cDNA Microarray Analysis of a Glucocorticoid Treated Acute Lymphoblastic Leukemia Cell Line

Emmanouil G. Sifakis, *Student Member, IEEE*, George I. Lambrou, Andriana Prentza, *Senior Member, IEEE*, Dimitris Koutsouris, *Senior Member, IEEE* and Fotini Tzortzatou-Stathopoulou

***Abstract*—The objective of the present study was the analysis of microarray data from a T-cell leukemia cell line (CCRF-CEM), treated with two different prednisolone concentrations, using four different pre-processing methods, within the Matlab® computing environment. We have compared these methods using hierarchical clustering. The gene expression patterns revealed by hierarchical clustering were used to draw probable conclusions on the question whether resistance to glucocorticoids is inherent or acquired, in this type of cells. Although different algorithmic approaches have concluded different results, the set of genes examined manifested an opposing pattern in their expression profile between low and high prednisolone concentrations. This opposing behavior seems to be related to glucocorticoid receptor-related gene repression or activation, leading to the activation of resistance mechanisms within the cell system studied.**

I. Introduction

MICROARRAY is a high-throughput technology allowing the simultaneous screening of the expression levels (differences in labeling between the two fluorescent dyes usually used – Cy3(Green) for the reference sample and Cy5 (Red) for the treated sample) of thousands of genes in one experiment. Yet, there is a substantial question with microarray data analysis on how to draw conclusions out of this vast amount of data, sometimes reaching the size of the whole genome, and how to extract meaningful information about the biological system studied.

The data obtained from the microarray are noisy due to different sources of variation in the experiments. In order to obtain reliable gene expression data for further analysis (e.g. clustering analysis) and interpretation, experimental procedures need to be rigorously controlled to minimize noise and irrelevant variation and different pre-processing methods have been developed for background noise correction, filtering to eliminate poor quality spots and data normalization. However different pre-processing methods [1]-[3] produce different results making the analysis of microarray data mostly conflicting.

Several works have been conducted implementing different pre-processing tasks, i.e. normalization methods, within the Matlab® (The MathWorks, Inc., USA) programming environment [4]-[6], since it comprises a powerful and user-friendly tool for the development and application of algorithms for microarray analysis.

Also, several works have been conducted using microarray technology for high-throughput screening including studies of glucocorticoid action in pediatric Acute Lymphoblastic Leukemia (ALL) both *in vivo* and *in vitro* [7]-[9]. Resistance or sensitivity to glucocorticoids is considered to be of crucial importance for ALL disease prognosis [10], [11]. A common model used for glucorticoid resistance is the T-cell leukemia cell line CCRF-CEM. It has been reported that this cell line is resistant to glucocorticoids [12]-[14], which makes it an ideal model for glucocorticoid resistance. Specifically, the glucocorticoid receptor-mediated apoptosis consists of a very complex and not well understood system. It has been reported that cells possess an intrinsic property for glucocorticoid resistance *in vivo* [8]. It is apparent that the same question can be asked for *in vitro* systems.

In the present study, we have used microarray technology data focusing on the development of several different analysis approaches for the study of the effects of glucocorticoids on leukemic cells. In order to be able to draw conclusions based on different analyses, we have compared four pre-processing methods, using hierarchical clustering as the end result classification. Emphasis was given to the gene expression patterns revealed by different hierarchical clustering methods and not the individual genes. The different analysis methods have finally been utilized to give interpretation to the question whether the system studied possesses an intrinsic property for glucocorticoid resistance or not and how this is reflected from the different analysis methods.

Manuscript received July 5, 2008. This work was supported in part by the national program EPEAEK II, PYTHAGORAS, Ministry of National Education and Religious Affairs.

E. G. Sifakis is with the Biomedical Engineering Laboratory, National Technical University of Athens, Zografou, Athens, 15773, Greece (tel: +30-210-7721142; fax: +30-210-7722431; e-mail: sifakise@biomed.ntua.gr).

G. I. Lambrou is with the Hematology/Oncology Unit, University of Athens, "Aghia Sophia" Children's Hospital, Goudi, Athens, 11527, Greece (e-mail: glamprou@med.uoa.gr).

