions turned fluorescent as they took up dextran molecules, while the cells seeded at other regions of the slide chamber remained dextran-negative. This unintentional (indirect) optoinjection effect was strongest on cells immediately bordering the ablated regions and gradually weakened towards the slide periphery. We observed an inverse correlation between the level of indirect optoinjection and the cells' distance from the targeted areas (Figure 5.12.B, Region 1)**.** We found indirect optoinjection affected cells in an approximately 8-12 cell diameter-wide (60-120µm) band surrounding the targeted zone**.** This distance depended on the laser energies used, which were different for each cell type.

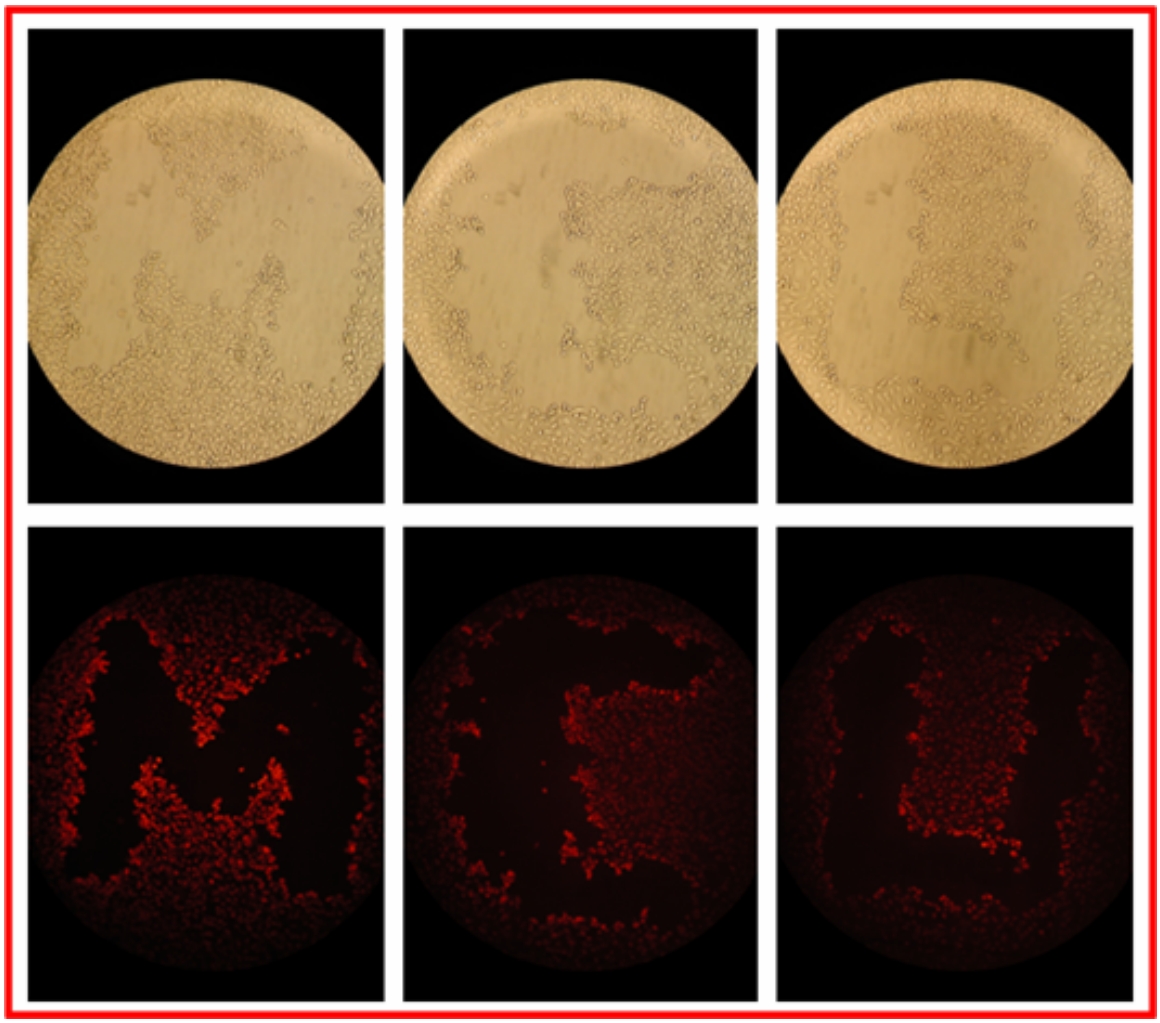

Figure 5.11. Indirect optoinjection during LEAP-mediated ablation from a cell monolayer

HeLa cells were plated in a semi-confluent monolayer and processed live by LEAP. Different regions of the cell monolayer in the shape of the letters M, C, and U were LEAP-ablated. Fluorescent dextran was added to the medium prior to LEAP ablation to visualize indirect optoinjection effects. Grid-pattern shooting method with high laser power (100%) and 2 repeats was used. Brightfield (top) and darkfield (bottom) images of the same views of the targeted areas after LEAP-ablation were compared.

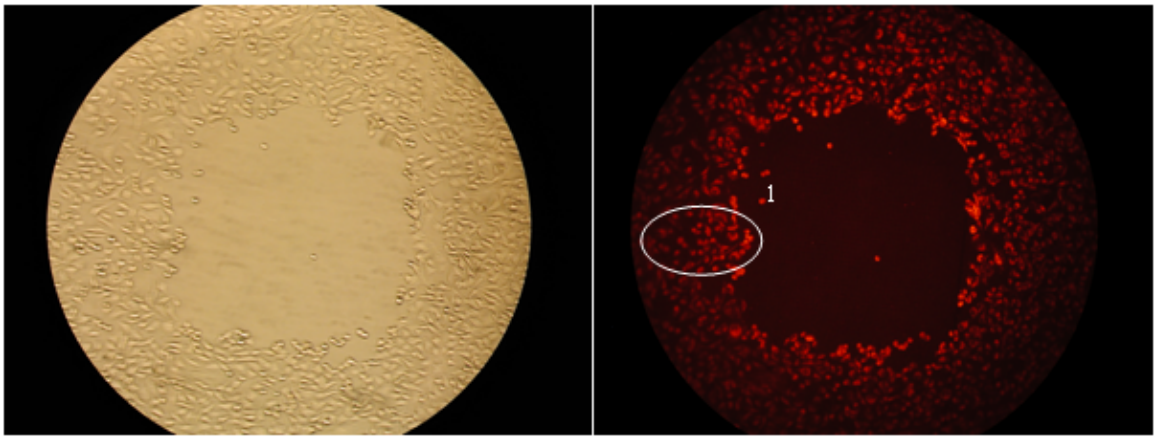

Figure 5.12. Range of indirect optoinjection effects during LEAP-mediated ablation from a cell monolayer

HeLa cells were plated in a semi-confluent monolayer and processed live by LEAP. A square region of the cell monolayer was LEAP-ablated. Fluorescent dextran was added to the medium prior to LEAP ablation to visualize the range of indirect optoinjection effects. Grid-pattern shooting method with high laser power (100%) and 2 repeats was used. Brightfield (left) and darkfield (right) images of the same view after LEAP-ablation were compared. Region 1 displays optoinjected cells in a distance of up to 10-12 cell diameters from the ablated region.

To study direct LEAP-optoinjection of confluent/semi-confluent HeLa cells, we used the laser at 20-40% of its full power, as recommended by Clark et al. (88). We introduced fluorescent dextrans of different sizes as deliverable macromolecules, into the medium before LEAP-shooting to visualize optoinjection effects and to assess the size range of optoinjectable molecules. Figure 5.13. demonstrates one of these experiments where the culture medium of a confluent HeLa cell monolayer was loaded with 10kD TMR-conjugated dextran prior to LEAP. Figure 5.13.A displays a defined square region targeted by LEAP. Figure 5.13.B demonstrates that the targeted region did get optoinjected by the fluorescent dextran from the medium using the grid-pattern shooting method. We found that shooting a single shot-grid at 25-50% laser power, repeated 2-3 times with a 4-second delay time between the series, optoinjected 100% of the targeted cells with very little cell loss (Figure 5.13.B, Region 1). The cells were approximately evenly fluorescent within the targeted region. The indirect optoinjection effect was found to be much weaker than with full power shooting (used in ablation experiments) and was limited to about a 4-6-cell diameter (30-60μm) region around the targeted area.

At a higher magnification (40x) the optoinjected cells were found to be structurally intact with a bright, dextran-positive cytoplasm and a darker, but still dextran-positive nucleus suggesting that most of the optoinjected dextran remained in the cytoplasm and only a small portion of the dextran molecules penetrated the nuclear membrane. (Figure 5.14.) The optoinjected dextran remained inside the cells for several days; detectable fluorescence gradually disappeared after 2-3 days depending on the original brightness of the optoinjected cells. In parallel experiments, we compared the optoinjectability of different sized TMR-conjugated dextrans using the same mass/volume final concentration of 3kD, 10kD, 40kD, and 70kD dextran molecules. We observed that using dextrans up to 40kD in size approximately the same level of fluorescence could be achieved. Optoinjecting HeLa cells with 70kD dextran still resulted in detectable fluorescence but its level was significantly lower than with smaller dextrans.

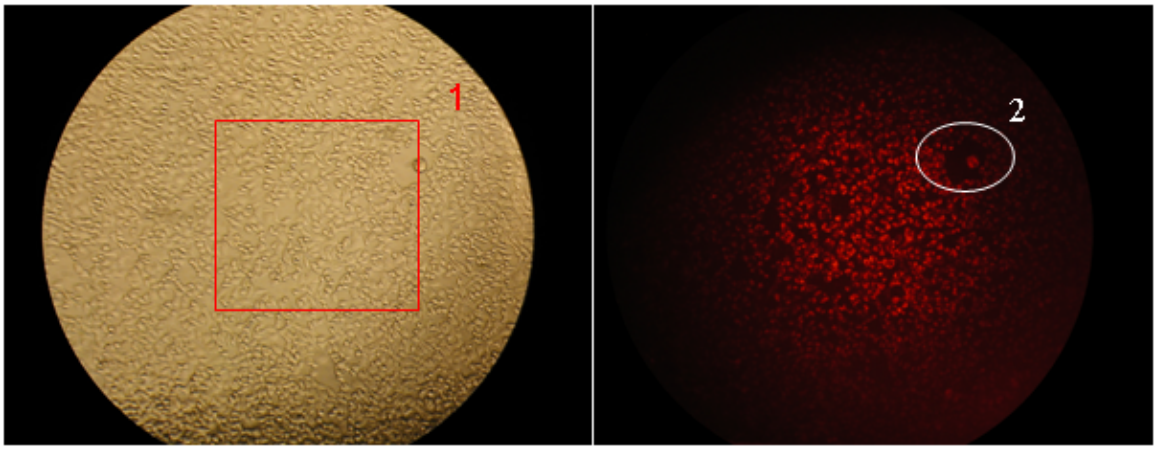

Figure 5.13. Direct, LEAP-mediated optoinjection of a cell monolayer

HeLa cells were plated in a semi-confluent monolayer and processed live by LEAP. A square region of the cell monolayer was LEAP-processed using grid-pattern shooting method with low laser power (25%) and 2 repeats 4 seconds apart. Tetra methyl rhodamine (TMR) conjugated, MW=10kD sized, fluorescent dextran was added to the medium prior to LEAP processing to visualize direct optoinjection effects. Brightfield (left) and darkfield (right) images of the same view after optoinjection were compared. Region 1 displays the targeted area. Region 2 displays a small ablated region with a few damaged cells.

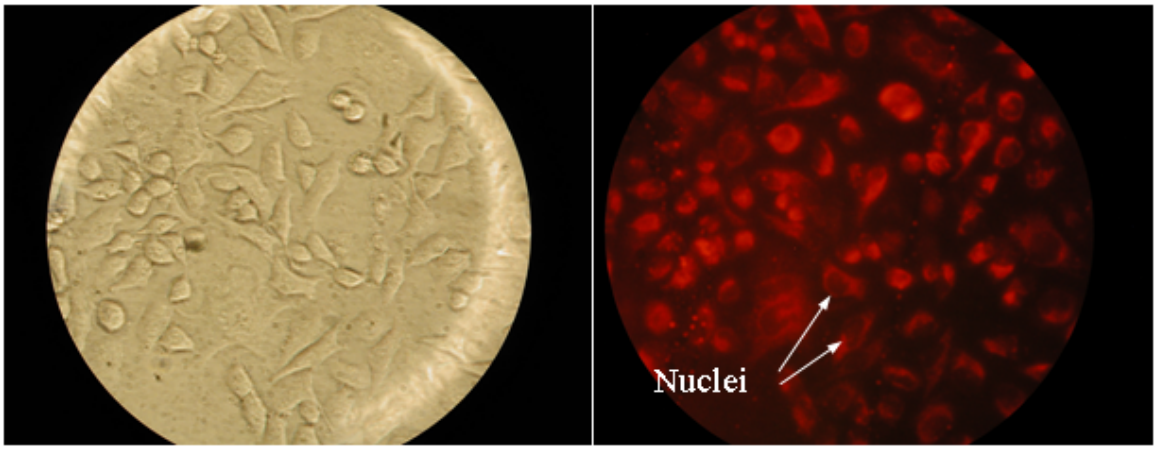

Figure 5.14. Distribution of fluorescent dextran within LEAP-mediated optoinjected HeLa cells

HeLa cells were plated in a semi-confluent monolayer and processed live by LEAP. A square region of the cell monolayer was LEAP-processed using grid-pattern shooting method with low laser power (25%) and 2 repeats 4 seconds apart. Tetra methyl rhodamine (TMR) conjugated, MW=10kD sized, fluorescent dextran was added to the medium prior to LEAP processing to visualize direct optoinjection effects. Brightfield (left) and darkfield (right) images of the same view after optoinjection were compared. To visualize the distribution of fluorescent dextran within optoinjected cells 40x magnification was used.

Optoinjection of Live Suspension cells

Figure 5.15. illustrates an experiment where a square region of densely plated KG-1a cells was optoinjected with 10kD fluorescent dextran to study the effectiveness of LEAP optoinjection on temporarily semi-attached suspension cells. The surrounding cells that were not targeted by the laser beam served as negative controls. The brightfield and darkfield confocal microscopy (10x magnification, 543nm excitation/625nm emission) results display a portion of the targeted square region (bottom, right, square area of this view) and the surrounding, non-targeted cells after the shooting. Using optimized shooting conditions (single shot-grid shooting at 20% laser power; repeated twice with a 4-second delay time between the series), 100% of the targeted cells became optoinjected. The overall level of fluorescence in the targeted region was noticeably lower compared to HeLa cells. Indirect optoinjection effects were also visually weaker and less extensive than with HeLa cells; only a 1-4 cell diameter (5-30μm) wide, faintly optoinjected region surrounded the targeted area. The edges of the fluorescent targeted area were much sharper than with adherent cell optoinjection.

Confocal microscopy results at higher (63x) magnification (Figure 5.15.) demonstrated that every targeted cell was successfully optoinjected and that the level of optoinjection varied from KG-1a cell to KG-1a cell in a much wider range (Figure 5.16.) than in the previous experiment with HeLa cells (Figure 5.14.) To confirm that the fluorescent dextran molecules were indeed inside the targeted cells after LEAP-optoinjection, a series of confocal images of the same view at consecutive vertical planes were taken (Figure 5.17.). Examining the optoinjected cells at a single vertical plane at the mid-section of the cells, we found that most of the optoinjected dextran molecules resided in the cells' cytoplasm with the nuclei also containing some, but much less dextran (Figure 5.18.). Cells in both the targeted region and the surrounding areas were found to be morphologically intact after LEAP-optoinjection.

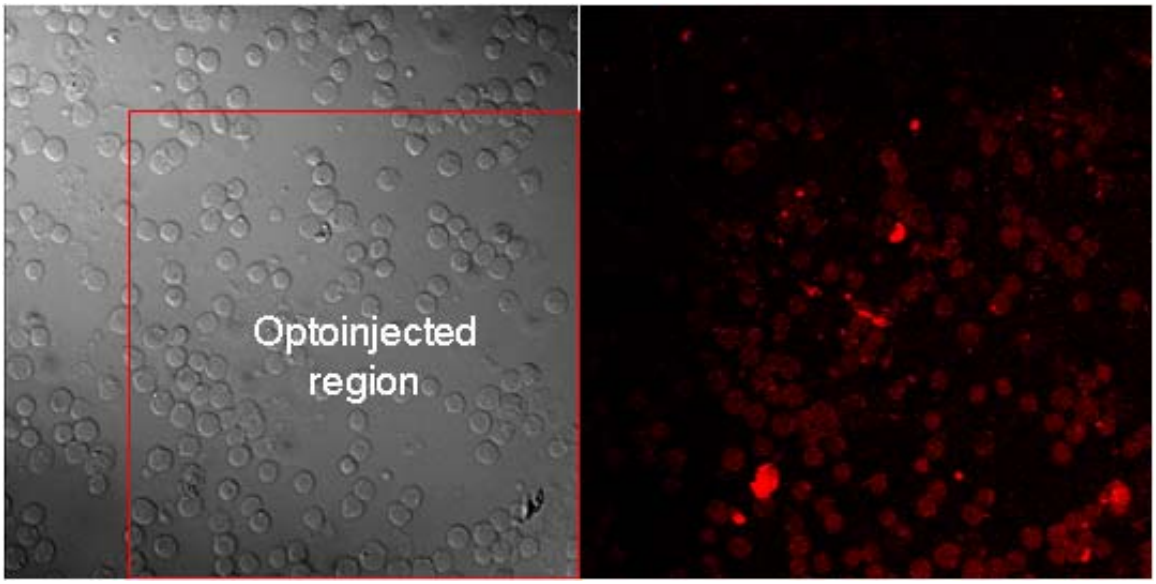

Figure 5.15. Direct, LEAP-mediated optoinjection of KG-1a cells

Live, unlabeled KG-1a cells were plated in 8-chambered slides at high density. The slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. A square region of the slide was LEAP-processed using grid-pattern shooting method with low laser power (20%) and 2 repeats 4 seconds apart. Tetra methyl rhodamine (TMR) conjugated, MW=10kD sized, fluorescent dextran was added to the medium prior to LEAP processing to visualize direct optoinjection effects. Brightfield (left) and darkfield (right) confocal images of the same view after optoinjection was compared at 10x magnification.

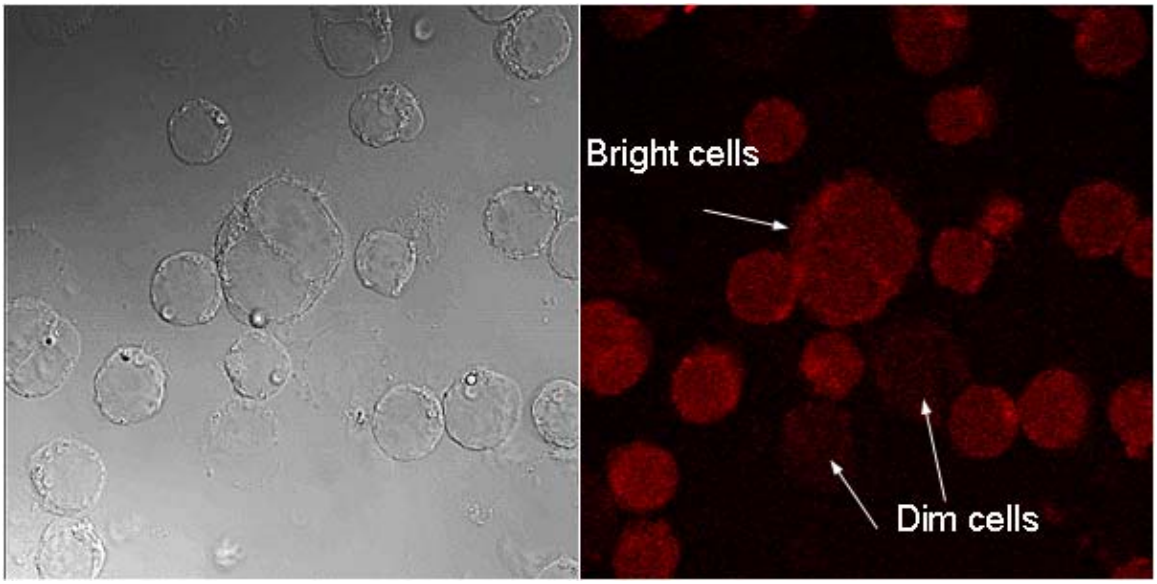

Figure 5.16. Differences in dextran uptake among LEAP-optoinjected KG-1a cells

Live, unlabeled KG-1a cells were plated in 8-chambered slides at high density. The slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. A square region of the slide was LEAP-processed using grid-pattern shooting method with low laser power (20%) and 2 repeats 4 seconds apart. Tetra methyl rhodamine (TMR) conjugated, MW=10kD sized, fluorescent dextran was added to the medium prior to LEAP processing to visualize direct optoinjection effects. Brightfield (left) and darkfield (right) confocal images of the same portion of the targeted region after optoinjection was compared at 63x magnification to visualize differences in brightness among optoinjected cells.

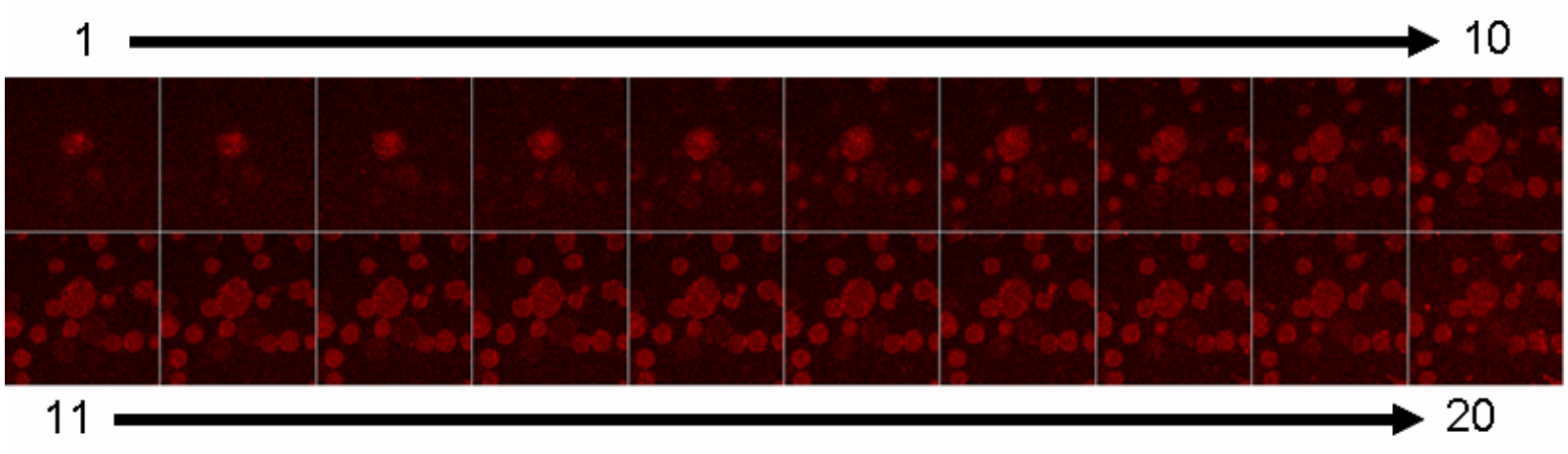

Figure 5.17. Serial confocal images of optoinjected KG-1a cells

Live, unlabeled KG-1a cells were plated in 8-chambered slides at high density. The slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. A square region of the slide was LEAP-processed using grid-pattern shooting method with low laser power (20%) and 2 repeats 4 seconds apart. Tetra methyl rhodamine (TMR) conjugated, MW=10kD sized, fluorescent dextran was added to the medium prior to LEAP processing to visualize direct optoinjection effects. A portion of the optoinjected area was inspected at 20 consecutive planes using confocal microscopy at 63x magnification to verify that the fluorescent dextran was located inside the cells after LEAP-optoinjection.

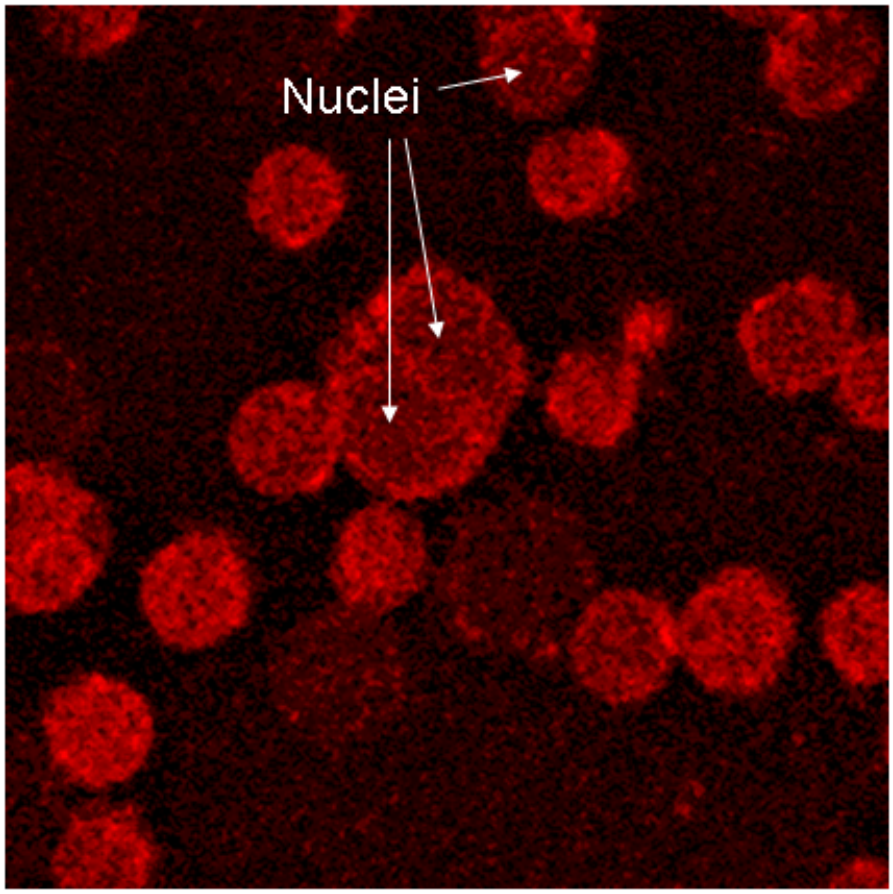

Figure 5.18. Distribution of optoinjected fluorescent dextran inside KG-1a cells

Live, unlabeled KG-1a cells were plated in 8-chambered slides at high density. The slides were gently centrifuged to settle all cells on the slide surface in a semi-attached state. A square region of the slide was LEAP-processed using grid-pattern shooting method with low laser power (20%) and 2 repeats 4 seconds apart. Tetra methyl rhodamine (TMR) conjugated, MW=10kD sized, fluorescent dextran was added to the medium prior to LEAP processing to visualize direct optoinjection effects. Confocal image of a portion of the optoinjected region at 63x magnification is presented to visualize differences in brightness between the cytoplasm and nucleus of optoinjected cells.

The main characteristics of both types of LEAP-optoinjection experiments are summarized in Table 2.

	Percent* Optoinjection	Delivery Efficacy**	Indirect*** Optoinjection	Visible Damage	Laser Power
Adherent Cells	100	High	30-60µm	None	25-50%
Suspension Cells	100	Low	5-30µm	None	10-20%

Table 5.2. LEAP-mediated optoinjection of adherent and suspension cells

Success rate and efficacy of optoinjection, visible damage of the optoinjected cells as well as the laser power used for the optoinjection and the width of the unintentionally optoinjected zone of cells are compared for each cell type tested.

* Percent Optoinjection: Percentage of optoinjected cells out of all the cells that were targeted
** Delivery Efficacy: Relative visual brightness of fluorescent dextran-optoinjected cells
*** Indirect Optoinjection: Width of the annular zone of cells unintentionally optoinjected

DISCUSSION

We have assessed the capabilities of LEAP, a new scanning cytometry technology in two fields of application: cell purification and targeted macromolecule delivery. In the purification studies, we only explored methods resulting in the immediate removal of unwanted cells. A more "patient" approach utilizing the reported apoptosis and necrosis inducing effect of LEAP irradiation (88) may lead to even better results in terms of the yield/purity ratio. However, the goal of this study was to characterize LEAP-based cell purification for immediate further cell processing (e.g. by microarray analysis).

We found that unwanted cells could be removed from attached cultures with almost "surgical" accuracy causing minimal damage to neighboring cells within 2-3 cell diameters. These initial studies only assessed any possible major morphological and/or structural damage to the processed cells; further analysis of LEAP purification effects on the finer structure and functions of the purified cells will be necessary.

A major strength of LEAP technology is its ability to deal with "problematic" cells where flow cytometry, laser capture microscopy (LCM), or magnetic bead sorting are not feasible options. Primary hepatocytes, as an example are extremely sensitive to manipulation with conventional techniques. Maintaining structurally and functionally

intact hepatocytes throughout the purification process is crucial for further analysis, culturing and experimentation (51). We demonstrated that LEAP provides the unique ability to purify these fragile cells without trypsinization and without causing any visible damage to them.

Processing live, suspension cells is always a challenge in scanning cytometry since the targeted cells need to be held in the focal plane of analysis and manipulation. To purify suspension cells by LEAP, we found that cells could be gently centrifuged to the slide/well surface and afterwards they remained in a semi-attached state with no additional attachment material needed and without any apparent damage to the cells. Another observation was that targeting settled suspension cells with LEAP, often resulted in these cells "bouncing" off the surface of the tissue culture well apparently unharmed rather than ablating these unwanted cells. These buoyant cells appeared in the recovered, purified cell population as contaminating cells. At the same time shooting with high laser energy required for ablation caused some non-targeted neighboring cells to detach as well. For these reasons, we changed our strategy and utilized this "bouncing effect" of LEAP shooting (requiring less laser power) instead of aiming for ablation. We centrifuged the cell mixture onto the slide surface without causing any apparent damage to the cells and bounced off the unwanted cells from this semi-attached state by LEAP shooting. We observed that many of the targeted cells still were ablated indicating that the amount of energy required for ablation of CEM cells varied in a wide range. After LEAP processing, we first removed the floating cells with a very gentle wash and then recovered the purified cells using a more aggressive wash. With this approach, we achieved much higher purity (above 90% in most applications) in the recovered cell population than when we were aiming for ablation of the contaminating cells alone. In many applications where the targeted cells are highly LEAP-irradiation sensitive, ablation might be the better approach for purification. In other cases – like in our suspension cell model – bouncing off contaminating cells using sub-lethal laser power may be more advantageous, especially since using less power allows for higher cell plating density that ultimately results in less time needed to purify a given number of

cells. This aspect may be important in studies where a large number of purified cells are required (e.g. microarray analysis).

We established the optimal cell plating densities needed for a given experiment based on the ratio of contaminating cells, the required end-purity, and the affordable cell loss. We demonstrated that when recovery of most purified cells is an important issue as from a small sample, it is possible to achieve above 95% purity and above 80% cell recovery with LEAP even from a 50% cell mixture. This requires low density cell plating that ultimately results in slower cell processing, but with small cell samples this is usually not an issue since the process still only takes a few minutes after the initial setup. When large cell samples are purified where 30-50% cell loss is acceptable, the cells can be plated denser significantly increasing the processing speed and still maintaining above 90% purity. We observed no apparent damage to the purified cells with either method.

We have also found that the combination of one round magnetic bead –pre-sorting from a large, 50% cell mix yielded a 75%-80% pure sample and a second round, high-density LEAP-purification from this sample resulted in 90% purity or better, similarly to 2 rounds of LEAP-purification or 2 rounds of magnetic bead sorting. The combined approach was much faster than handling a large sample with 2 rounds of LEAP purification alone and resulted in much less cell loss than 2 rounds of magnetic bead sorting alone. Since both methods can be carried out in a closed environment this combination method might be the best way to purify large, biohazard samples to high purities prior to microarray analysis.

We found an indirect optoinjection effect at the edges of the ablation zone in confluent adherent cultures that might be a sign of altered membrane and other functions in the purified cells. However, these effects are likely to be transient, lasting for a few seconds only since these cells did not display any signs of “leakiness” or morphological damage a few minutes after the indirect optoinjection as determined by confocal, brightfield, and fluorescent microscopy. The width of this indirectly optoinjected zone was found to depend on the cell type and the laser power used in the experiment. As

discussed earlier, further studies will be needed to elucidate the effects of laser mediated cell purification on different levels of the cell's cytome.

LEAP-mediated optoinjection is a novel tool for targeted macromolecule delivery (88,89). We studied optoinjection effects adding fluorescent dextrans of four different sizes as deliverable macromolecules into the medium before LEAP-shooting to visualize optoinjection effects. We found that optoinjection works with all cell types we studied (adherent and suspension cells) with no exception. It is even more promising that for every cell type we studied, it could be optimized to achieve literally 100% optoinjection of the targeted cells. When low laser power was used, we found no apparent morphological damage to any of the cell types. Optimized conditions, the concentration, size, and nature of macromolecules deliverable, and the optoinjection effects on cells close to the targets varied according to the different applications used.

We noted that adherent cells were easier to manipulate with LEAP since it did not require extra effort to keep them in the focal plane. The optoinjection effect appeared to be more diffuse on adherent cells than on suspension cells. Indirect optoinjection of a cell monolayer visibly affected cells up to 4-6 cell diameters away from the targeted zone. This indirect optoinjection effect as well as the amount of macromolecules delivered to the targeted cells (measured by the visible level of fluorescence) directly correlated with the laser power and the number of pulses applied. Higher laser power applied in fewer pulses and lower power applied in more pulses resulted in similar levels of fluorescence. The more sum energy we delivered to the cells the higher fluorescence level and the wider zone of indirect optoinjection we achieved – up to a certain point. Both of these effects were found to reach a plateau at a certain level, depending on the cell type. Delivering more energy above this plateau (by raising laser power or applying more pulses per cell) did not raise the level of fluorescence significantly, but caused visible cell damage. Based on these observations, we could optimize LEAP-shooting conditions for each cell type and application maximizing optoinjection effects without causing any visible damage to the targeted cells.

Gently centrifuged suspension cells could be optoinjected similarly to attached cells; the laser energy required for optoinjection did not remove the cells from the slide surface. The level of fluorescence was found to be higher even within the targeted region when we optoinjected adherent cells than with suspension cells. A possible reason could be that adherent cells are flat, all of them lying exactly in the focal plane while suspension cells even in their semi-attached state keep a certain vertical diameter depending on their size which causes them to stick out of the slide surface –thus the focal plane - unevenly. Alternatively, flat cell membrane may be more susceptible to optoinjection than spherical shaped. This could also explain the fact that we found much less indirect optoinjection effect only (1-4 cell diameter wide zone) with suspension cells than with adherent cells (4-6 cell diameter wide zone).

Confocal microscopy results confirmed that the fluorescent dextran molecules were indeed inside the cells after optoinjection. Most of the optoinjected molecules were found in the cells' cytoplasm with the nuclei remaining relatively dextran-negative, although, in some cases, visibly optoinjected. This analysis also confirmed that optoinjection did not cause any apparent alteration in cellular morphology.

The maximum size of optoinjectable macromolecules and the efficacy of optoinjection for macromolecules with different size and chemical structure still need to be determined. We recorded very similar results with 3kD-40kD dextrans and significantly reduced (~50%) fluorescence of the optoinjected cells when using 70kD dextran. These findings might mean that the limits of the underlying mechanism are not much above the size of a 70kD dextran molecule. The limiting size of optoinjected macro molecules remained to be determined.

Laser irradiation mediated cell membrane permeability changes and optoinjection effects have been reported by several studies, but the mechanism of optoinjection is unknown (88,123,124). It is not likely that the laser directly punches holes into the membrane of the targeted cell, because this theory could not explain the indirect optoinjection phenomenon we observed in this study. Shock waves created in the culture medium by laser shooting might contribute to the effects (125), but according to

calculations, they do not have enough energy to achieve macromolecule delivery through the cell membrane by themselves (88). The distance-dependent diffuse optoinjection phenomenon and the results with different laser energy/pulse number combinations suggest that the level of optoinjection depends on the sum energy communicated to the cells. Based on our results we propose the following possible mechanistic theory for the optoinjection phenomenon, which will require further testing beyond the scope of this dissertation.

During laser irradiation, laser energy gets absorbed by different molecules in the medium around the cells, in the cell membrane, and inside the cells warming up the cells in the targeted area. This heating effect of laser irradiation on cells and tissues is well-characterized and used in therapeutic applications (126-128). As a result of warming up beyond a certain threshold, the cell membrane goes through a phase change becoming more liquid-like than gel-like (129-131)**.** The uneven warming and the low energy shockwaves caused by the pulsing laser (125) generate waves in the fluid cell membrane. These waves may result in opening transient holes in the cell membrane and/or in transiently opening and enlarging existing pores and channels. If the sum laser energy is large enough to create membrane ruptures beyond the cells healing capabilities or even cause the cytoplasm to explode (increase its volume beyond the membranes flexibility) the end result will be ablation or permanent damage. If the sum laser energy is not enough for the above-described effects, but enough to cause transient membrane disturbances, the end result will be sum energy dependent optoinjection. If the sum laser energy absorbed stays below a certain threshold the cell membrane will remain intact and no optoinjection will occur. Further experiments will be needed to confirm our theory and to elucidate all underlying mechanisms for optoinjection.

In summary, LEAP offers a novel approach in analyzing, purifying and manipulating live cells. It is capable of purifying large or small, live cell samples in a closed system achieving 90% and higher purities. LEAP might become especially useful in areas where other sorting methods are seriously challenged as in purifying large numbers of live, adherent cells; or very small samples of live, suspension cells; or live

cells that are highly sensitive to traditional processing; or for safe, live processing of biohazardous cells. This capability might be especially useful in our current project. Optoinjection is an exciting capability LEAP offers. The 100% efficacy, highly accurate selectivity, and zero toxicity we observed are unmatched features by any other existing method for macromolecule delivery into cells.

Our beta prototype LEAP instrument did not have the capability of automated high speed cell processing; therefore we were not able to process large samples with it. The amount of RNA we could isolate from LEAP-purified and LEAP-optoinjected samples was not sufficient for direct microarray analysis. To analyze the GEP of these relatively small samples RNA amplification was necessary. These experiments are described in Chapter 6.

CHAPTER 6.
MICROGENOMICS

INTRODUCTION

The objective of the experiments described in this chapter was to test whether it was possible to reconstruct a meaningful GEP from small cell samples using existing RNA amplification technologies. We demonstrated that to analyze the GEP of a minor cell subset in a heterogeneous cell sample, purification of these cells is not only necessary but also very much achievable with carefully selected cell fixation and purification techniques. We showed that the originally hidden profile of these cells could be recovered without any significant distortion (63). After successfully purifying small cell samples using Laser Enabled Analysis and Processing we concluded that the purified cells did not yield enough RNA for direct microarray analysis (72). To analyze the GEP of such small cell samples RNA amplification was necessary. We tested the applicability of two commonly used RNA amplification methods (one of each major type) for consecutive microarray analysis. We studied their capacities and limitations in amplifying aRNA or cDNA from miniscule amounts of total RNA isolated from well defined numbers of cells even from single cells. We investigated the effects of these methods on the amplified GEP with special concern about any possible GEP distortion.

Microarray analysis requires microgram amounts of total RNA. This constraint narrows down the field of applicability for this powerful technology considerably, because without RNA amplification such amounts are obtainable only from millions of cells (44,45,92,93). Biological samples of this size are usually heterogeneous cell mixtures (with the exceptions of tumors and immortalized cell lines) and after the necessary cell purification (63) they often yield much fewer cells of the purified cell type than a million. In other cases like in small biopsies, fine needle aspirates, rare- or micromanipulated cells, the original tissue sample is already too small for direct microarray analysis (94-96). These challenges spawned the new field of microgenomics

that has already started to provide the means to analyze the GEP of these small biological samples with the goal to extend the benefits of microarray technology in defining sub-classes of diseases (other than tumors), predicting responses to different treatments and, ultimately, designing patient-tailored therapies (24-28). To be able to increase the amount of RNA we isolated from small samples to 10^3-10^6 times of its original size, as our first objective we tested selected RNA amplification methods for their productivity.

Single cell biology is a rapidly growing field of research (44,45,93,97-99). Most tissues are complex mixtures of heterogeneous cells, all of which respond to physiological, pathological, and experimental conditions in a unique fashion characteristic to the cell type. Many of these responses are reflected in the transcriptional activity of the given cell. In the context of the cellular diversity of tissues the ability to study the entire transcriptome of single cells using gene expression microarray technology could yield valuable insight into the biochemistry, physiology, and pathology of biological systems (97). For many diseases the GEP alterations in single cells may be more informative about the underlying pathomechanism than regional expression patterns of tissues (96). In an effort to apply gene expression microarray technology to small samples and single cells, several methods were developed to amplify picogram amounts of RNA available from these samples to microgram quantities required by microarrays (45,93,100-103). New methods for single cell identification, separation, and handling have been developed as well as RNA extraction and preservation techniques adapted to single cell biology (44,45,93,94,98,101,104-106). To this end, as our second objective we developed protocols to sort single cells, isolate RNA from them and amplify this few picograms of RNA 1-5x10^6 times.

Most RNA amplification techniques belong to one of two basic method types, they can be either exponential (also called logarithmic) or linear technologies (93,96). The principle of exponential RNA amplification is illustrated by Figure 1.6. First, the carefully isolated and preserved RNA undergoes a reverse transcription reaction primed by oligo-dT primer to amplify mRNA species only from total RNA. At the same time, this reaction is also used to label the cDNA with a universal tag sequence for the next

step. The resulting cDNA is amplified by PCR reaction using the universal tag sequence for priming. The resulting product is double stranded cDNA. The method is called exponential because the amplifying enzyme (usually Taq polymerase) uses the products of one round as templates in the next cycle resulting in exponential increase in the amount of produced cDNA. In contrast, linear amplification applies RNA polymerases (usually T7 RNA polymerase) as the amplifying enzyme which does not use the products of one round as templates in the next round directly. The principle of exponential RNA amplification is shown in Figure 1.7. The isolated RNA is also reverse transcribed by oligo-dT primer, but the additional sequence tag introduced by this method is a T7 promoter sequence. In the next step the T7 polymerase uses this sequence as starting point for transcription. The resulting product is antisense RNA (usually called aRNA or cRNA). Currently, the advantages and disadvantages of each RNA amplification technology, especially their effect on the amplified GEP, are not fully understood, but are of great importance for the field of microgenomics (63,93,96). As our third objective we analyzed the effects of selected linear and exponential methods on the GEP with special concern to any possible alterations they might introduce into the relative expression ratios of individual genes.

The model experiments described in our first three objectives enabled us to develop feasible methods to purify, fixate, and label selected cells from small samples, to isolate RNA from the few purified cells, amplify the isolated RNA, and analyze the amplified GEP. As a proof of principle experiment, our fourth objective was, to analyze the GEP of LEAP purified and optoinjected cells we discussed in chapter 5. This experiment expanded our evaluation of Laser Enabled Analysis and Processing technology since now we could investigate its effects at the transcriptome level of the processed cells.