A. Prentza is with the Institute of Communication and Computer Systems, Biomedical Engineering Laboratory, National Technical University of Athens, Zografou, Athens, 15773, Greece (e-mail: aprentza@biomed.ntua.gr).

D. Koutsouris is with the Biomedical Engineering Laboratory, National Technical University of Athens, Zografou, Athens, 15773, Greece (e-mail: dkoutsou@biomed.ntua.gr).

F. Tzortzatou-Stathopoulou is with the Hematology/Oncology Unit, University of Athens, "Aghia Sophia" Children's Hospital, Goudi, Athens, 11527, Greece (e-mail: ftzortza@med.uoa.gr).

II. MATERIALS AND METHODS

A. Data Collection

The T-cell leukemia cell line was exposed to different concentrations of prednisolone for 4h, and then the gene expression profiles have been analyzed using 1,2k cDNA microarrays (IntelliGene® Human Cancer CHIP Ver. 4.0) obtained from TAKARA BIO Inc., Japan. This microarray chip consists of a set of genes known to be involved in cancer progression, apoptosis, anti-apoptosis, cell cycle and tumour suppression.

The microarray chips were hybridized with cDNA from untreated cells and cells treated with 10nM, and 701μM prednisolone. The following pairs of samples were compared in triplicate experiments:

1) Untreated cells vs. treated with 10nM prednisolone (designated as "*0vs1*").

2) Treated with 10nM vs. treated with 701μM prednisolone (designated as "*1vs3*").

3) Untreated cells vs. 701μM prednisolone (designated as "*0vs3*").

Slides were then scanned with the ScanArray 4000XL microarray scanner, and images were generated with the ScanArray microarray acquisition software (GSI Lumonics, USA). ImaGene v6.0 (Biodiscovery Inc., USA) was then used to analyze the generated images and to produce the raw fluorescent foreground and background signal intensities for each channel Cy3 and Cy5.

B. Preprocessing

Our analysis of the raw intensities was performed within Matlab® v.7.6.0 computing environment, since Matlab®, along with its numerous available Toolboxes, and especially the Bioinformatics Toolbox™, comprises an integrated software environment for gene expression analysis that enables user to perform specific algorithm development from scratch with ease and flexibility adjusted to its sole needs [15], compared to other dedicated gene expression analysis tools.

Throughout our analysis we utilized the spot median intensity values in each channel, as the median is known to be a robust measure of central tendency of the data comparing to the mean, which is not a robust statistic [16].

The following preprocessing methods were developed:

1) Background Correction: Two alternatives for background correction were followed according to [3], [17]:

a) Local background subtracted median value.

b) No background correction.

In the first case, the local background subtracted median value is calculated by using (1), in order to remove the effects of non-specific binding or spatial heterogeneity across the array [3],

$$I_{median} = I_{median}^{foreground} - I_{median}^{background} \quad (1)$$

where $I_{median}^{foreground}$ and $I_{median}^{background}$ correspond to the median intensity values measured for pixels identified in the foreground and the local background by the image processing software for each gene and for each channel, respectively.

In the second case, no background correction is applied (therefore the median value is calculated as in (2)),

$$I_{median} = I_{median}^{foreground} \quad (2)$$

due to several concerns regarding the effectiveness of subtracting the machine-generated background noise measure [3].

2) Intensity-based Filtering of Array Elements: This method was used to filter out low intensity data [18], which is related to bad spots. Low intensity data correspond to spots where there is no significant difference between the foreground and the background intensities. For this case, the absolute lower threshold value of 10 was used in either channel. This specific threshold value was selected in order to reject negative and very small intensity values resulted from background subtraction (first method followed for background correction).

3) Normalization: The background corrected signal intensities were further normalized in order to compensate for systematic variations in the measured gene expression levels of the two co-hybridized mRNA samples, so that meaningful biological comparisons emerge and differences can be more easily distinguished [19].