RESULTS

Exponential RNA Amplification

We selected the same two cell lines (CEM and A2780) with characteristic similarities and differences in their GEPs that we used in the cell purification experiments. This enabled us to monitor the effects of the amplification process on the already characterized GEP of these cells. We used the commercially available SMART-PCR RNA amplification kit (Clontech, Palo Alto, CA) to investigate exponential RNA amplification. The method uses Taq-polymerase as its main amplifier enzyme as described in Chapters 1 and 2. We analyzed the GEP of amplified samples using Atlas Trial arrays (Clontech, Palo Alto, CA). Figure 6.1. displays the results of an RNA amplification experiment comparing the GEPs of 10^6 CEM cells, unamplified, 10^4 CEM cells, amplified by 18 rounds of PCR, and 10^3 CEM cells, amplified by 24 rounds of PCR to yield the required 1microgram cDNA. Approximately equal amounts of labeled cDNA were hybridized to all three Atlas Trial microarrays and the results were visualized by Storm 860 phosphorimager (Molecular Dynamics, Sunnyvale, CA). The scanned images were further analyzed by Scanalyze software (Stanford University, Stanford, CA) to enhance and quantify signal intensities. We found this amplification method relatively easy to handle and very robust, it produced the required amounts of cDNA in a few hours. The images presented display the full array with 96 spotted gene sequences in duplicates. The average signal strength was noticeably lower on the amplified arrays showing a gradually diminishing pattern throughout the amplification. Within this general tendency, individual genes behaved in different fashions, some diminished faster than others. A gene that was greatly over-amplified and pulled out of the background ("emerging gene") is marked by white oval. A rapidly diminishing ("dropped" gene) gene is marked with yellow oval. The overall profile (or what was left of it) after 24 PCR cycles did not resemble the original GEP. Figure 6.2. demonstrates the results of the same experiment using A2780 cells. The outcome of this experiment was very similar to the

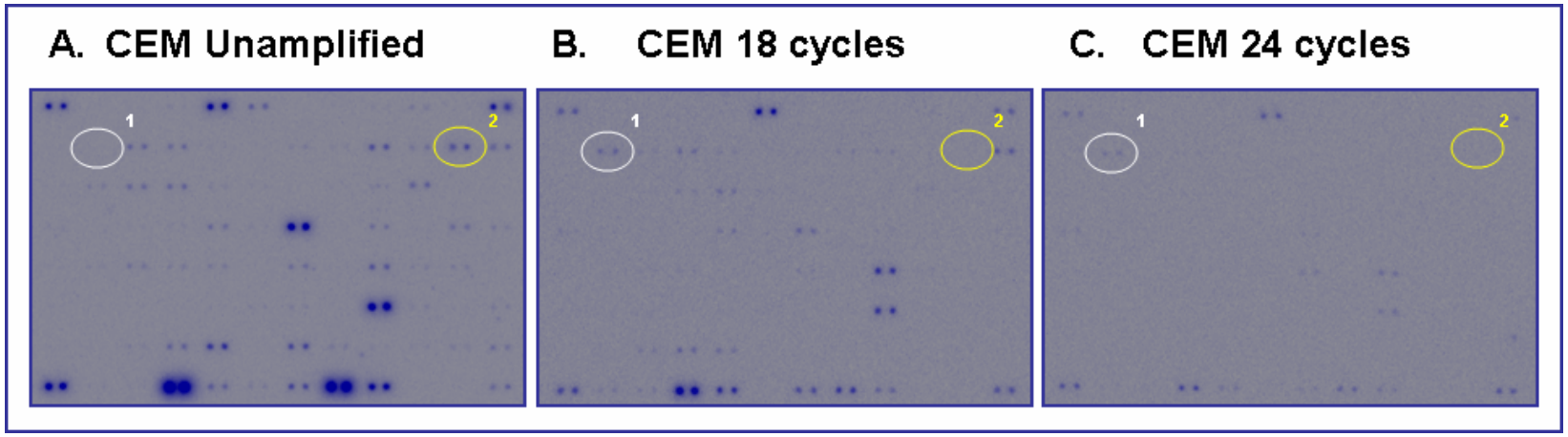

Figure 6.1. Atlas array images of exponential GEP amplification – CEM cells

Approximately equal amounts of labeled cDNA were hybridized to all three Atlas Trial microarrays. Each gene is spotted in duplicates on the array. A:10^6 CEM cells, unamplified. B:10^4 CEM cells, amplified by 18 rounds of PCR. C:10^3 CEM cells, amplified by 24 rounds of PCR. The average signal strength is gradually diminishing. Individual genes behave in different fashion. 1: A gene that is greatly overamplified and pulled out of the background is marked by white oval. 2: A rapidly diminishing gene is marked with yellow oval.

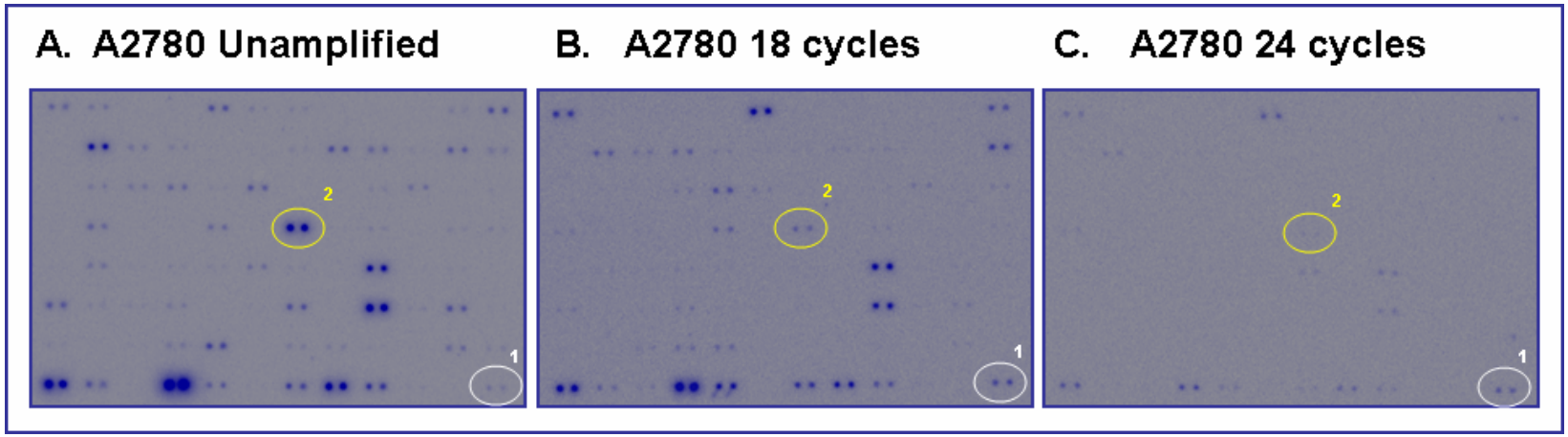

Figure 6.2. Atlas array images of exponential GEP amplification - A2780 cells

Approximately equal amounts of labeled cDNA were hybridized to all three Atlas Trial microarrays. Each gene is spotted in duplicates on the array. A:10^6 A2780 cells, unamplified. B:10^4 A2780 cells, amplified by 18 rounds of PCR. C:10^3 A2780 cells, amplified by 24 rounds of PCR. The average signal strength is gradually diminishing. Individual genes behave in different fashion. 1: A gene that is moderately overamplified is marked by white oval. 2: A rapidly diminishing gene is marked with yellow oval.

previous one, the overall signal level diminished quickly, and the characteristic GEP of A2780 cells disappeared. A gene that was moderately overamplified is marked by white oval. A rapidly diminishing gene is marked with yellow oval.

Figure 6.3. compares exponential amplification results for the two cell types. This comparison revealed that after 18 PCR cycles of amplification the characteristic GEP of the two cell types was not only mostly lost, but the two samples started to resemble to each other. Figure 6.4. confirmed that after 24 PCR cycles the original GEPs were completely lost and the two samples looked virtually identical (uniform profile). To verify these findings we compared the GEPs of unamplified and amplified samples using the scatterplot comparison method we developed and described in chapter 3. Figure 6.5. displays these scatterplots and confirmed that for both cell types the amplified profiles were very different from the original, unamplified GEP. Based on the R^2 values (0.61 for CEM cells and 0.76 for A2780 cells) representing the overall similarity (or dissimilarity) we concluded that exponential PCR significantly distorts the GEP. Furthermore, as Figure 6.6. demonstrates this amplification method diminishes differences between samples and by overamplifying a few genes and underamplifying the rest of them it produces a "uniform profile" that is almost identical for each sample, regardless its origin.

Linear RNA Amplification

After the disappointing results with exponential amplification we tested a linear amplification method using the same two cell lines (CEM and A2780 cells). We selected the commercially available MessageAmp RNA amplification kit (Ambion, Austin, TX) that utilizes T7 RNA polymerase as its main amplifier enzyme as described in Chapters 1 and 2. We analyzed the results by GeneChip® Human Genome U95Av2 (Affymetrix, Santa Clara, CA) oligonucleotide arrays that enabled us to monitor the expression levels of over 12,000 genes. For direct visualization we applied the method we developed and

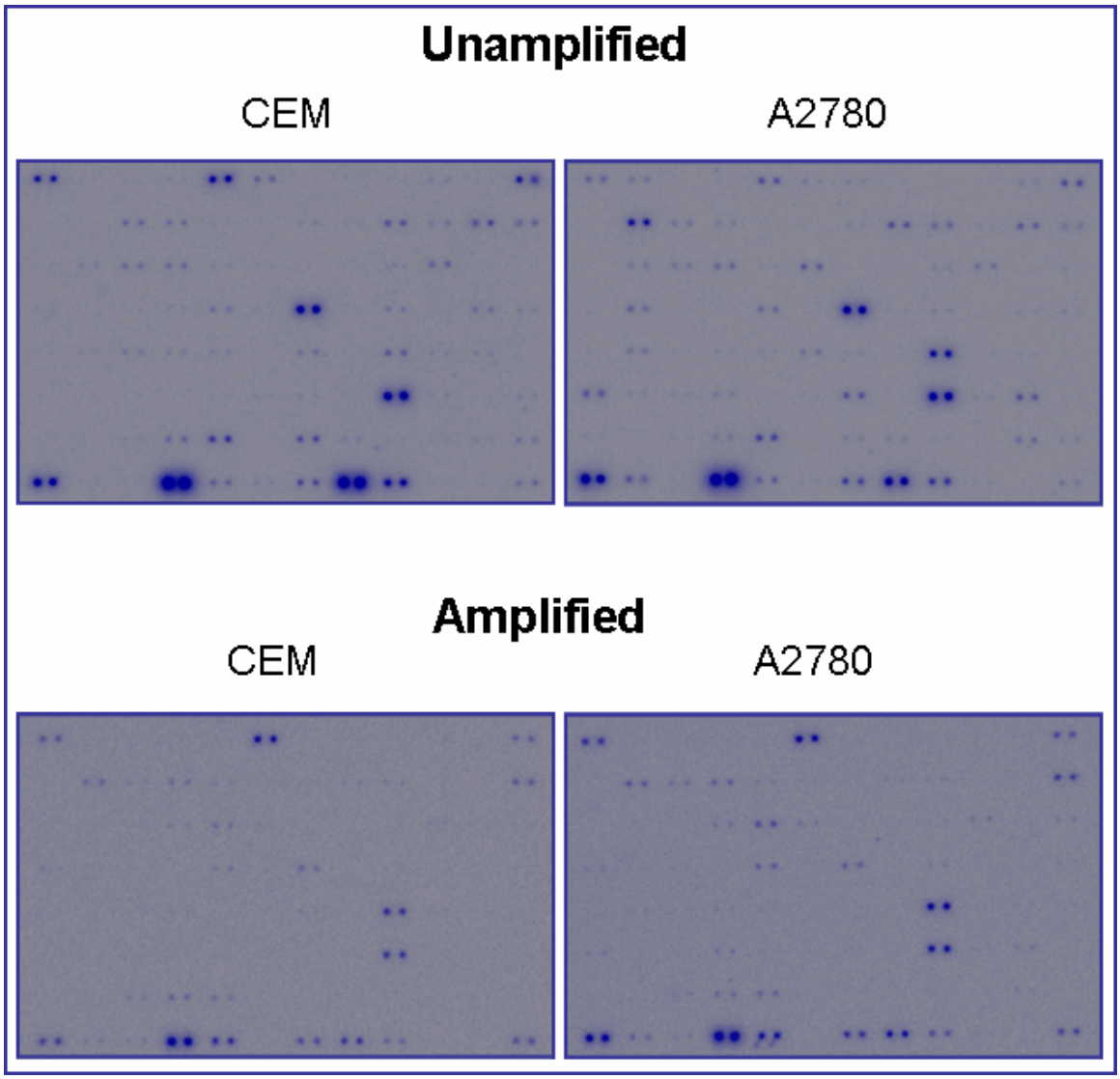

Figure 6.3. Effects of exponential amplification on the GEP of CEM and A2780 cells

Approximately equal amounts of labeled cDNA were hybridized to all Atlas Trial microarrays. Each gene is spotted in duplicates on the array. After amplification average signal intensity decreased, and the cell type specific GEP was mostly lost. The GEP of the two cell types became similar. Unamplified samples: 10^6 cells, Amplified samples: 10^4 cells amplified by 18 PCR cycles

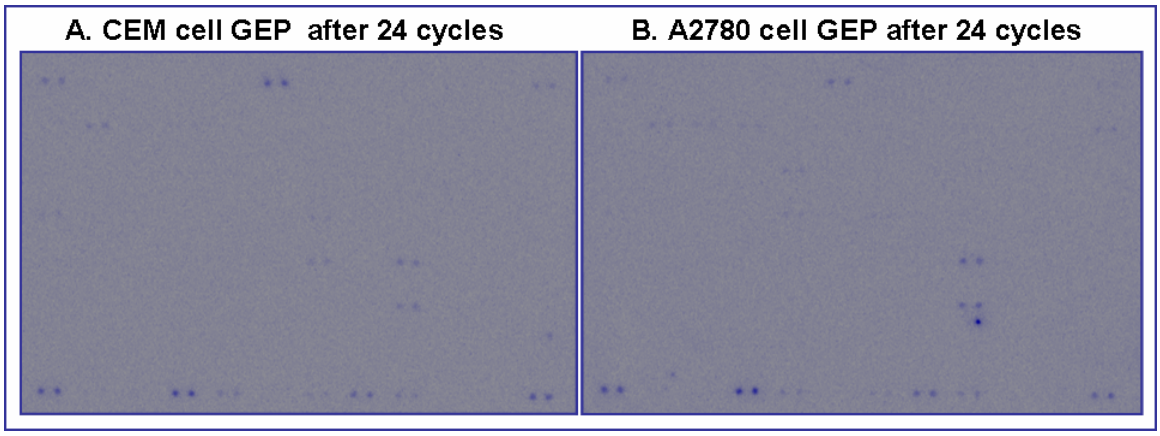

Figure 6.4. Atlas array images of amplified CEM and A2780 cell GEP

Approximately equal amounts of labeled cDNA were hybridized to both Atlas Trial microarrays. Each gene is spotted in duplicates on the array. A: 10^3 CEM cells, amplified by 24 rounds of PCR. B:10^3 A2780 cells, amplified by 24 rounds of PCR. The average signal strength is very low. The remaining signals are very similar between the two arrays. The two samples look almost identical.

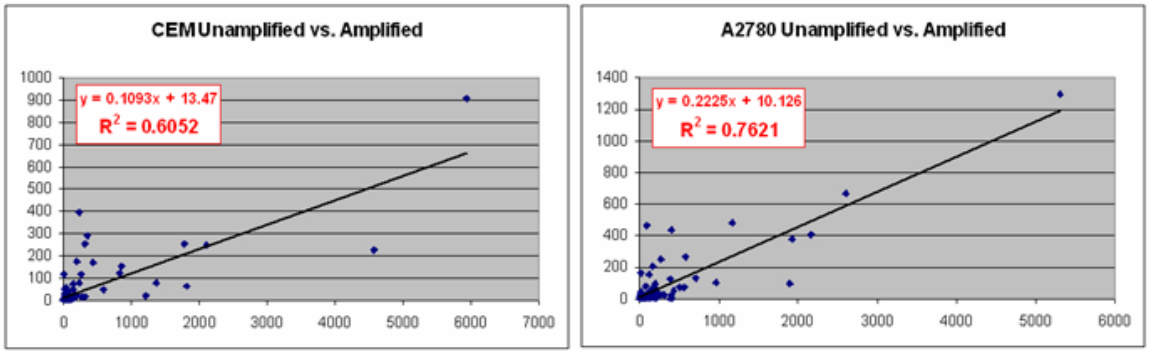

Figure 6.5. Exponential RNA amplification distorts the GEP

Scatterplots of the unamplified and amplified GEP of the same cell type. Both cell types suffered significant GEP distortion during amplification as indicated by the low R^2 values. Unamplified samples: 10^6 cells, Amplified samples: 10^4 cells amplified by 18 PCR cycles.

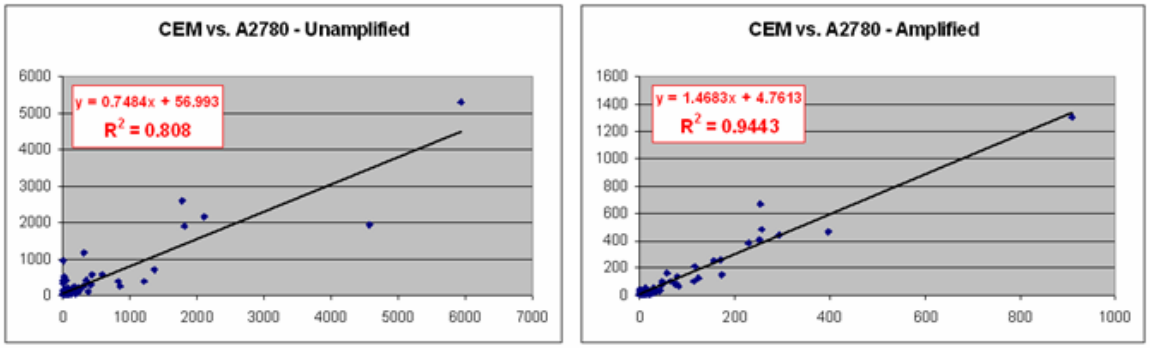

Figure 6.6. Exponential RNA amplification diminishes differences between samples

Scatterplots comparing the GEPs of CEM and A2780 cells, first unamplified then amplified. The significant difference between the two cell types diminished after amplification as indicated by the higher R^2 value between the amplified GEPs. Unamplified samples: 10^6 cells, Amplified samples: 10^4 cells amplified by 18 PCR cycles

described in the purification experiments (Chapter 3). We found the T7 RNA amplification method relatively difficult to handle, time consuming, and moderately robust producing approximately 1,000-fold increase in aRNA quantity per round. Figure 6.7. displays the pseudocolored images of the same cropped and magnified portion of each microarray we used in Chapter 3. On these images each small square-shaped feature represents a gene sequence.

The GEPs of 10^6 cells unamplified, 10^4 cells after one round, and a single cell after two rounds of T7 amplification are compared. For isolation of single cells we used our custom-built High Resolution Cell Sorter (HiReCS) system (described in chapter 2) set up for single cell sort. For RNA isolation from single cells we used our regular RNA isolation protocol described in chapter 2, using the maximum recommended volumes rather than scaling down the protocol. This way we could ensure that we did not lose the miniscule amounts of RNA (possibly sticking to the walls of the tubes and pipette tips) during the process. Since we could not quantitate picogram amounts of RNA, we processed three samples of RNA isolated from single cells separately throughout the first round of linear amplification without knowing whether we had any RNA in the tubes. We used all the products from the first round as starting material for a second round of T7 amplification and measured the amplified RNA afterwards. We found that all three single cell samples produced more than the required amount of RNA (1 microgram aRNA) for Affymetrix microarray analysis. After analyzing the cropped images we found the visual patterns of all three arrays very similar (with noticeable differences); even a single cell produced an analyzable GEP. We concluded that 2 rounds of T7 RNA amplification were sufficient to produce enough aRNA for microarray analysis even from a single cell.

Figure 6.8. demonstrates that the amplification process was highly reproducible, the R^2 values of triplicate experiments were comparable to unamplified samples; all of them were in the 0.98-0.99 range. As Figure 6.9. demonstrates, scatterplot analysis comparing unamplified and one-round amplified samples revealed that linear RNA amplification (just like exponential amplification) significantly distorts the GEP.

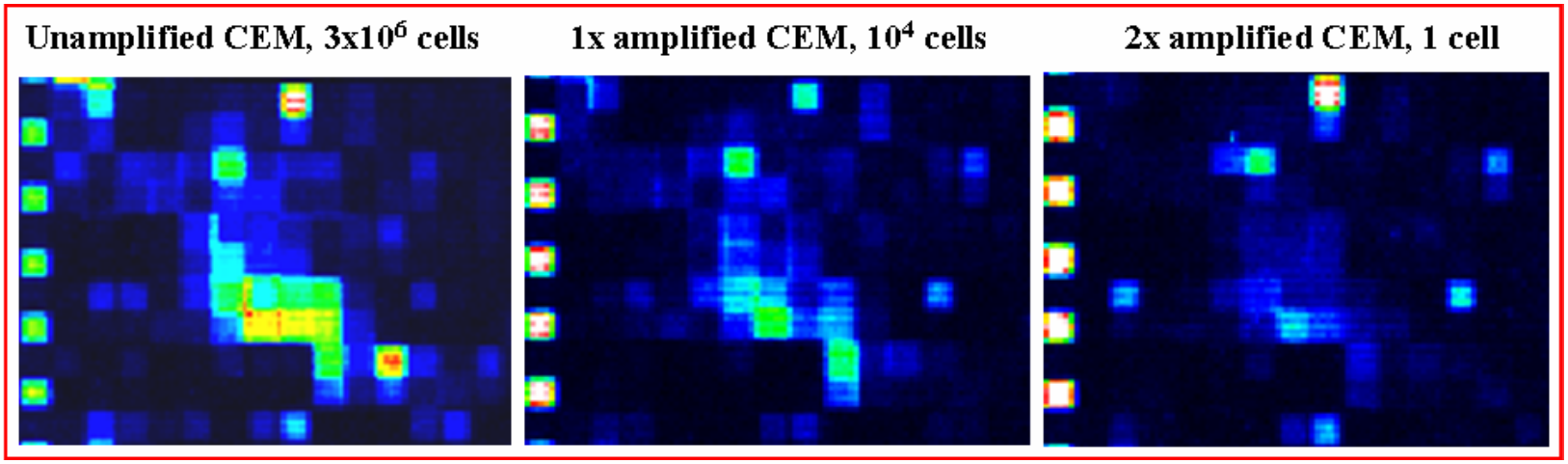

Figure 6.7. Affymetrix array images of linear GEP amplification - CEM cells

The GEPs of 10^6 cells unamplified, 10^4 cells after one round, and a single cell after two rounds of T7 amplification were compared. Each small square represents the expression level of an individual gene sequence. To visualize the raw image data of microarray analysis, corresponding segments of Affymetrix images were cropped, magnified and pseudo-colored using the Affymetrix Microarray Suite 5.0 software. Even a single cell produced an analyzable GEP. The pattern of expressed genes was similar at each stage.

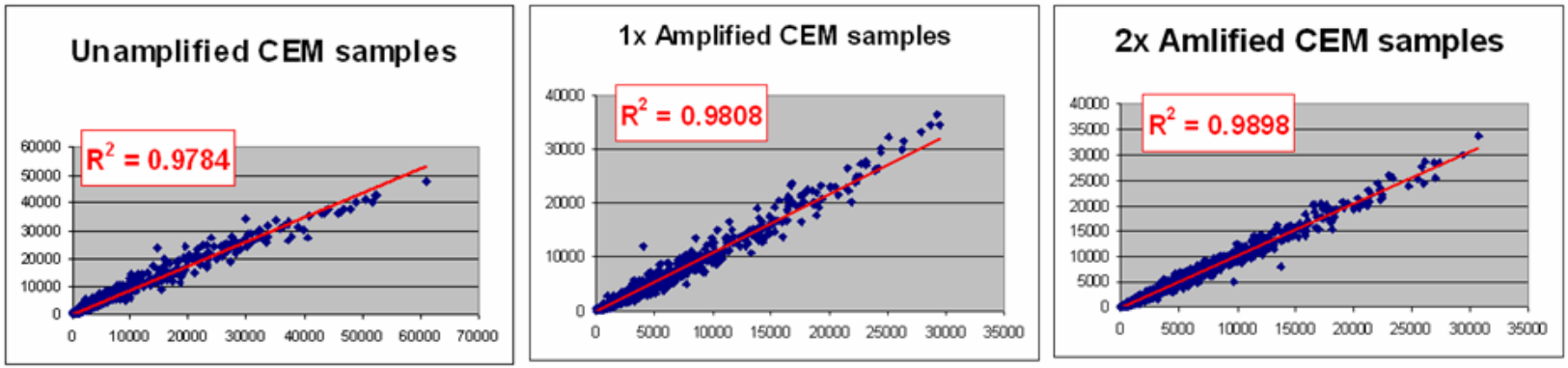

Figure 6.8. Linear amplification is highly reproducible

Scatterplots of replicate samples after each T7 amplification round. High R^2 values indicate high reproducibility.

The R^2 values for both cell types (0.69 for CEM and 0.73 for A2780 cells) were comparable to the ones produced by linear amplification. When we further investigated this GEP distorting effect of linear RNA amplification we found that both rounds distort the GEP and their effect virtually adds up (Figure 6.10.) We noted that the second round introduced less distortion than the first round. The overall GEP distortion after two rounds of linear RNA amplification was fairly high with R^2 values of approximately 0.5 for both cell types.

Figure 6.11. demonstrates that the differences between the overall GEPs of CEM and A2780 cells were not diminished by linear RNA amplification. The R^2 values of comparing the two cell types remained about the same (approximately 0.8) after each round. We evaluated if there was any loss in overall signal strength caused by linear amplification. Figure 6.12. compares the average signal levels for both methods throughout the amplification process and demonstrates that while with the exponential method this level diminished rapidly, with linear amplification it remained steady with less than 10% alterations between readings.

GEP Distortion after Linear Amplification

After the initial results showing that although linear amplification distorts the overall GEP, this distortion is very reproducible, does not lead to loss in average signal level, and does not diminish GEP differences between samples, we further explored the characteristics of this phenomenon. We demonstrated earlier that exponential amplification selectively picked up and greatly overamplified a few genes even if they only had background level expression level in the unamplified sample ("emerging genes"). To test if linear amplification did the same we studied the behavior of genes with very low raw signal levels in both cell types. Figure 6.13 shows the amplified versus unamplified GEP scatterplots of genes that were absent in the unamplified samples (raw expression value below 50 intensity units on Affymetrix microarrays). In both A2780 and

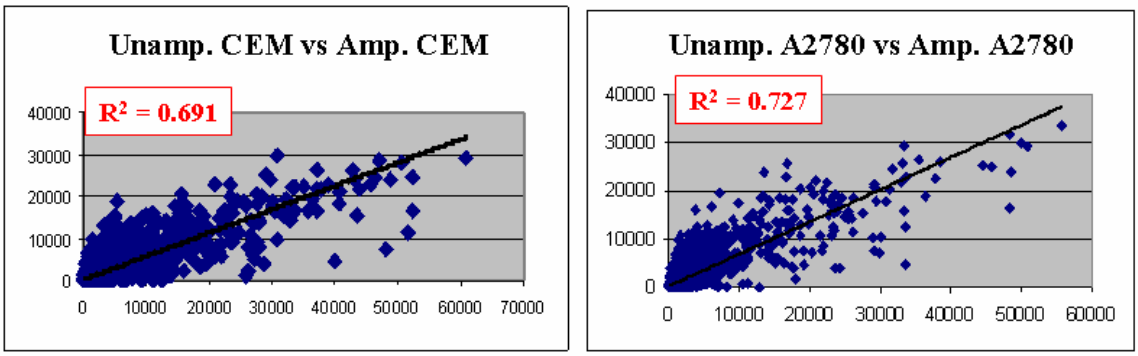

Figure 6.9. Linear RNA amplification distorts the GEP

Scatterplots of the unamplified and amplified GEP of the same cell type. Both cell types suffered significant GEP distortion during amplification as indicated by the low R^2 values. Unamplified samples: 10^6 cells, Amplified samples: 10^4 cells amplified by 1 round of T7 RNA amplification.

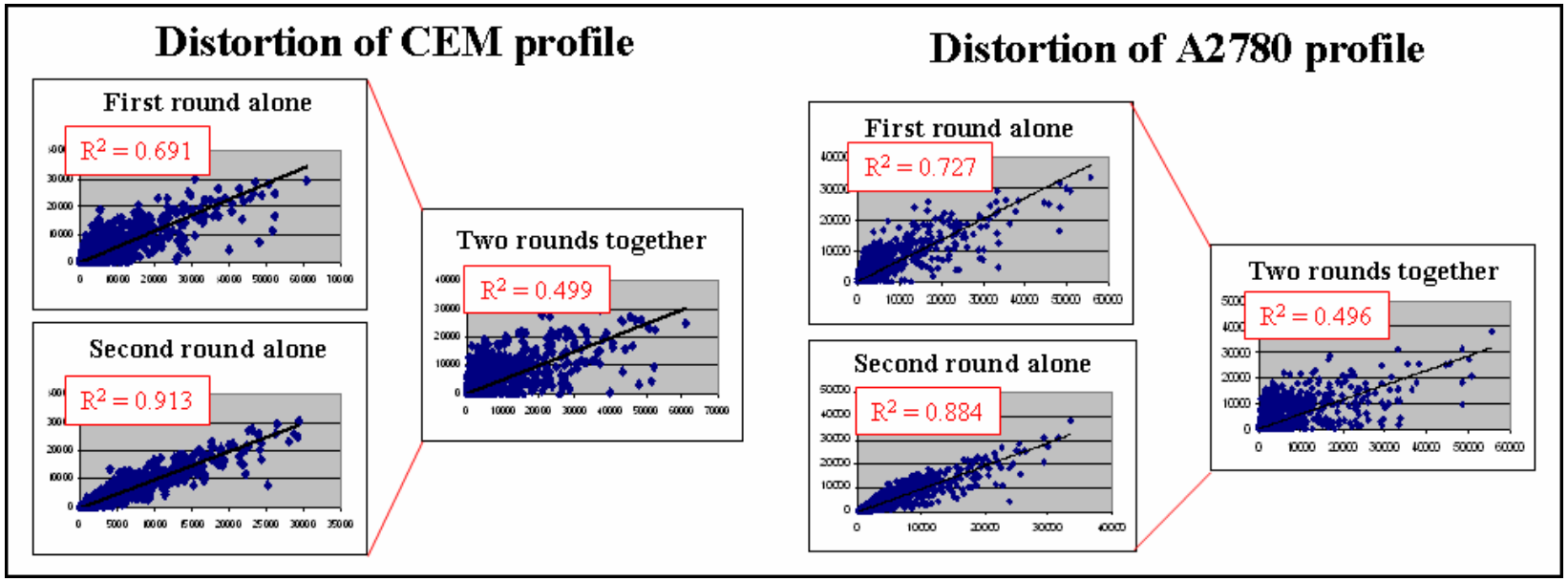

Figure 6.10. Linear RNA amplification distorts the GEP in each round

Each amplification round is plotted individually and together. CEM and A2780 cells suffered very similar GEP distortion. Lower R^2 values correspond to lower correlation.

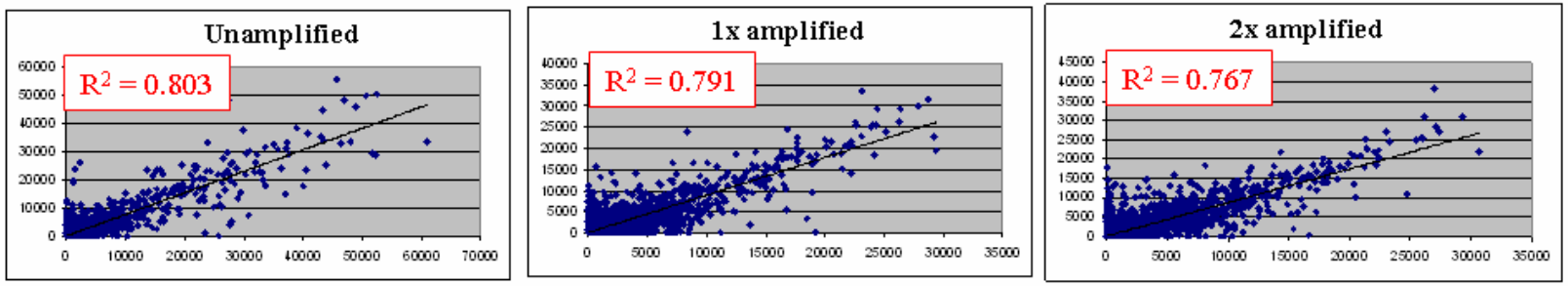

Figure 6.11. Linear RNA amplification preserves differences between samples

The GEPs of CEM and A2780 cells were plotted against each other after each round of T7 amplification. Similar R^2 values indicate similar correlations between the two cell types after each round. The R^2 value did not increase after amplification.

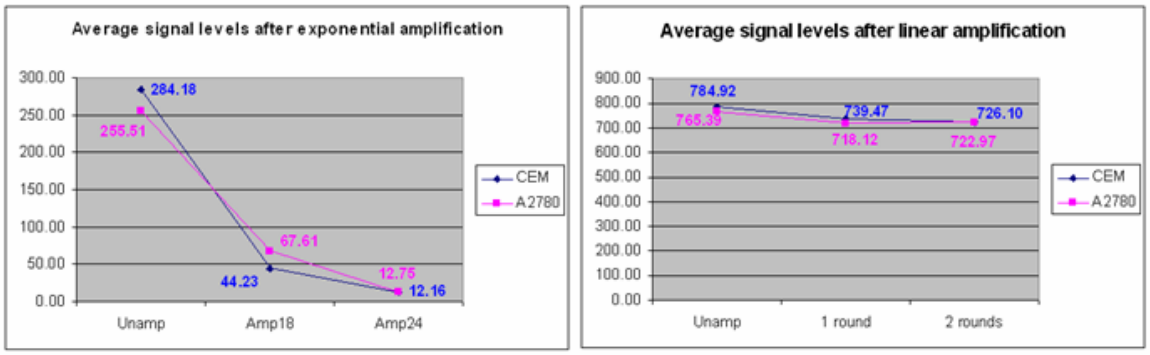

Figure 6.12. Average gene expression levels after amplification

Average signal levels over background dropped rapidly during exponential amplification but remained stable during linear amplification. Ordinate: gene expression level above background in raw data units of the reader. Abscissa: level of amplification. Exponential: 0, 18, 24 PCR cycles; Linear: 0, 1, 2 rounds.

CEM cells very few genes "emerged" from the background and posed as expressed genes after amplification (11 out of 3309 genes in CEM cells and 23 out of 2866 genes in A2780 cells). The amplified samples in this experiment went through one round of T7 RNA amplification.

To study the extent of the distortion we investigated the genes that were surely expressed in both CEM and A2780 cells in the unamplified samples. We found 5209 genes above 200 raw expression intensity units and followed their behavior after the first round of linear amplification (Figure 6.14.). When we established a fairly forgiving threshold of 3.0–fold for overamplification and 0.33-fold for underamplification we found only a sum of 307 genes overamplified in either cell type and 1129 underamplified ones. The overamplified genes constituted about 5% of all common genes and the underamplified ones were approximately 20%. Altogether, approximately 75% of all commonly expressed genes stayed within acceptable range with these thresholds. Figure 6.14. was constructed by plotting the amplified/unamplified ratios of each gene in a given cell type against the same ratios in the other cell type. This way we could compare "amplified behavior" for each gene between the two cell types. The figure shows that the overamplified genes behaved very similarly between the two cell types. Only very few of them differed by more than 2-fold from the mean "amplified behavior" as represented by the parallel green lines. In other words, the genes behaved similarly between the two cell types even the "misbehaving" ones.

To further study if GEP distortion mainly depends on the gene sequence itself (and not on the initial expression level of the given gene in the sample, as one would intuitively think) we examined the 1129 underamplified genes from the main sample. Figure 6.15 was constructed similarly to Figure 6.14. but only shows the genes that were underamplified in at least one of the samples. This approach revealed that only 17 genes out of 1129 were overamplified in the other sample (discordant genes), only 5 of them were outside the 2x zone (surely discordant genes), and none outside the 2.5x zone (seriously discordant genes). This result corroborated our notion that a gene tends to

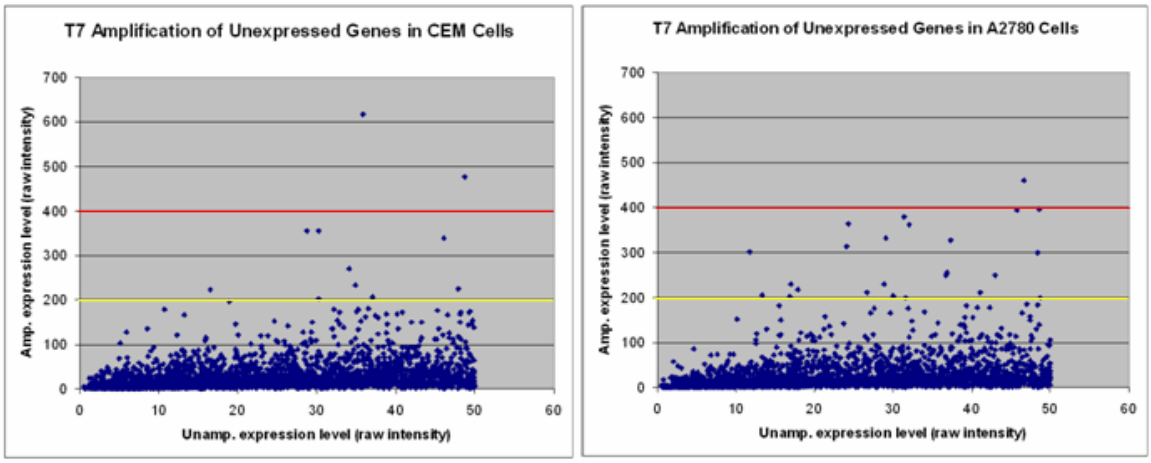

Figure 6.13. Linear amplification does not overamplify absent genes

Amplified versus unamplified GEP scatterplots of genes that were absent in the unamplified samples (raw expression value below 50 on Affymetrix microarrays). Both in A2780 and CEM cells very few genes "emerge" as expressed genes (11 out of 3309 genes in CEM cells and 23 out of 2866 genes in A2780 cells). Amplified samples went through one round of T7 RNA amplification.

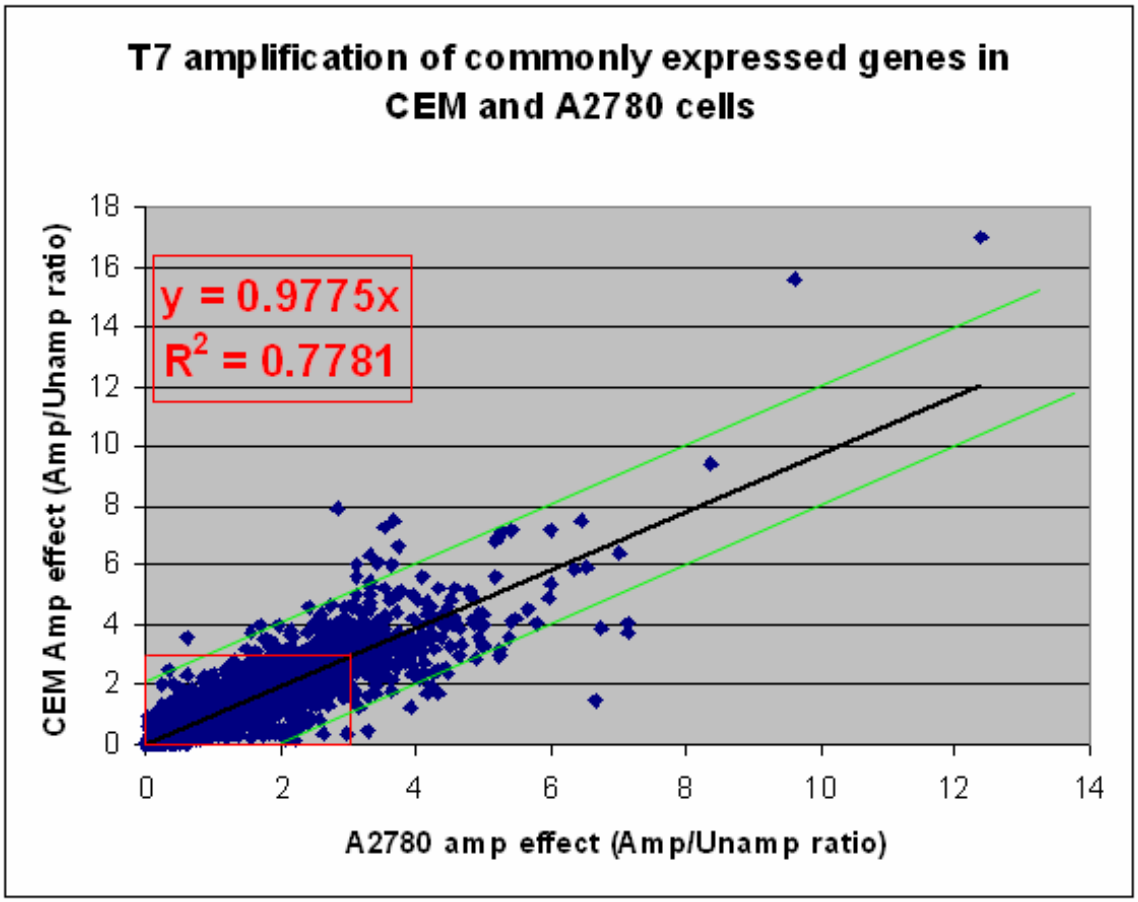

Figure 6.14. Linear amplification distorts genes similarly in different samples

Genes that were expressed in both unamplified CEM or A2780 cells were plotted. Only 307 genes out of 5209 were overamplified in either one of the samples (genes outside the red box). Only a few genes were overamplified in a discordant fashion (genes outside the green zone). Amplified samples went through one round of T7 RNA amplification.

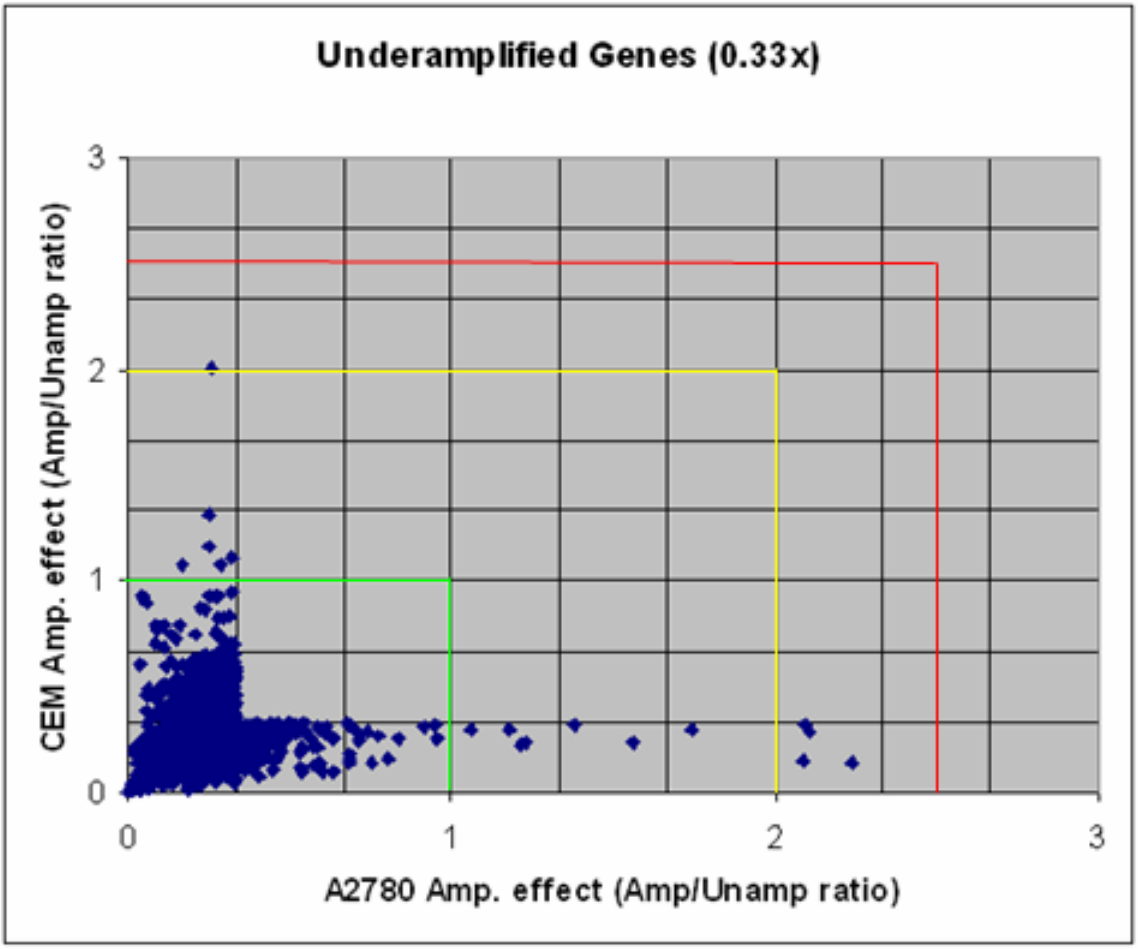

Figure 6.15. Linear amplification underamplifies the same genes from different samples

Genes that were underamplified (0.33 amplified/unamplified expression ratio) in either CEM or A2780 cells were plotted. Only 17 genes out of 1129 were overamplified in the other sample (discordant genes), only 5 of them were outside the 2x zone (surely discordant genes), and none outside the 2.5x zone (seriously discordant genes). Amplified samples went through one round of T7 RNA amplification.