In the present study, all genes in the array were utilized for normalization based on the hypothesis that only a small proportion of the genes' mRNA will vary significantly in expression between the samples. The normalization approaches followed here were:

a) Intensity-dependent normalization [18], [19], [20] in which we applied the robust scatter-plot smoother LOWESS (LOcally WEighted Scatter-plot Smoother) [21] with the parameter (the fraction of the data used for smoothing at each point) set equal to 40% [19].

b) Logarithm base two transformation which is the most widely used transformation [18].

For the weighted least squares regression applied to our data in the first approach, during the application of the LOWESS method, a linear polynomial model is used, as it provides adequate smoothed points and computational ease [21]. The robust fitting procedure guards against deviant points within the scatter-plot distorting the smoothed points [21]. The regression weights (w_{re_i}) and the robust weights (w_{ro_i}) for each data point within the span (defined by parameter f) were computed by using (3), and (4), respectively [15].

$$w_{re_i} = \left(1 - \left|\frac{x - x_i}{d(x)}\right|^3\right)^3 \quad (3)$$

where x is the predictor value associated with the response value to be smoothed, x_i are the nearest neighbors of x as defined by the span, and $d(x)$ is the distance along the abscissa from x to the most distant predictor value within the span.

$$w_{ro_i} = \begin{cases} (1-(r_i/6MAD)^2)^2, & |r_i| < 6MAD \\ 0, & |r_i| \geq 6MAD \end{cases} \quad (4)$$

where r_i is the residual of the i^{th} data point produced by the regression smoothing procedure, and MAD is the Median Absolute Deviation of the residuals r, $MAD = median(|r|)$.

Throughout our analysis, concerning the LOWESS regression, the independent parameter (predictor value x) was the overall intensity, while the dependent one (response value) was the log-ratio, due to the reported intensity-dependent dye bias [18], [19].

4) Further Selection of Array Elements: As far as, the processing including background correction is concerned, local background intensities vary across each slide and for each channel, thus a significant number of the background corrected spot intensities are either negative or zero. Thereby, a significant number of missing values appears in the LOWESS normalized or log-transformed data set, and therefore we have decided to exclude these spots from further analysis.

Moreover, due to the exclusion of different genes in each of the three experimental set-ups, our analysis included genes present in all our set-ups. This had as a result the rejection of a plethora of genes since only few were present simultaneously in all three experiments.

C. Clustering

After processing the raw data with the various methods referred above, we calculated the expression ratio for each spot in the array for each experimental set-up. As we had no *a priori* knowledge of the complete repertoire of expected gene expression patterns for all experimental set-ups, we have favored an unsupervised method [22]. Thus, the expression ratios were further classified using hierarchical clustering, a method which is widely used and offers simplicity and ease of visualization [1], [2], [22]. More specifically, the clustering method applied to the pre-processed gene intensities was the average-linkage hierarchical clustering algorithm, with the Unweighted Pair-Group Method with Arithmetic mean (UPGMA) as the linkage rule, and the Euclidean distance as the gene similarity metric.

We have also enabled the optimal leaf ordering calculation for the dendrogram display of data, which determines the leaf order that maximizes the similarity between neighboring leaves, since it can reveal biological structure that is not observed with the common applied ordering method [23].

III. Results

1) Ratio vs. Intensity Plots: We first visualized pre-processed data using IR-plots (Ratio-Intensity plots). These are plots of the log-ratio $log_2(I^R/I^G)$ for each element on the array as a function of the product intensities $log_{10}(I^R * I^G)$, where I^R and I^G correspond to the intensity values measured in the Red (Gy5) and the Green (Cy3) channel, respectively. IR-plots can reveal systematic intensity-dependent effects in the measured log-ratio values [18].

IR-plots have been implemented for all experimental set-ups and all combinations between background correction and normalization. As an example, an IR-plot in Fig. 1 is shown, which represents the *0vs3* experiment with median background corrected intensities and further LOWESS normalization applied.