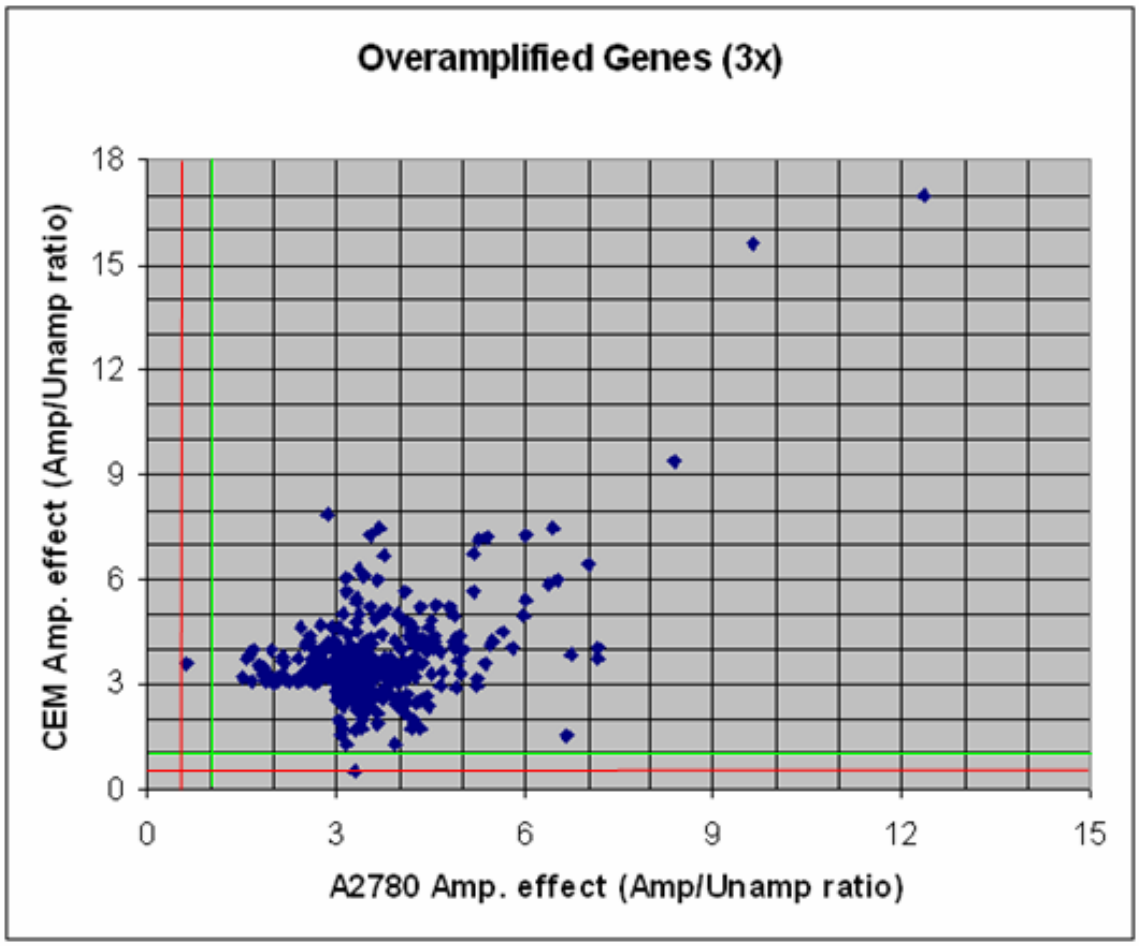

Figure 6.16. Linear amplification overramplifies the same genes from different samples

Genes that were overamplified (3.0 amplified/unamplified expression ratio) in either CEM or A2780 cells were plotted. Only 2 genes out of 307 were underamplified in the other sample (discordant genes), only one of them was outside the 0.5x zone (surely discordant genes), and none was outside the 0.4x zone (seriously discordant genes). Amplified samples went through one round of T7 RNA amplification.

behave the same way when amplified; the genes that were underamplified in one sample did not get overamplified in the other.

We investigated the 307 overamplified genes in a similar fashion (Figure 6.16.) Genes that were overamplified (3.0 amplified/unamplified expression ratio) in either CEM or A2780 cells were plotted. Only 2 genes out of 307 were underamplified in the other sample (discordant genes), only one of them was outside the 0.5x zone (surely discordant genes), and none was outside the 0.4x zone (seriously discordant genes). Again, genes that were overamplified in one sample did not underamplify in the other. We noted that this similar "amplified behavior" of genes between samples did not depend on similar expression levels, many of them started from several fold higher level in one sample than in the other and still behaved the same way (got overamplified or underamplified) in both samples. We concluded that the main factor that determines the direction and extent of distortion a gene suffered by amplification was the gene (and its sequence) itself, not the original expression level it started with or the gene environment (GEP) around it.

Finally, we studied whether changing the amplifying enzyme changed the distortion pattern. The MessageAmp-2 kit contains a modified T7 RNA polymerase giving us the opportunity to compare the effects of two (slightly) different linear amplifier enzymes. We found that the modified enzyme produced just as highly reproducible results as the old enzyme and distorted the GEP of amplified cells to about the same extent (not shown). However, the overall GEP distortion patterns introduced by the two enzymes were different, as shown in Figure 6.17. We concluded that the GEP distortion was enzyme-specific.

Microarray Analysis of LEAP Purification and Optoinjection

Using our conclusions from the model experiments described above we analyzed the GEP of the samples we produced by Laser Enabled Analysis and Processing in Chapter 5. We used linear RNA amplification on all samples. Figures 6.18. and 6.19. show that LEAP purification did not distort the GEP of purified KG-1 cells. The slightly

low R^2 value of 0.966 when compared to the pure KG1 GEP could be due to the fact that only 85% purity was achieved. Nevertheless, this R^2 value was still significantly higher than that of the 50% CEM/KG1 cell mixture (0.898). Figure 6.19 displays a computer-software generated heat map that was prepared similarly to the one described in chapter 3. It shows all three replicates of the pure KG1 sample, the 50% CEM/KG1 cell mix and the purified KG1 cells. It demonstrates that the purified KG1 cells recovered most of the original GEP of the pure KG1 cells. Triplicate results were found to be virtually identical.

Figures 6.20. and 6.21. display the results of an optoinjection experiment where KG1 cells were optoinjected with fluorescent dextrans. The figures demonstrate that despite the fact that we found no microscopic evidence of any cellular damage after optoinjection, LEAP optoinjection did distort the GEP of optoinjected KG-1 cells with or without the dextran. The R^2 values of 0.91 and 0.93 are indicating significant GEP distortion, and the heat map confirmed that the GEP of these samples were indeed different. Triplicate results were found to be virtually identical proving that the distortion was not an artifact of a single, failed microarray analysis. Interestingly, the mock- and dextrane-optoinjected samples were also different when compared to each other, and the dextran-optoinjected cells were slightly more similar to control cells than the mock optoinjected ones.

Prediction of Unamplified GEP from Amplified GEP – GEP Reconstruction

Based on the results of our model RNA amplification experiments we concluded that the GEP distortion introduced by linear amplification was not random, it was highly reproducible and affected the same genes in different samples similarly, regardless of their original expression level or the GEP they were part of, while preserving the GEP differences between different samples. (Figures 6.8.-6.16.) These characteristics of linear GEP distortion enabled us to use the known distortion factor (amplified/unamplified

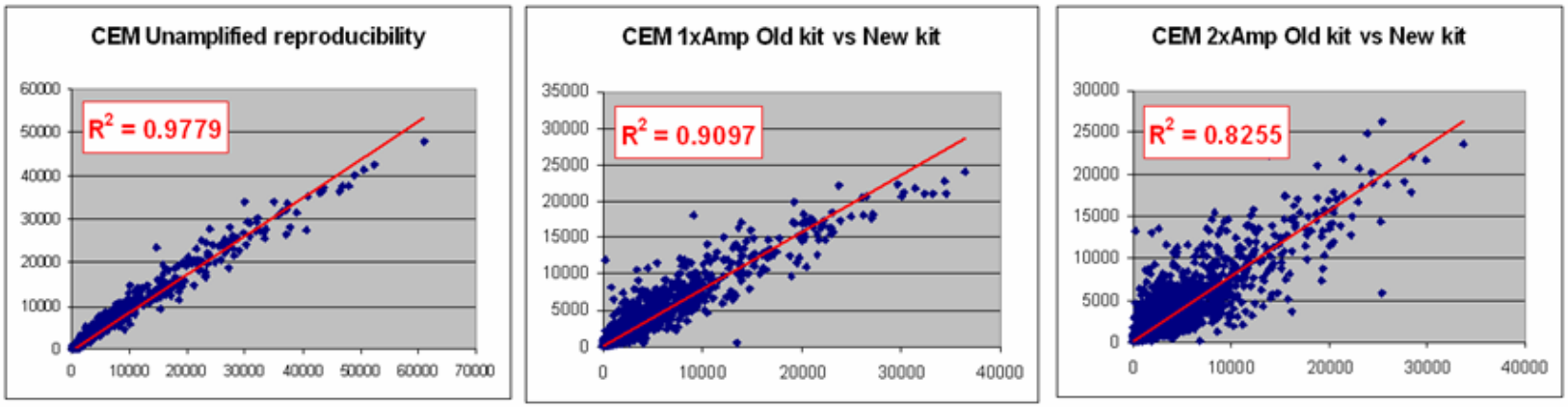

Figure 6.17. Modified T7 RNA polymerase introduces different distortion pattern from the original enzyme

CEM GEP was amplified with the original T7 RNA polymerase and with a modified version of it. The GEPs were plotted after each round of amplification. Each round increases the difference in the GEP of the amplified samples.

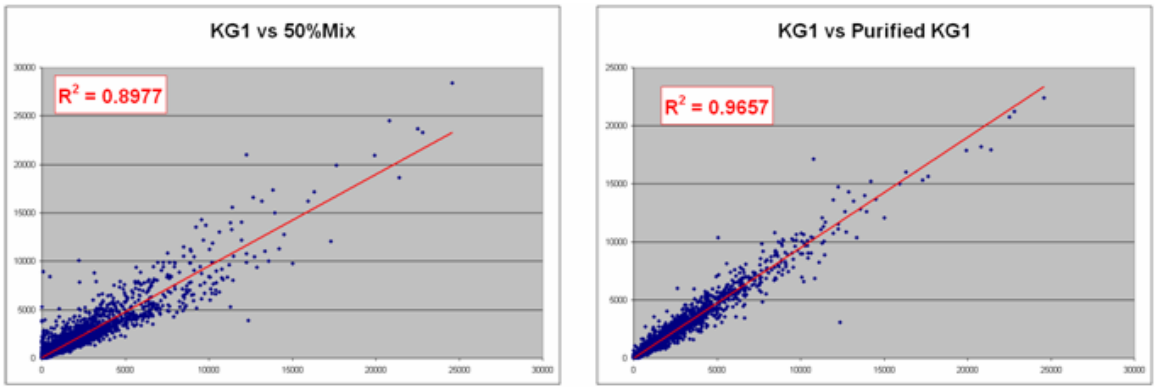

Figure 6.18. LEAP purification preserves the GEP

Scatterplots representing a heterogeneous sample before and after LEAP sort. KG1 cells were purified from a 50% KG1/CEM cell mixture to approximately 85% purity. One round of linear amplification was applied to the purified cells and to the control cells as well. The GEP of the purified KG1 cells was almost perfectly recovered indicated by increased R^2 value.

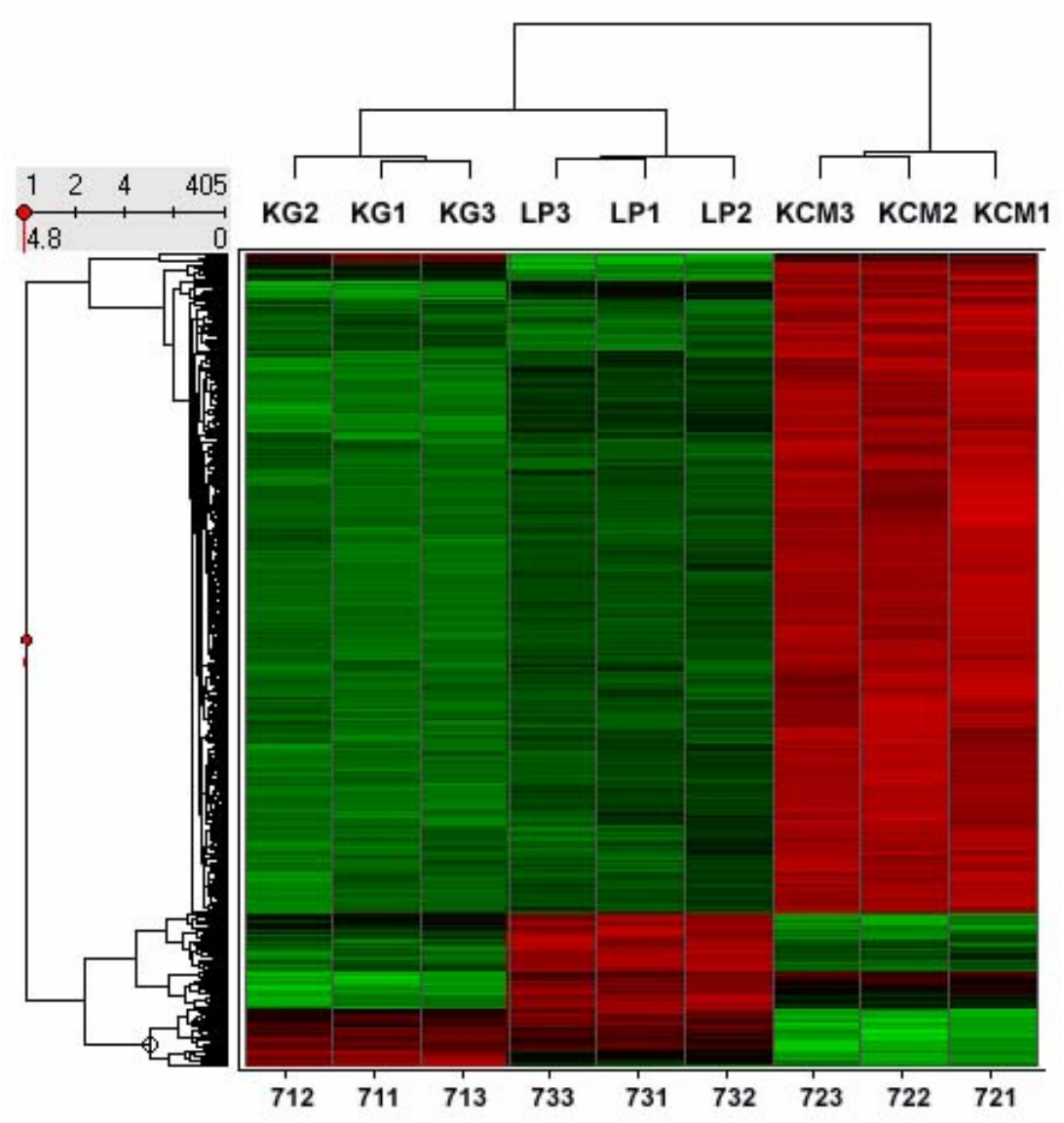

Figure 6.19. LEAP purification preserves the GEP – Heat map

Scatterplots representing a heterogeneous sample before and after LEAP sort. KG1 cells were purified from a 50% KG1/CEM cell mixture to approximately 85% purity. One round of linear amplification was applied to the purified cells and to the control cells as well. The GEP of the purified KG1 cells was very close to the pure profile. Replicate samples are virtually identical. KG: pure KG1 cells, LP: LEAP-purified KG1 cells, KCM: KG1/CEM mix.

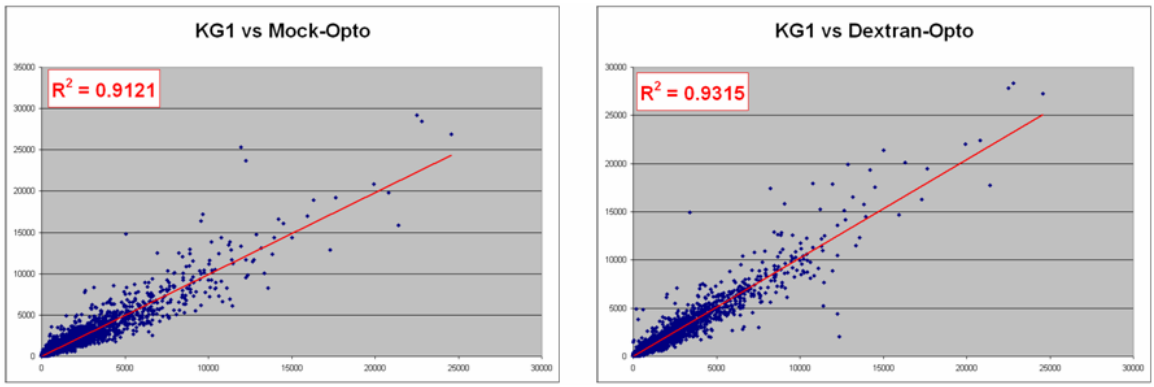

Figure 6.20. LEAP optoinjection moderately distorts the GEP

Scatterplots plotting unprocessed KG1 cells against mock-otoinjected, and dextran-otoinjected cells. Both optoinjection significantly reduced the R^2 value indicating GEP distortion. One round of linear amplification was applied to the purified cells and to the control cells as well.

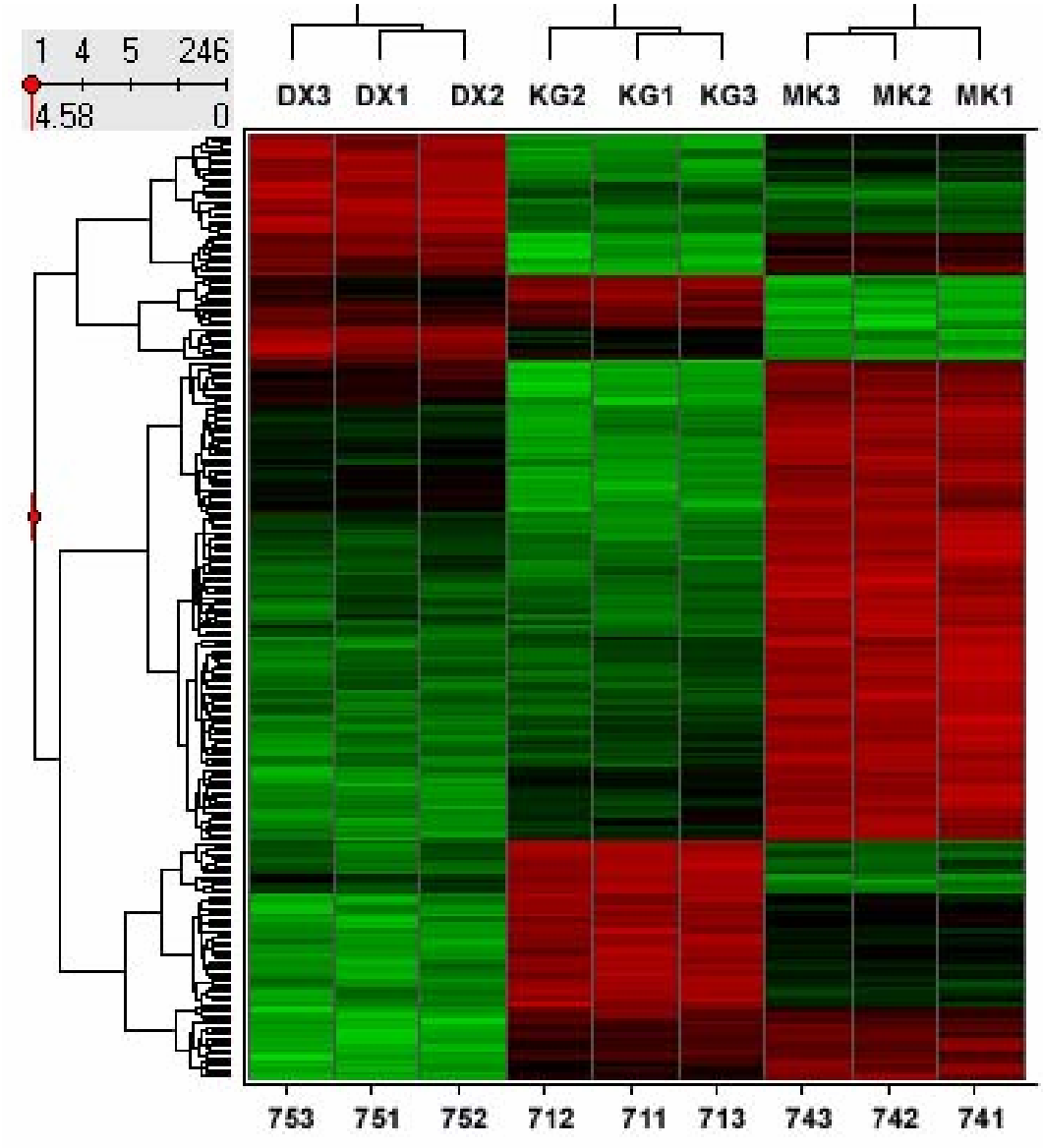

Figure 6.21. LEAP optoinjection alters the GEP – Heat map

Scatterplots plotting unprocessed KG1 cells against mock-otoinjected, and dextran-otoinjected cells. Both types of optoinjection significantly altered the GEP. One round of linear amplification was applied to the purified cells and to the control cells as well. Replicate samples are very similar. KG: pure KG1 cells, DX: dextran-otoinjected KG1 cells, MK: mock-otoinjected KG1 cells.

expression ratio for a given gene) from one sample (A2780 cells) to predict the original GEP from the distorted, amplified GEP of another sample (CEM cells). To do this GEP reconstruction, first we selected a set of genes that were surely expressed in the unamplified profile of both cell types. We used the same set of 5,209 genes we selected earlier on the bases of their unamplified expression level being above 200 raw expression intensity values. As Figure 6.22. shows this set of genes displays the same amplification pattern as the full array of 12,559 genes we used in earlier analyses. Using this gene set, the unamplified GEPs between A2780 and CEM cells were still different (R^2=0.82), and the difference between their amplified GEPs was preserved (R^2=0.81). The GEP distortion introduced by one round of RNA amplification was also similar (R^2=0.71 for A2780 cells and R^2=0.74 for CEM cells). Figure 6.23 demonstrates that without any additional modification, just by applying the distortion factor from the A2780 data set to the distorted GEP of amplified CEM cells, the predicted CEM profile partially reconstructed the original GEP to the level of the R^2=0.88. This prediction can be further improved just by eliminating the few genes (8 genes altogether) that had a predicted value of 50,000 or more (extremely high predictions). As Figure 6.24 presents, the predicted GEP came closer to the original GEP to a difference of only R^2=0.93. We noted that 7 out of the 8 discarded genes were not much distorted either, shown by the scatterplots of Figure 6.23. Further improvements can be achieved in the prediction of unamplified profiles after "training" the prediction method. As an example, Figure 6.25. displays the predicted GEP after eliminating genes with distortion factor ratios more than 2.0 or less than 0.5 between CEM and A2780 cells. This rather strict "training" still left 4,724 genes for analysis and achieved an R^2=0.97 reconstruction of the original profile. Similar methods lead to even more impressive results starting from the seriously distorted GEP of samples that were amplified by 2 rounds of T7 amplification. As Figure 6.26. demonstrates from a GEP with R^2=0.53 distortion the predicted GEP reconstructed the original GEP to a level of R^2=0.94.

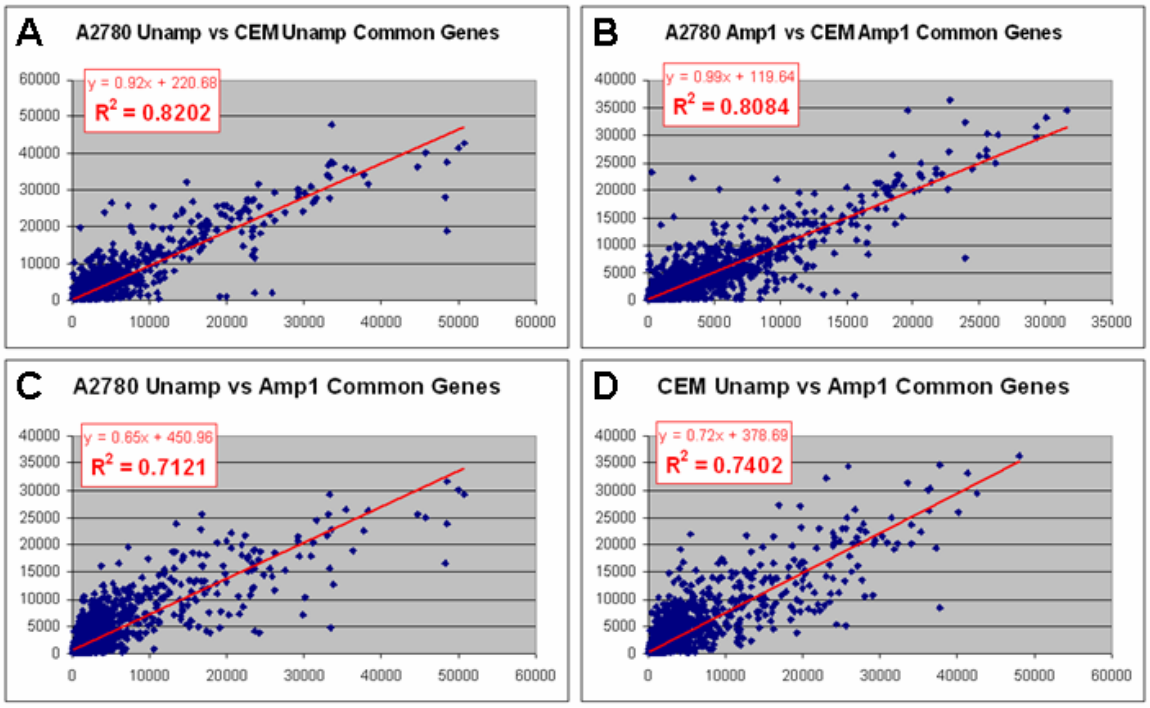

Figure 6.22. Scatterplots of genes expressed in both A2780 and CEM cells – Linear amplification

Genes that were expressed in both unamplified CEM or A2780 cells were plotted. The 5,209 commonly expressed genes show the same patterns like the full array of 12,559 genes did. The GEPs of CEM and A2780 cells are different before (A) and after (B) amplification (R^2=0.80-0.82). Amplified GEPs of these genes are distorted to R^2=0.71-0.74 (C and D). Amplified samples went through one round of T7 RNA amplification.

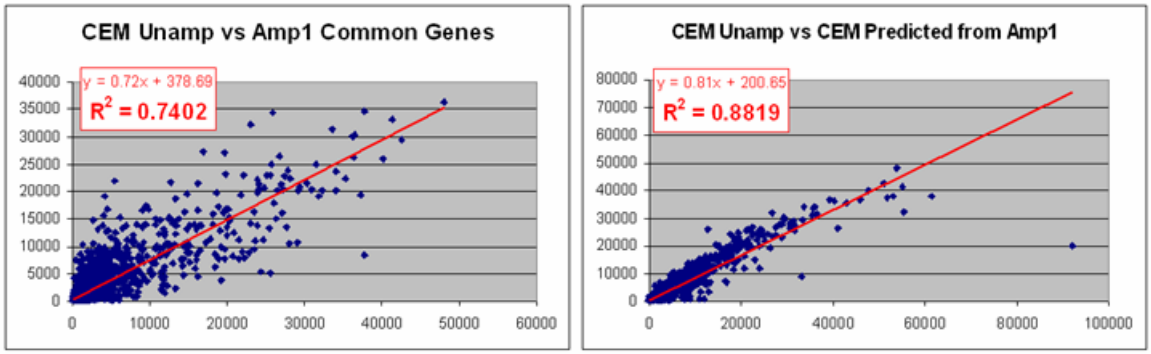

Figure 6.23. Reconstructed GEP of CEM cells after linear amplification

Genes that were expressed in both unamplified CEM or A2780 cells were plotted. The GEP of CEM cells was predicted from amplified GEP using the distortion factors of the same genes in A2780 cells. This prediction reconstructed the CEM GEP to R^2=0.88. Amplified samples went through one round of T7 RNA amplification.

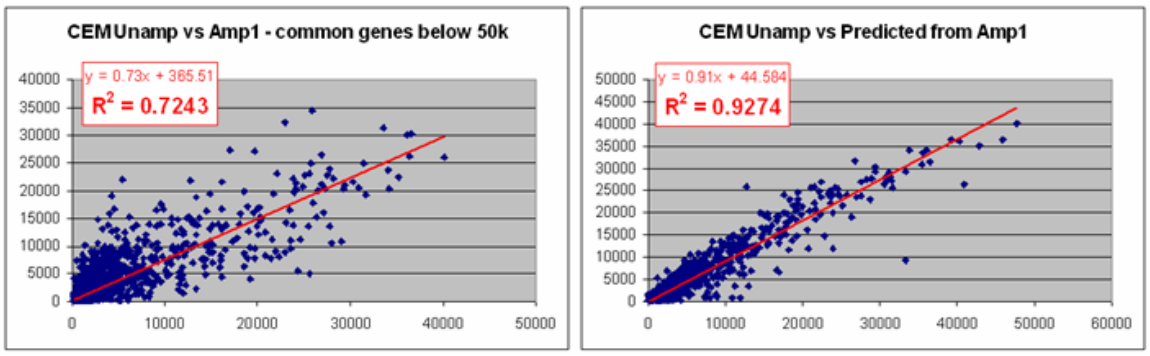

Figure 6.24. Reconstructed GEP of CEM cells after linear amplification – genes with predicted expression level below 50,000

Genes that were expressed in both unamplified CEM or A2780 cells were plotted. The GEP of CEM cells was predicted from amplified GEP using the distortion factors of the same genes in A2780 cells. Genes with predicted expression levels above 50,000 units were discarded (only 8 genes). This prediction reconstructed the CEM GEP to R^2=0.93. Amplified samples went through one round of T7 RNA amplification. 5,201 genes were plotted.

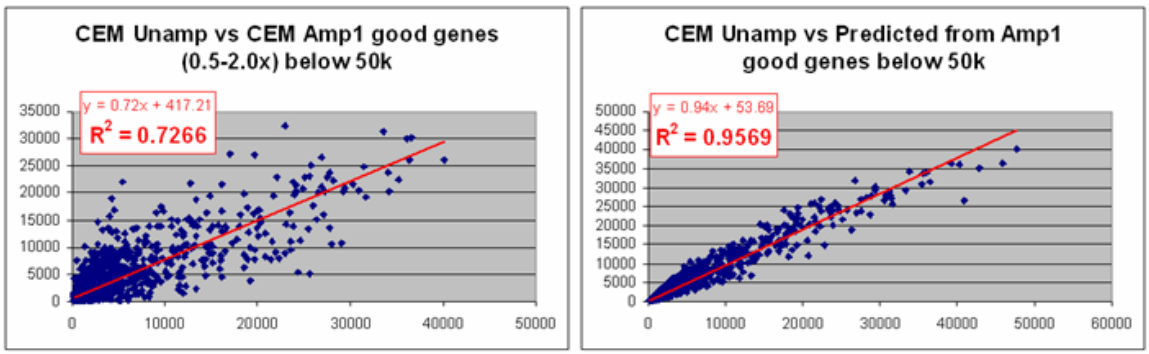

Figure 6.25. Reconstructed GEP of CEM cells after linear amplification – "good" genes with predicted expression level below 50,000

Genes that were expressed in both unamplified CEM or A2780 cells were plotted. The GEP of CEM cells was predicted from amplified GEP using the distortion factors of the same genes in A2780 cells. Genes with predicted expression levels above 50,000 units were discarded. Only genes with 0.5-2.0 distortion factor ratio between CEM and A2780 cells were plotted. This prediction reconstructed the CEM GEP to R^2=0.96. Amplified samples went through one round of T7 RNA amplification. 4,724 genes were plotted.

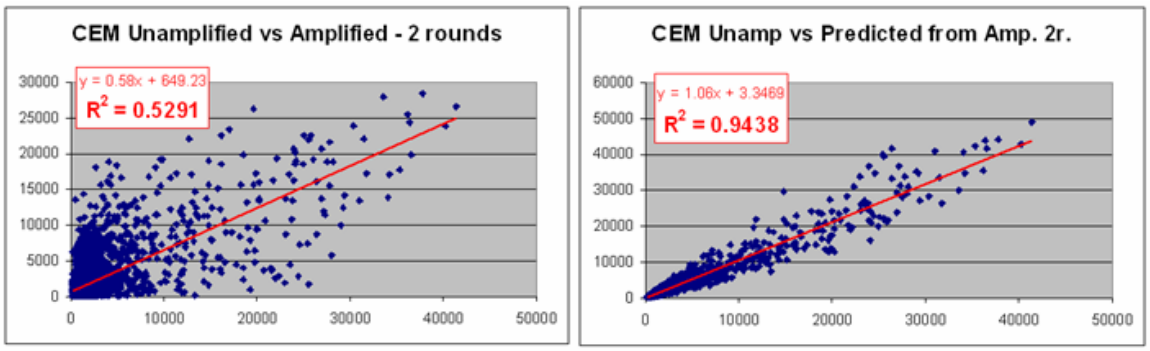

Figure 6.26. Reconstructed GEP of CEM cells after 2 rounds of linear amplification – "good" genes, below 40,000 predicted level

Genes that were expressed in both unamplified CEM or A2780 cells were plotted. The GEP of CEM cells was predicted from amplified GEP using the distortion factors of the same genes in A2780 cells. Genes with predicted expression levels above 40,000 units were discarded. Only genes with 0.5-2.0 distortion factor ratio between CEM and A2780 cells were plotted. This prediction reconstructed the CEM GEP to R^2=0.94. Amplified samples went through two rounds of T7 RNA amplification. 4,505 genes were plotted.

DISCUSSION

We have studied the capabilities and limitations of both exponential and linear RNA amplifications to analyze the GEP of small, purified cell samples. Our earlier conclusions, that for meaningful microarray data a reasonable purification of the target cell subpopulation is both necessary and achievable, lead to the problem that real biological samples, especially after cell purification, often do not provide the necessary amounts of RNA, for direct gene expression microarray analysis. Furthermore, in some cases – purified rare cell subsets, like stem-progenitor blood cells, aspiration biopsies, micromanipulated cells, like LEAP-optoinjected individual cells – GEP analysis of a few or even single cells would be desirable.

Since for meaningful GEP analysis it is not enough to obtain the required amount of RNA or cDNA, it is also extremely important to keep the relative frequencies of each mRNA species at the original ratios throughout the amplification process, we analyzed the amplified GEPs of model cell lines and compared them to the GEPs of unamplified samples and to each other to investigate the extent of possible GEP distortions. We used two different cell lines, A2780 and CEM cells, with characteristically different GEPs to evaluate how accurately the two methods preserved the GEP differences between the two cell types. We established that the unamplified GEPs of the two cell lines were significantly different with R^2-value of 0.808 when plotting them against each other.

When testing our selected exponential RNA amplification method that uses Taq-polymerase as amplifier enzyme, we found that from 10^4 cells 18 PCR-cycles generated equivalent amounts of cDNA to an unamplified, 10^6 cell sample. From 10^3 cells 25 PCR cycles were necessary. We found this technology relatively easy to handle and very robust - it produced the required microgram amounts of cDNA within a few hours. When we analyzed if there was any significant GEP distortion introduced by this method we found that several genes expressed at a high or medium level in the unamplified GEP disappeared from the amplified GEP of the same cell type (“dropped genes”). Other genes that were expressed at a low level or were undetectable in the unamplified profile

appeared to be expressed at a high or medium level in the amplified profile (over-amplified genes). Ratios of differentially expressed genes were often found to be reversed when comparing unamplified and amplified GEPs of the same cell type. Comparing the GEPs of an amplified sample to the GEP of the same sample without amplification showed significant GEP distortion with R^2-values of 0.60-0.76 after 18 cycles and 0.40-0.71 after 25 cycles. Furthermore, the original significant differences between the unamplified GEPs of A2780 and CEM cells were almost entirely diminished after PCR amplification. The R^2-value increased to 0.944 between CEM and A2780 cells. Finally, after 18 cycles the average relative expression level of the expressed genes dropped to one-fifth to one-seventh of the unamplified level. After 25 cycles this tendency continued (reaching a 20 to 30-fold drop) to the point where only a few genes showed significant expression values above background. Based on these results of the exponential RNA amplification experiments, we concluded that this method distorted the overall GEP beyond recognition, at the same time greatly diminished GEP differences between samples and resulted in a huge loss of signal over background ratio. For these reasons we concluded that while PCR might be a good method for genome-wide cDNA library construction where the relative ratios of the individual cDNA species are not important it was not suitable for gene expression microarray profiling.

When we tested a linear RNA amplification method that uses T7 RNA-polymerase for amplification, we found that from 10^3 cells or more one round of T7 polymerase-based, linear RNA amplification (T7-amplification) provided enough RNA for microarray analysis. Smaller samples required two rounds to produce equivalent amounts of RNA to an unamplified sample of 10^6 cells confirming the notion that one round of T7 amplification results in approximately 1000-fold RNA increase. We demonstrated that 2 rounds of T7-amplification were enough for even a single cell microarray and that a single cell provided a valid GEP. We found this technology relatively difficult to handle and time consuming since two rounds of amplification normally took 3 days to produce microarray ready RNA. To evaluate if there was any significant GEP distortion introduced by this method, we compared the T7-amplified

GEP of CEM and A2780 cells to their original, unamplified profiles. We found that some genes expressed at a high or medium level in the unamplified GEP showed decreased expression levels ("under-amplified genes") or completely disappeared from the amplified GEP of the same cell type ("dropped genes"). This inefficient amplification affected about 20% of the GEP even after 2 rounds of amplification and the average signal level did not drop significantly (as opposed to exponential amplification). Furthermore, about 5% of expressed genes were found to be over-amplified. We found that T7-amplification results were highly reproducible with R^2-values of 0.98 between triplicate samples after one round that is similar to triplicates of unamplified samples. After two rounds of T7-amplification the reproducibility seemed to be even higher approaching R^2-values of 0.99. When we compared the GEPs of an amplified sample to the GEP of the same sample without amplification we found significant GEP distortion with R^2-values of 0.609-0.73 after one round and 0.50 after two rounds of linear amplification. However, as opposed to exponential RNA amplification, linear amplification did not diminish the original overall GEP difference between the two cell types. The original value of R^2=0.80 between the unamplified A2780 and CEM cells remained almost exactly the same with R^2=0.79 after one round and R^2=0.77 after two rounds of T7-amplification. This suggested that, unlike exponential RNA amplification that created a "universal amplified profile" regardless of the original, unamplified GEPs, linear RNA amplification preserved the differences of overall GEPs between samples despite the GEP distortion it introduced. While after exponential amplification all GEPs became virtually indistinguishable, after T7-amplification the individual GEPs remained identifiable. After evaluating the possible causes of GEP distortion introduced by linear RNA amplification we concluded that it was not caused by sample handling problems or other protocol related difficulties and it was not only greatly independent from the initial expression levels of the individual mRNA species, but also from the entire original GEP. We concluded that the most probable source of distortion in linear amplification processes was the amplifying enzyme that amplified some genes differentially in a sequence specific manner. The introduced distortion affected a much smaller portion of

the GEP than with exponential amplification. Future studies will examine the possible sequence specific nature of RNA amplification related GEP distortion.

Based on all our conclusions we found the linear amplification methods much more suitable to analyze the LEAP purified and optoinjected samples. After one round of linear RNA amplification prior to microarray analysis we found that LEAP purification did not distort the GEP (R^2=0.97) of the purified cells. On the other hand, optoinjection introduced a significant GEP distortion with or without dextrans in the medium (R^2=0.93 and R^2=0.91 respectively). Additionally, the mock-optoinjected GEP also differed significantly from the dextran-optoinjected GEP of the same cells (R^2=0.91). Further analysis of the optoinjection effects including the genes and pathways affected, is beyond the scope of this study.

Finally, based on our results we demonstrated that some of the distortion introduced into the GEP of an unknown sample by linear amplification could be eliminated by applying the distortion factor of another (known) sample. We hypothesize that different amplifying enzymes distort different portions of the GEP. Complementary use of different linear amplification methods may increase the reconstructable portion of the amplified GEP close to 100%. This could mean that in the near future microgenomics could provide equally accurate GEP analysis from small samples or even from a single cell to traditional, large-sample microarray analysis.

CHAPTER 7.
CONCLUSIONS AND OVERALL DISCUSSION

In conclusion, our data indicate that to obtain meaningful results from gene expression microarray analysis of biological samples containing mixed cell populations it is necessary to purify the studied cell subset. We modeled biological samples using cell mixtures of two cell types, CEM cells and A2780 cells with characteristically different overall gene expression profiles (GEPs). Using well-defined mixtures with different cell ratios of CEM and A2780 cells we demonstrated that the overall GEPs of mixed cell populations are, as expected, the combined expression profiles for each cell subpopulation weighted according to its relative frequency in the cell mixture. We found this to be true for both major types of microarrays - spotted arrays, and short oligonucleotide arrays - each of which generated highly reproducible results with a 3 to 4 log dynamic range. We showed that without applying any cell separation the cell type in majority dominated the overall GEP of the sample while the GEPs of minor cell subsets got washed out. We also demonstrated that alterations in the GEP of a minor cell subset could only be studied for those genes that were not expressed by the contaminating cell type. We found that, in our model, the overall GEP of a more than 75% pure sample was practically indistinguishable from a 100% pure sample. Based on this result we determined that the functional threshold for the necessary purity of a cell type in a sample to produce virtually identical overall GEP to a pure sample was 75%. We noted that the presence of a few outlier genes indicated that purity requirements could be very different for monitoring individual genes, depending on whether or not those same genes are expressed at high or low levels in the contaminating cell types. While analyzing GEPs of model cell mixtures we found that “housekeeping genes” that are often used to standardize samples in Northern blots and other technologies studying gene expression, were differentially expressed between cell samples. This result warned against using any single “housekeeping gene” as standard in these experiments.

Based on our results for the purity requirements of a sample for meaningful microarray analysis we developed a method for sample purification that did not distort the overall GEP of the purified cell subset. We demonstrated that after antibody labeling, methanol fixation and cell sorting by flow cytometry or magnetic bead based cell separation technologies, the overall GEP of the purified cell subset remained unaltered and even omitting steps traditionally used to improve RNA quality did not have any significant effect on the overall GEP. Again, the presence of a few outlier genes indicated that individual genes might be much more affected by certain processing steps, e.g. antibody labeling of a surface receptor on a live cell obviously, might trigger certain pathways altering the expression levels of the genes involved. Nevertheless, we were able to conclude that the overall GEP of a sample (representing the vast majority of all genes) is more robust and resistant to sample processing than it has been generally appreciated. We demonstrated that starting with a 10-50% cell mixture greater than 75% purity could be achieved by two rounds of magnetic bead sorting. One round typically resulted in about 70% purity (just below the 75% threshold), two rounds raised the purity to approximately 90%, while after 3 rounds it was generally above 95% even from samples in which the purified cell subset was originally 1% or less (e.g. human stem/progenitor cell subsets in cord blood). One round of magnetic bead sorting followed by one round of flow cytometry/cell sorting resulted in about 90-98% purity as well.

To test just how much of the hidden profile of a minority cell subset could be revealed by cell purification, we purified the minority cell subset from a model CEM/A2780 cell mixture containing only 10% of CEM cells where the GEP of these cells was almost completely covered by the background cells. Using both magnetic bead cell purification and flow cytometric cell sorting we managed to recover the ‘hidden’ GEP virtually perfectly, also proving that the sort process itself did not distort the profile. As a proof-of-principle experiment, we analyzed the GEP of purified, CD34+ cord blood stem/progenitor cells. Since these cells are present in cord blood in less than 1% minority of all mononuclear cells, their GEP had been heavily masked by the overwhelming presence of mature contaminating cell types. We showed that the recovered GEP of these

cells was characteristically different from both CBMCs and KG-1a cells. We concluded that for meaningful gene expression microarray profiling a minor cell subset of a cell mixture, purification of these cells was not only necessary but also very much achievable.

To be able to purify cell subsets from small, biohazardous samples where flow cytometry and column based immuno-magnetic cell sorting were not feasible methods, we studied the capabilities of multistage magnetic sorting (Magsort), a new, closed-system cell sorting technology that separated magnetic particles and magnetically labeled cells based on their magnetophoretic mobility. After successfully separating two different types of magnetic microparticles to 75-98% purity from approximately even mixtures we used model cell mixtures – similarly to those described in the first section - to test the cell sorting capabilities and limitations of this technology. We tested Magsort using several different cell types, antibodies, magnetic bead types, loading methods and sort programs. Out of these variables the magnetic bead type and the loading method used proved to be the most important factors that determined the outcome of the sort experiments. We found that using the original protocols Magsort was not able to selectively pull magnetically labeled cells from the bottom of the sample chamber into the sort chamber. The vast majority (88-98%) of all loaded cells remained in the sample insertion chamber. The few cells that did end up in the sort chambers were not delivered there selectively based on their immuno-magnetic labeling, because the cell mixture ratio of these fractions remained unaltered compared to the initial sample. This was invariably the case with all cell types, and magnets tested using any user definable sort program. We concluded that the cells that settled on the bottom of the sample insertion chamber were too far from the magnets to get sorted.

To shorten the distance between the cells and the sorting magnets and also reduce the necessary force the cells needed to start moving, we layered Percoll underneath the sample to bring the cells closer to the magnets and to create a surface where the cells can be pulled off from easier than from the plastic surface of the bottom of the sample chamber. The overall effect of this approach depended on the magnetic bead type used for labeling. Cells labeled with Dynal beads (1 μm diameter) or with Bangs beads (0.8-1

μm diameter) selectively moved up into the sort chambers in the increasing magnetic field. Miltenyi beads with their 50 nm diameter were too small to lift the cells up even under the modified conditions. Dynal beads with 4.5 μm diameter formed large, multicell/multibead aggregates during sorting that prevented individual cell sorting. We demonstrated that using Bangs beads and optimized Percoll and sample loading protocols Magsort was capable of sorting up to 4 distinguishable fractions with different labeled/unlabeled cell ratios of immuno-magnetically labeled cells. The maximum sample purity we could achieve with this method was 75-80% which was slightly over the purity requirement threshold we established for microarray experiments.

To sort cell subsets from small and/or biohazardous samples with higher accuracy we evaluated Laser Enabled Analysis and Processing (LEAP), a new scanning cytometry technology that applies laser energy to ablate or catapult unwanted cells to achieve sample purification. We used model cell lines and cell mixtures – similarly to those described earlier - to test the cell sorting capabilities and limitations of this technology. We found that unwanted cells could be removed from attached cell cultures by laser ablation with almost "surgical" accuracy causing minimal damage to neighboring cells within 2-3 cell diameters. As a proof of principle experiment we demonstrated that LEAP provided the unique ability to purify primary hepatocytes that were extremely sensitive to conventional cell sorting methods, without causing any visible damage to these fragile cells. To purify suspension cells by LEAP, we found that cells could be gently centrifuged to the slide/well surface and afterwards they remained in a semi-attached state with no additional attachment material needed and without any apparent damage to the cells. For suspension cells we developed a purification method of bouncing off contaminating cells using sub-lethal laser power rather than ablating them. With this approach, after LEAP targeting and washing away the floating unwanted cells, we achieved much higher purity (above 90% in most applications) in the recovered cell population than when we were aiming for ablation of the contaminating cells alone. Furthermore, with a second round of LEAP-purification virtually 100% purity could be achieved when this was necessary. We established optimal cell plating densities needed

for a given experiment based on the ratio of contaminating cells, the required end-purity, and the affordable cell loss. We showed that when recovery of most purified cells was an important issue as from a small sample, it was possible to achieve above 95% purity and above 80% cell recovery with LEAP even from a 50% cell mixture. This required low density cell plating that ultimately resulted in slower cell processing, but with small cell samples this was usually not an issue since the process still only took a few minutes after the initial setup. Large cell samples, where 30-50% cell loss was not a problem, could be plated at higher cell densities significantly increasing the processing speed and still maintaining above 90% purity. We observed no apparent damage to the purified cells with either method. We have also shown that the combination of one round magnetic bead –pre-sorting from a large, 50% cell mix yielded a 75%-80% pure sample and a second round, high-density LEAP-purification from this sample resulted in 90% purity or better, similarly to 2 rounds of LEAP-purification or 2 rounds of magnetic bead sorting. The combined approach was much faster than handling a large sample with 2 rounds of LEAP purification alone and resulted in much less cell loss than 2 rounds of magnetic bead sorting alone. Since both methods can be carried out in a closed environment this combination method might be the best way to purify large, biohazard samples to high purities prior to microarray analysis.

We also studied LEAP-mediated optoinjection, a novel tool for targeted macromolecule delivery where laser targeted cells take up macromolecules from the culture medium. We examined optoinjection effects by adding fluorescent dextrans of different sizes as deliverable macromolecules into the medium before LEAP-shooting to visualize optoinjection effects. We developed optimized LEAP-shooting conditions for each cell type and application maximizing optoinjection effects without causing any visible damage to the targeted cells. We found that optoinjection worked with all cell types we studied (adherent and suspension cells) with no exception and it could be optimized to achieve literally 100% optoinjection of the targeted cells. We found that adherent cells were easier to manipulate with LEAP than suspension cells since it did not require extra effort to keep them in the focal plane, however the optoinjection effect

appeared to be more diffuse on adherent cells than on suspension cells. Gently centrifuged suspension cells could be optoinjected similarly to attached cells; the laser energy required for optoinjection did not remove the cells from the slide surface. The level of fluorescence we could achieve with these cells was lower than with adherent cells. Confocal microscopy results confirmed that the fluorescent dextran molecules were indeed inside the cells after optoinjection. Most of the optoinjected molecules were found in the cells' cytoplasm with the nuclei remaining relatively negative, although also visibly optoinjected. This analysis also confirmed that optoinjection did not cause any apparent alteration in cellular morphology. The maximum size of optoinjectable macromolecules and the efficacy of optoinjection for macromolecules with different size and chemical structure still need to be determined. We observed very similar results with 3kD-40kD dextrans and significantly reduced (~50%) fluorescence of optoinjected cells when using 70kD dextrans. These observations might mean that the limits of the underlying mechanism are not much above the size of a 70kD dextran molecule. During cell ablation experiments we found an indirect optoinjection effect at the edges of the ablation zone - both in confluent adherent cultures and with suspension cells - that might be a sign of altered membrane and other functions in the purified cells. However, these effects are likely to be transient, lasting for a few seconds only, since these cells did not show any signs of "leakiness" or morphological damage a few minutes after the indirect optoinjection as determined by confocal, brightfield, and fluorescent microscopy. We found much less indirect optoinjection effect (1-4 cell diameter wide zone) with suspension cells than with adherent cells (4-6 cell diameter wide zone).

Based on our results, we developed a theory to explain the underlying mechanism of optoinjection, an unexplained phenomenon to date. According to our theory, laser energy is absorbed by different molecules in the medium around the cells, in the cell membrane, and inside the cells warming up the cells in the targeted area. As a result of warming up beyond a certain threshold, the cell membrane goes through a phase change becoming more liquid-like than gel-like. The uneven warming and the low energy shockwaves caused by the pulsing laser generate waves in the fluid cell membrane. These

waves may result in opening transient holes in the cell membrane and/or in transiently opening and enlarging existing pores and channels. If the sum laser energy is large enough to create membrane ruptures beyond the cells healing capabilities or even cause the cytoplasm to explode (increase its volume beyond the membranes flexibility) the end result will be ablation or permanent damage. If the sum energy is not enough for the above-described effects, but enough to cause transient membrane disturbances, the end result will be sum energy dependent optoinjection. If the sum absorbed energy stays below a certain threshold the cell membrane will remain intact and no optoinjection will occur. Further experiments will be needed to test this theory and to elucidate the underlying mechanism(s) for optoinjection.

Based on our results we concluded that for cell sorting LEAP might become especially useful in areas where other sorting methods are seriously challenged as in purifying large numbers of live, adherent cells; very small samples of live, suspension cells; live cells that are highly sensitive to traditional processing; and safe, live processing of biohazardous cells. However, the amount of RNA we could isolate from LEAP-purified and LEAP-optoinjected samples was not sufficient for direct microarray analysis. To analyze the GEPs of these relatively small samples, RNA amplification was necessary.

Our earlier conclusions, that for meaningful microarray data a reasonable purification of the target cell subpopulation is both necessary and achievable, lead to the problem that real biological samples, especially after cell purification, often do not provide the necessary amounts of RNA, for direct gene expression microarray analysis. Furthermore, in some cases – purified rare cell subsets (like stem-progenitor blood cells), aspiration biopsies, micromanipulated cells (like LEAP-optoinjected individual cells) – GEP analysis of a few or even single cells would be desirable. To be able to analyze the GEP of small, purified cell samples we studied the capabilities and limitations of exponential and linear RNA amplification methods. Since for meaningful GEP analysis it is not enough to obtain the required amount of RNA or cDNA, it is also extremely important to keep the relative frequencies of each mRNA species at the original ratios

throughout the amplification process, we analyzed the amplified GEPs of model cell lines and compared them to the GEPs of unamplified samples and to each other to investigate the extent of possible GEP distortions. We used two different cell lines, A2780 and CEM cells, with characteristically different GEPs to evaluate how accurately the two methods preserved the GEP differences between the two cell types. We established that the unamplified GEPs of the two cell lines were significantly different with R^2-value of 0.808 when plotting them against each other.

When testing our selected exponential RNA amplification method that uses Taq-polymerase as amplifier enzyme, we found that from 10^4 cells 18 PCR-cycles generated equivalent amounts of cDNA to an unamplified, 10^6 cell sample. From 10^3 cells 25 PCR cycles were necessary. We found this technology relatively easy to handle and very robust - it produced the required microgram amounts of cDNA within a few hours. When we analyzed if there was any significant GEP distortion introduced by this method we found that several genes expressed at a high or medium level in the unamplified GEP disappeared from the amplified GEP of the same cell type ("dropped genes"). Other genes that were expressed at a low level or were undetectable in the unamplified profile appeared to be expressed at a high or medium level in the amplified profile (over-amplified genes). Ratios of differentially expressed genes were often found to be reversed when comparing unamplified and amplified GEPs of the same cell type. Comparing the GEPs of an amplified sample to the GEP of the same sample without amplification showed significant GEP distortion with R^2-values of 0.60-0.76 after 18 cycles and 0.40-0.71 after 25 cycles. Furthermore, the original significant differences between the unamplified GEPs of A2780 and CEM cells were almost entirely diminished after PCR amplification. The R^2-value increased to 0.944 between CEM and A2780 cells. Finally, after 18 cycles the average relative expression level of the expressed genes dropped to one-fifth to one-seventh of the unamplified level. After 25 cycles this tendency continued (reaching a 20 to 30-fold drop) to the point where only a few genes showed significant expression values above background. Based on these results of the exponential RNA amplification experiments, we concluded that this method distorted the overall GEP

beyond recognition, at the same time greatly diminished GEP differences between samples and resulted in a huge loss of signal over background ratio. For these reasons we concluded that while PCR might be a good method for genome-wide cDNA library construction where the relative ratios of the individual cDNA species are not important it was not suitable for gene expression microarray profiling.

When we tested a linear RNA amplification method that uses T7 RNA-polymerase for amplification, we found that from 10^3 cells or more, one round of T7 polymerase-based, linear RNA amplification (T7-amplification) provided enough RNA for microarray analysis. Smaller samples required two rounds to produce equivalent amounts of RNA to an unamplified sample of 10^6 cells confirming the notion that one round of T7 amplification results in approximately 1000-fold RNA increase. We demonstrated that 2 rounds of T7-amplification were enough for even a single cell microarray. We found this technology relatively difficult to handle and time consuming since two rounds of amplification normally took 3 days to produce microarray-ready RNA. To evaluate if there was any significant GEP distortion introduced by this method, we compared the T7-amplified GEP of CEM and A2780 cells to their original, unamplified profiles. We found that some genes expressed at a high or medium level in the unamplified GEP showed decreased expression levels ("under-amplified genes") or completely disappeared from the amplified GEP of the same cell type ("dropped genes"). This inefficient amplification affected only 10-20% of the GEP even after 2 rounds of amplification and the average signal level did not drop significantly (as opposed to exponential amplification). Furthermore, hardly any genes were found to be over-amplified. We found that T7-amplification results were highly reproducible with R^2-values of 0.98 between triplicate samples after one round that is similar to triplicates of unamplified samples. After two rounds of T7-amplification the reproducibility seemed to be even higher approaching R^2-values of 0.99. When we compared the GEPs of an amplified sample to the GEP of the same sample without amplification we found significant GEP distortion with R^2-values of 0.6-0.75 after one round and 0.5 after two rounds of linear amplification. However, as opposed to exponential RNA amplification,

linear amplification did not diminish the original overall GEP difference between the two cell types. The original value of R^2=0.80 between the unamplified A2780 and CEM cells remained almost exactly the same with R^2=0.79 after one round and R^2=0.77 after two rounds of T7-amplification. This suggested that, unlike exponential RNA amplification that created a "universal amplified profile" regardless of the original, unamplified GEPs, linear RNA amplification preserved the differences of overall GEPs between samples despite the GEP distortion it introduced. While after exponential amplification all GEPs became virtually indistinguishable, after T7-amplification the individual GEPs remained identifiable. After evaluating the possible causes of GEP distortion introduced by linear RNA amplification we concluded that it was not caused by sample handling problems or other protocol related difficulties and it was not only greatly independent from the initial expression levels of the individual mRNA species, but also from the entire original GEP. We concluded that the most probable source of distortion in linear amplification processes was the amplifying enzyme that amplified some genes differentially in a sequence specific manner. The introduced distortion affected a much smaller portion of the GEP than with exponential amplification. Further studies will be needed to examine the possible sequence specific nature of RNA amplification related GEP distortion.

Based on all our conclusions we found the linear amplification methods much more suitable to analyze the LEAP purified and optoinjected samples. After one round of linear RNA amplification prior to microarray analysis we found that LEAP purification did not distort the GEPs (R^2=0.97) of the purified cells. On the other hand, optoinjection introduced a significant GEP distortion with or without dextrans in the medium (R^2=0.93 and R^2=0.91 respectively). Additionally, the mock-optoinjected GEP also differed significantly from the dextran-optoinjected GEP of the same cells (R^2=0.91). Further analysis of the optoinjection effects including the genes and pathways affected, will be necessary.

Finally, we demonstrated that parts of the unamplified GEP could be predicted from the distorted, amplified GEP. Based on our results we hypothesize that complementary use of different linear amplification methods may increase the

reconstructable portion of the amplified GEP close to 100%. This could mean that in the near future microgenomics could provide equally accurate GEP analysis of small samples or even of a single cell to traditional, large-sample microarray analysis.

REFERENCES

1. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001;409(6822):860-921.

2. Venter JC, et al. THE HUMAN GENOME: Science Genome Map. Science 2001;291(5507):1304-1351.

3. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF and others. The Genome Sequence of Drosophila melanogaster. Science 2000;287(5461):2185-2195.

4. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM and others. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 1995;269(5223):496-512.

5. The C. elegans Sequencing Consortium. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science 1998;282(5396):2012-2018.

6. Howbrook DN, van der Valk AM, O'Shaughnessy MC, Sarker DK, Baker SC, Lloyd AW. Developments in microarray technologies. Drug Discovery Today. 2003;8(14):642-651.

7. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray.[see comment]. Science. 1995;270(5235):467-470.

8. Chaussabel D, Sher A. Mining microarray expression data by literature profiling. Genome Biology. 2002;3(10):55.1-55.16.

9. Shoemaker DD, Linsley PS. Recent developments in DNA microarrays. Current Opinion in Microbiology. 2002;5(3):334-337.

10. Lander ES. Array of hope. Nature Genetics. 1999;21(1 Suppl):3-4.

11. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. Nature Genetics. 1999;21(1 Suppl):33-37.

12. Bowtell DD. Options available--from start to finish--for obtaining expression data by microarray.[erratum appears in Nat Genet 1999 Feb;21(2):241]. Nature Genetics. 1999;21(1 Suppl):25-32.

13. Holloway AJ, van Laar RK, Tothill RW, Bowtell DD. Options available--from start to finish--for obtaining data from DNA microarrays II. Nature Genetics. 2002;32(Suppl):481-489.

14. Leung YF, Cavalieri D. Fundamentals of cDNA microarray data analysis. Trends in Genetics. 2003;19(11):649-659.

15. Copland JA, Davies PJ, Shipley GL, Wood CG, Luxon BA, Urban RJ. The use of DNA microarrays to assess clinical samples: the transition from bedside to bench to bedside. Recent Progress in Hormone Research. 2003;58:25-53.

16. Petricoin EF, 3rd, Hackett JL, Lesko LJ, Puri RK, Gutman SI, Chumakov K, Woodcock J, Feigal DW, Jr., Zoon KC, Sistare FD. Medical applications of microarray technologies: a regulatory science perspective. Nature Genetics. 2002;32(Suppl):474-479.

17. Shaughnessy J, Jr. Primer on medical genomics. Part IX: scientific and clinical applications of DNA microarrays--multiple myeloma as a disease model. Mayo Clinic Proceedings. 2003;78(9):1098-1109.

18. Nakeff A, Sahay N, Pisano M, Subramanian B. Painting with a molecular brush: genomic/proteomic interfacing to define the drug action profile of novel solid-tumor selective anticancer agents. Cytometry. 2002;47(1):72-79.

19. Los G, Yang F, Samimi G, Manorek G, Guerorguieva IM, Howell S, van Erp N, Breaux JK. Using mRNA expression profiling to determine anticancer drug efficacy. Cytometry. 2002;47(1):66-71.

20. Staunton JE, Slonim DK, Coller HA, Tamayo P, Angelo MJ, Park J, Scherf U, Lee JK, Reinhold WO, Weinstein JN and others. Chemosensitivity prediction by transcriptional profiling. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(19):10787-1092.

21. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA and others. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-752.

22. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A and others. Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature. 2000;406(6795):536-540.

23. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT and others. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002;415(6871):530-536.

24. Abdullah-Sayani A, Bueno-de-Mesquita JM, van de Vijver MJ. Technology Insight: tuning into the genetic orchestra using microarrays--limitations of DNA microarrays in clinical practice. Nature Clinical Practice Oncology 2006;3(9):501-516.

25. Chen C-N, Lin J-J, Chen JJW, Lee P-H, Yang C-Y, Kuo M-L, Chang K-J, Hsieh F-J. Gene expression profile predicts patient survival of gastric cancer after surgical resection. Journal of Clinical Oncology 2005;23(29):7286-7295.

26. Chen Y, Knosel T, Kristiansen G, Pietas A, Garber ME, Matsuhashi S, Ozaki I, Petersen I. Loss of PDCD4 expression in human lung cancer correlates with tumour progression and prognosis. Journal of Pathology. 2003;200(5):640-646.

27. Murphy N, Millar E, Lee CS. Gene expression profiling in breast cancer: towards individualising patient management. Pathology 2005;37(4):271-277.

28. Jones J, Libermann TA. Genomics of renal cell cancer: the biology behind and the therapy ahead. Clinical Cancer Research 2007;13(2 Pt 2):685s-692s.

29. Simon R, Radmacher MD, Dobbin K. Design of studies using DNA microarrays. Genetic Epidemiology 2002;23(1):21-36.

30. Tan PK, Downey TJ, Spitznagel EL, Jr., Xu P, Fu D, Dimitrov DS, Lempicki RA, Raaka BM, Cam MC. Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Research 2003;31(19):5676-5684.

31. Spruill SE, Lu J, Hardy S, Weir B. Assessing sources of variability in microarray gene expression data. Biotechniques. 2002;33(4):916-920.

32. Mirnics K. Microarrays in brain research: the good, the bad and the ugly. Nature Reviews Neuroscience. 2001;2(6):444-447.

33. Kuo WP, Jenssen T-K, Butte AJ, Ohno-Machado L, Kohane IS. Analysis of matched mRNA measurements from two different microarray technologies. Bioinformatics 2002;18(3):405-412.

34. Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. Nature Methods 2005;2(5):337-344.

35. Knight J. When the chips are down. Nature. 2001;410(6831):860-861.

36. Rogojina AT, Orr WE, Song BK, Geisert EE, Jr. Comparing the use of Affymetrix to spotted oligonucleotide microarrays using two retinal pigment epithelium cell lines. Molecular Vision. 2003;9:482-496.

37. Harrington CA, Rosenow C, Retief J. Monitoring gene expression using DNA microarrays. Current Opinion in Microbiology 2000;3(3):285-291.

38. Newton SS, Bennett A, Duman RS. Production of custom microarrays for neuroscience research. Methods 2005;37(3):238-246.

39. Pollock JD. Gene expression profiling: methodological challenges, results, and prospects for addiction research. Chemistry & Physics of Lipids 2002;121(1-2):241-256.

40. Chu TT, Fink MY, Mong JA, John G, Auger AP, Ge Y, Sealfon SC. Effective use of microarrays in neuroendocrine research. Journal of Neuroendocrinology 2007;19(3):145-161.

41. Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, Frank BC, Gabrielson E, Garcia JGN, Geoghegan J, Germino G and others. Multiple-laboratory comparison of microarray platforms. Nature Methods 2005;2(5):345-350.

42. Weis BK, et al. Standardizing global gene expression analysis between laboratories and across platforms. Nature Methods 2005;2(5):351-356.

43. Packeisen J, Korsching E, Herbst H, Boecker W, Buerger H. Demystified...tissue microarray technology. Molecular Pathology. 2003;56(4):198-204.

44. Hartmann CH, Klein CA. Gene expression profiling of single cells on large-scale oligonucleotide arrays. Nucleic Acids Research 2006;34(21):e143.

45. Kurimoto K, Yabuta Y, Ohinata Y, Ono Y, Uno KD, Yamada RG, Ueda HR, Saitou M. An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Nucleic Acids Research 2006;34(5):e42.

46. Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JC, Trent JM, Staudt LM, Hudson J, Jr., Boguski MS and others. The transcriptional program in the response of human fibroblasts to serum. Science. 1999;283(5398):83-87.

47. Ntzani EE, Ioannidis JP. Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment. Lancet. 2003;362(9394):1439-1444.

48. Russo G, Zegar C, Giordano A. Advantages and limitations of microarray technology in human cancer. Oncogene. 2003;22(42):6497-6507.

49. Rhodes DR, Chinnaiyan AM. Integrative analysis of the cancer transcriptome. Nature Genetics 2005;37:s31-s37.

50. Segal E, Friedman N, Kaminski N, Regev A, Koller D. From signatures to models: understanding cancer using microarrays. Nature Genetics 2005;37:s38-s45.

51. Boess F, Kamber M, Romer S, Gasser R, Muller D, Albertini S, Suter L. Gene expression in two hepatic cell lines, cultured primary hepatocytes, and liver slices compared to the in vivo liver gene expression in rats: possible implications for toxicogenomics use of in vitro systems. Toxicological Sciences. 2003;73(2):386-402.

52. Kaminski N, Allard JD, Pittet JF, Zuo F, Griffiths MJ, Morris D, Huang X, Sheppard D, Heller RA. Global analysis of gene expression in pulmonary fibrosis reveals distinct programs regulating lung inflammation and fibrosis. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(4):1778-1783.

53. Lock C, Hermans G, Pedotti R, Brendolan A, Schadt E, Garren H, Langer-Gould A, Strober S, Cannella B, Allard J and others. Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. Nature Medicine. 2002;8(5):500-508.

54. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA, Brown PO. Individuality and variation in gene expression patterns in human blood. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(4):1896-1901.

55. Lee S, Zhou G, Clark T, Chen J, Rowley JD, Wang SM. The pattern of gene expression in human CD15+ myeloid progenitor cells. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(6):3340-3345.

56. Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, Lemischka IR. A stem cell molecular signature. Science. 2002;298(5593):601-604.

57. Ramalho-Santos M, Yoon S, Matsuzaki Y, Mulligan RC, Melton DA. "Stemness": transcriptional profiling of embryonic and adult stem cells. Science. 2002;298(5593):597-600.

58. Auer H, Lyianarachchi S, Newsom D, Klisovic MI, Marcucci, Kornacker K. Chipping away at the chip bias: RNA degradation in microarray analysis. Nature Genetics. 2003;35(4):292-293.

59. Schuchhardt J, Beule D, Malik A, Wolski E, Eickhoff H, Lehrach H, Herzel H. Normalization strategies for cDNA microarrays. Nucleic Acids Research. 2000;28(10):E47.

60. Goldsworthy SM, Stockton PS, Trempus CS, Foley JF, Maronpot RR. Effects of fixation on RNA extraction and amplification from laser capture microdissected tissue. Molecular Carcinogenesis. 1999;25(2):86-91.

61. Su JMF, Perlaky L, Li X-N, Leung H-CE, Antalffy B, Armstrong D, Lau CC. Comparison of ethanol versus formalin fixation on preservation of histology and RNA in laser capture microdissected brain tissues. Brain Pathology 2004;14(2):175-182.

62. Miltenyi S, Muller W, Weichel W, Radbruch A. High gradient magnetic cell separation with MACS. Cytometry 1990;11(2):231-238.

63. Szaniszlo P, Wang N, Sinha M, Reece LM, Van Hook JW, Luxon BA, Leary JF. Getting the right cells to the array: Gene expression microarray analysis of cell mixtures and sorted cells. Cytometry. 59A 2004;2:191-202.

64. Chalmers JJ, Zborowski M, Sun L, Moore L. Flow through, immunomagnetic cell separation. Biotechnology Progress 1998;14(1):141-148.

65. Cooper RP, Doyle J. F., Dunn D. S. , Vellinger J. C. and Todd P. Multistage magnetic particle separator 2. Classification of ferromagnetic particles. Separation Science and Technology 2004;39:2809-2825.

66. Todd PC, R. P.; Doyle, J. F.; Dunn, S.; Vellinger, J.; Deuser, M. S. Multistage Magnetic Particle Separator. Journal of Magnetism and Magnetic Materials 2001(225):294-300.

67. Vellinger JC, Todd, P. W., Barton, K., Dunn, D. S., and Deuser, M. S.; Multistage Electromagnetic Separator for Purifying Cells, Chemicals and Protein Structures. U. S. Patent. 2004.

68. Vellinger JT, P.; Barton, K.; Dunn, S.; Deuser, M. S.; Multistage Electromagnetic Separator for Purifying Cells, Chemicals and Protein Structures. U. S. Patent. 2001.

69. Kamentsky LA. Laser scanning cytometry. Methods in Cell Biology 2001;63:51-87.

70. Luther E, Kamentsky, L, Henriksen, M, Holden, E. Next Generation Laser Scanning Cytometry. In: Darzynkiewicz Z, Roederer, M, Tanke, H, editor. Methods in Cell Biology. 2004;75:185-218.

71. Kamentsky LA, Kamentsky LD. Microscope-based multiparameter laser scanning cytometer yielding data comparable to flow cytometry data. Cytometry 1991;12(5):381-387.

72. Szaniszlo P, Rose WA, Wang N, Reece LM, Tsulaia TV, Hanania EG, Elferink CJ, Leary JF. Scanning cytometry with a LEAP: laser-enabled analysis and

processing of live cells in situ. Cytometry Part A: The Journal of the International Society for Analytical Cytology 2006;69(7):641-651.

73. Price J, Gough, DA; US Patent No, 5,548,661, assignee. Operator Independent Image Cytometer. U.S. Patent. 1996.

74. Price J, Gough, DA; US Patent No. 5,790,710, assignee. Autofocus System for Scanning Microscopy. U.S. Patent 1998.

75. Abraham VC, Taylor DL, Haskins JR. High content screening applied to large-scale cell biology. Trends in Biotechnology 2004;22(1):15-22.

76. Bedner E, Burfeind P, Gorczyca W, Melamed MR, Darzynkiewicz Z. Laser scanning cytometry distinguishes lymphocytes, monocytes, and granulocytes by differences in their chromatin structure. Cytometry 1997;29(3):191-196.

77. Bedner E, Ruan Q, Chen S, Kamentsky LA, Darzynkiewicz Z. Multiparameter analysis of progeny of individual cells by laser scanning cytometry. Cytometry 2000;40(4):271-279.

78. Clatch RJ, Walloch JL, Zutter MM, Kamentsky LA. Immunophenotypic analysis of hematologic malignancy by laser scanning cytometry. American Journal of Clinical Pathology 1996;105(6):744-755.

79. Clatch RJ, Foreman JR, Walloch JL. Simplified immunophenotypic analysis by laser scanning cytometry. Cytometry 1998;34(1):3-16.

80. Gerstner A, Laffers W, Bootz F, Tarnok A. Immunophenotyping of peripheral blood leukocytes by laser scanning cytometry. Journal of Immunological Methods 2000;246(1-2):175-185.

81. Gerstner AO, Lenz D, Laffers W, Hoffman RA, Steinbrecher M, Bootz F, Tarnok A. Near-infrared dyes for six-color immunophenotyping by laser scanning cytometry. Cytometry 2002;48(3):115-123.

82. Gorczyca W, Darzynkiewicz Z, Melamed MR. Laser scanning cytometry in pathology of solid tumors. A review. Acta Cytologica 1997;41(1):98-108.

83. Gorczyca W, Deptala A, Bedner E, Li X, Melamed MR, Darzynkiewicz Z. Analysis of human tumors by laser scanning cytometry. Methods in Cell Biology 2001;64:421-443.

84. Lenz D, Gerstner, AO, Laffers, W, SteinBrecher, M, Bootz, F, Tarnot, A, editor. Six and more color immunophenotyping on the slide by laser scanning cytometry (LSC). Society of Photographic Instrumentation Engineers Proceedings 2003;4962:364-374.

85. Tarnok A, Gerstner AO. Clinical applications of laser scanning cytometry. Cytometry 2002;50(3):133-143.

86. Bonner RF, Emmert-Buck M, Cole K, Pohida T, Chuaqui R, Goldstein S, Liotta LA. Laser capture microdissection: molecular analysis of tissue. Science 1997;278(5342):1481-1483.

87. Buican T, Neagley, DL., Morrison, WC., Upham, BD. Optical trapping, cell manipulation, and robotics. Proceedings of the Society of Photographic Instrumentation Engineers 1989;1063:190.

88. Clark IB, Hanania EG, Stevens J, Gallina M, Fieck A, Brandes R, Palsson BO, Koller MR. Optoinjection for efficient targeted delivery of a broad range of compounds and macromolecules into diverse cell types. Journal of Biomedical Optics 2006;11(1):14-34.

89. Koller M, Hanania, EG., Eisfeld, TM., Palsson, BO.; Optoinjection methods. U.S. Patent. 2004.

90. Koller MR, Hanania EG, Stevens J, Eisfeld TM, Sasaki GC, Fieck A, Palsson BO. High-throughput laser-mediated in situ cell purification with high purity and yield. Cytometry 2004;61A(2):153-161.

91. Palsson B, Koller M, Eisfeld T; Method and apparatus for selectively targeting specific cells within a cell population. U.S. Patent. 2003.

92. Luo L, Salunga RC, Guo H, Bittner A, Joy KC, Galindo JE, Xiao H, Rogers KE, Wan JS, Jackson MR and others. Gene expression profiles of laser-captured adjacent neuronal subtypes.[erratum appears in Nature Medicine 1999 Mar;5(3):355]. Nature Medicine 1999;5(1):117-122.

93. Ginsberg SD. RNA amplification strategies for small sample populations. Methods 2005;37(3):229-237.

94. Hemby SE, Ginsberg SD, Brunk B, Arnold SE, Trojanowski JQ, Eberwine JH. Gene expression profile for schizophrenia: discrete neuron transcription patterns in the entorhinal cortex. Archives of General Psychiatry 2002;59(7):631-640.

95. Bonaventure P, Guo H, Tian B, Liu X, Bittner A, Roland B, Salunga R, Ma X-J, Kamme F, Meurers B and others. Nuclei and subnuclei gene expression profiling in mammalian brain. Brain Research 2002;943(1):38-47.

96. McClain KL, Cai YH, Hicks J, Peterson LE, Yan XT, Che S, Ginsberg SD. Expression profiling using human tissues in combination with RNA amplification and microarray analysis: assessment of Langerhans cell histiocytosis. Amino Acids 2005;28(3):279-290.

97. Freeman TC, Lee K, Richardson PJ. Analysis of gene expression in single cells. Current Opinion in Biotechnology 1999;10(6):579-582.

98. Korobkova E, Emonet T, Vilar JMG, Shimizu TS, Cluzel P. From molecular noise to behavioural variability in a single bacterium. Nature 2004;428(6982):574-578.

99. Bahar R, Hartmann CH, Rodriguez KA, Denny AD, Busuttil RA, Dolle MET, Calder RB, Chisholm GB, Pollock BH, Klein CA and others. Increased cell-to-cell variation in gene expression in ageing mouse heart. Nature 2006;441(7096):1011-1014.

100. Eberwine J. Amplification of mRNA populations using aRNA generated from immobilized oligo(dT)-T7 primed cDNA. Biotechniques 1996;20(4):584-591.

101. Eberwine J, Kacharmina JE, Andrews C, Miyashiro K, McIntosh T, Becker K, Barrett T, Hinkle D, Dent G, Marciano P. mRna expression analysis of tissue sections and single cells. Journal of Neuroscience 2001;21(21):8310-8314.

102. Pabon C, Modrusan Z, Ruvolo MV, Coleman IM, Daniel S, Yue H, Arnold LJ, Jr. Optimized T7 amplification system for microarray analysis. Biotechniques 2001;31(4):874-879.

103. Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proceedings of the National Academy of Sciences of the United States of America 1990;87(5):1663-1667.

104. Vincent VAM, DeVoss JJ, Ryan HS, Murphy GM, Jr. Analysis of neuronal gene expression with laser capture microdissection. Journal of Neuroscience Research 2002;69(5):578-586.

105. Eberwine J. Single-cell molecular biology. Nature Neuroscience 2001;4 Suppl:1155-1156.

106. Ginsberg SD, Mirnics K. Functional genomic methodologies. Progress in Brain Research 2006;158:15-40.

107. Zaher H, Fernandez-Salguero PM, Letterio J, Sheikh MS, Fornace AJ, Jr., Roberts AB, Gonzalez FJ. The involvement of aryl hydrocarbon receptor in the activation of transforming growth factor-beta and apoptosis. Molecular Pharmacology 1998;54(2):313-321.

108. Huang G, Elferink CJ. Multiple mechanisms are involved in Ah receptor-mediated cell cycle arrest. Molecular Pharmacology 2005;67(1):88-96.

109. Leary JF, Reece,L.N., Szaniszlo,P., Prow,T. and Wang,N. High-throughput cell analysis and sorting technologies for clinical diagnostics and therapeutics. Proceedings of the Society of Photographic Instrumentation Engineers 2001;4255:16-27.

110. Lam LT, Pickeral OK, Peng AC, Rosenwald A, Hurt EM, Giltnane JM, Averett LM, Zhao H, Davis RE, Sathyamoorthy M and others. Genomic-scale measurement of mRNA turnover and the mechanisms of action of the anti-cancer drug flavopiridol. Genome Biology. 2001;2(10):41.1-41.11.

111. Hanauer DA, Rhodes DR, Sinha-Kumar C, Chinnaiyan AM. Bioinformatics approaches in the study of cancer. Current Molecular Medicine 2007;7(1):133-141.

112. Hu P, Bader G, Wigle DA, Emili A. Computational prediction of cancer-gene function. Nature Reviews Cancer 2007;7(1):23-34.

113. Sasik R, Calvo E, Corbeil J. Statistical analysis of high-density oligonucleotide arrays: a multiplicative noise model. Bioinformatics. 2002;18(12):1633-1640.

114. Hao QL, Shah AJ, Thiemann FT, Smogorzewska EM, Crooks GM. A functional comparison of CD34 + CD38- cells in cord blood and bone marrow. Blood. 1995;86(10):3745-3753.

115. Thoma SJ, Lamping CP, Ziegler BL. Phenotype analysis of hematopoietic CD34+ cell populations derived from human umbilical cord blood using flow cytometry and cDNA-polymerase chain reaction. Blood. 1994;83(8):2103-2114.

116. Todd P, Raghavarao KS, Sengupta S, Doyle JF, Vellinger J, Deuser MS. Multistage electrophoresis system for the separation of cells, particles and solutes. Electrophoresis 2000;21(2):318-324.

117. Todd P, Vellinger JC, Sengupta S, Sportiello MG, Greenberg AR, Krantz WB. Sliding-cavity fluid contactors in low-gravity fluids, materials, and biotechnology research. Annals of the New York Academy of Sciences 2002;974:581-590.

118. Despres D, Flohr T, Uppenkamp M, Baldus M, Hoffmann M, Huber C, Derigs HG. CD34+ cell enrichment for autologous peripheral blood stem cell transplantation by use of the CliniMACs device. J Hematother Stem Cell Res 2000;9(4):557-564.

119. Srour EF, Tong X, Sung KW, Plett PA, Rice S, Daggy J, Yiannoutsos CT, Abonour R, Orschell CM. Modulation of in vitro proliferation kinetics and primitive hematopoietic potential of individual human CD34+CD38-/lo cells in G0. Blood 2005;105(8):3109-3116.

120. Srour EF, Yoder MC. Flow cytometric analysis of hematopoietic development. Methods in Molecular Medicine 2005;105:65-80.

121. Tang H, Kuhen KL, Wong-Staal F. Lentivirus replication and regulation. Annual Review of Genetics 1999;33:133-170.

122. Moore JP, McKeating JA, Weiss RA, Sattentau QJ. Dissociation of gp120 from HIV-1 virions induced by soluble CD4. Science 1990;250(4984):1139-1142.

123. Mohanty SK, Sharma M, Gupta PK. Laser-assisted microinjection into targeted animal cells. Biotechnology Letters 2003;25(11):895-899.

124. Umebayashi Y, Miyamoto Y, Wakita M, Kobayashi A, Nishisaka T. Elevation of plasma membrane permeability on laser irradiation of extracellular latex particles. Journal of Biochemistry 2003;134(2):219-224.

125. Soughayer JS, Krasieva T, Jacobson SC, Ramsey JM, Tromberg BJ, Allbritton NL. Characterization of cellular optoporation with distance. Analytical Chemistry 2000;72(6):1342-1347.

126. Fatemi A, Weiss MA, Weiss RA. Short-term histologic effects of nonablative resurfacing: results with a dynamically cooled millisecond-domain 1320 nm Nd:YAG laser. Dermatologic Surgery 2002;28(2):172-176.

127. Gaon MD, Ho KH, Wong BJ. Measurement of the elastic modulus of porcine septal cartilage specimens following Nd: YAG laser treatment. Lasers in Medical Science 2003;18(3):148-153.

128. He X, Bischof JC. Quantification of temperature and injury response in thermal therapy and cryosurgery. Critical Reviews in Biomedical Engineering 2003;31(5-6):355-422.

129. Bloom M, Evans E, Mouritsen OG. Physical properties of the fluid lipid-bilayer component of cell membranes: a perspective. Quarterly Reviews of Biophysics 1991;24(3):293-397.

130. Hyslop PA, Kuhn CE, Sauerheber RD. Temperature optimum of insulin-stimulated 2-deoxy-D-glucose uptake in rat adipocytes. Correlation of cellular transport with membrane spin-label and fluorescence-label data. Biochemical Journal 1984;218(1):29-36.

131. Moriyama-Gonda N, Igawa M, Shiina H, Wada Y. Heat-induced membrane damage combined with adriamycin on prostate carcinoma PC-3 cells: correlation of cytotoxicity, permeability and P-glycoprotein or metallothionein expression. British Journal of Urology 1998;82(4):552-559.

VITA

Peter Szaniszlo was born on August 7, 1965 in Debrecen, Hungary to Erzsebet and Jozsef Szaniszlo. After completing high school in Debrecen, Hungary in 1983, he attended the Medical University of Debrecen and graduated as an M.D. in 1993. He married Marta Lorinczi on October 26, 1997. Peter was accepted in to the Department of Microbiology and Immunology Graduate Program at the University of Texas Medical Branch at Galveston, in August, 1998.

PUBLICATIONS

Szaniszlo, P., Rose, W.A., Wang, N., Reece, L.M., Tsulaia, T.V., Hanania, E.G., Elferink, C.J., Leary, J.F. "Scanning Cytometry with a LEAP: Laser-Enabled Analysis and Processing of Live Cells In Situ" Cytometry 69A: 641-651, 2005.

Smith, J.N., Reece, L.M., **Szaniszlo, P.**, Leary, R.C. Leary, J.F. "Subtractive Clustering Analysis: A Novel Data Mining Method for Finding Cell Subpopulations" Proc. of SPIE 5699: 354 – 361, 2005.

Szaniszlo, P., Wang, N., Sinha, M., Reece, L.M., Van Hook,J.W., Luxon, B.A., Leary, J.F. "Getting the Right Cells to the Array: Gene Expression Microarray Analysis of Cell Mixtures and Sorted Cells" Cytometry 59A: 191-202, 2004.

Leary, J.F., **Szaniszlo, P.**, Prow, T., Reece, L.M., Wang, N., Asmuth, D.M.: "The Importance of High-Throughput Cell Separation Technologies for Genomics/Proteomics-Based Clinical Therapeutics" Proc. of SPIE 4625: 1-8, 2002.

Leary, J.F. ,Reece, L.N., **Szaniszlo, P.**, Prow, T., Wang, N.: "High-Throughput Cell Analysis and Sorting Technologies for Clinical Diagnostics and Therapeutics. Proc. of SPIE 4255: 16-27, 2001.

Kong, L.Y., **Szaniszlo, P.**, Albrecht, T., Liehr, J.G. Frequency and molecular analysis of hprt mutations induced by estradiol in Chinese hamster V79 cells. International Journal of Oncology. 17(6):1141-9, 2000 Dec.

Boldogh, I., Bui, T.K., **Szaniszlo, P.**, Bresnahan, W.A., Albrecht, T., Hughes, T.K. Novel activation of gamma-interferon in nonimmune cells during human cytomegalovirus replication. Proceedings of the Society for Experimental Biology & Medicine. 215(1):66-73, 1997 May.

Boldogh, I., **Szaniszlo, P.**, Bresnahan, W.A., Flaitz, C.M., Nichols, M.C., Albrecht, T., Kaposi's sarcoma herpesvirus-like DNA sequences in the saliva of individuals infected with human immunodeficiency virus. Clinical Infectious Diseases. 23(2):406-7, 1996 Aug.