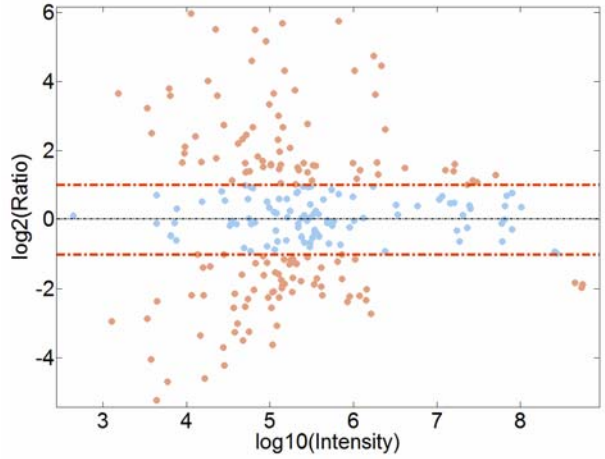


Fig. 1. An IR-plot representing experiment *0vs3* is shown. In this case, the raw intensities have been background corrected by subtracting the median value of the local background intensity and further normalized using the LOWESS algorithm. All dots colored blue are within the 2-fold lines. Dots in orange-red represent data outside the 2-fold lines meaning that they are either over- or under-expressed.

2) Hierarchical Clustering: The cluster analysis for each processing method and each experimental set-up is presented in Fig. 2 and Fig. 3 for the samples without and with background correction, respectively.

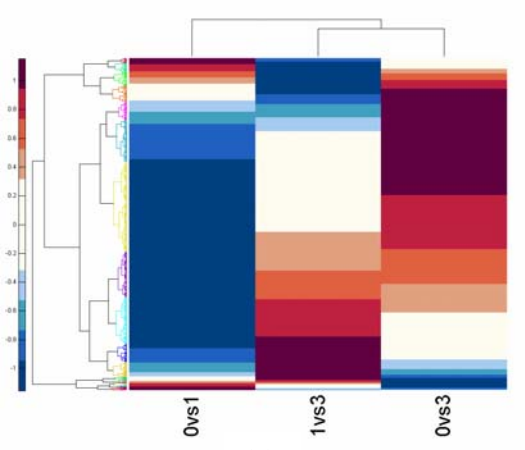
As it has been mentioned above, different processing methods of microarray data give different results. It is interesting to observe that without background correction, no spots (genes) are excluded from the analysis except for

the negative and positive control spots, while after background correction, spot (gene) number is reduced drastically, as described previously. This ensures with high enough confidence that those spots passing this analysis are good quality spots. However, a strict analysis would reject poor quality spots, rejecting at the same time, probably, spots corresponding to genes with biological significance. Also, as it is shown in Fig. 3, the experiment's analysis after background correction without (Fig. 3(a)) and with (Fig. 3(b)) LOWESS normalization revealed no common genes between them. This result made comparison between the two pre-processing methods not possible, yet we have used them separately in order to draw conclusions on the system's behavior.

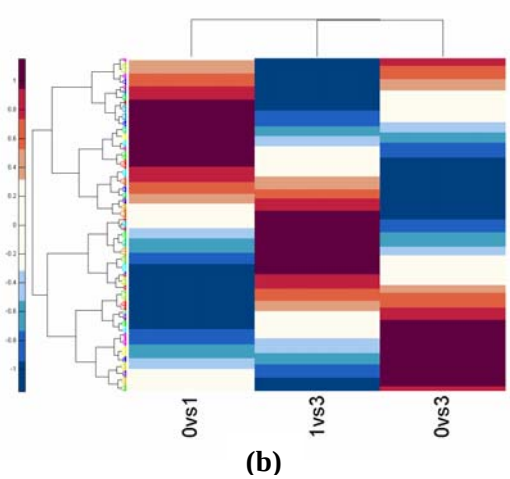
By row comparison of the expression matrices, the expression profiles of the genes can be compared for all pre-processing methods. At first, between the two clusters; with no background correction, with log-transformation and with LOWESS normalization, genes seem to follow a variant differential gene expression pattern (Fig. 2). More specifically, with log-transformation the majority of genes in the *0vs1* experiment seem to be underexpressed, while the opposite happens in the *0vs3* experiment (Fig. 2(a)).

With LOWESS normalization, gene expression of the up- and down-regulated genes seems to be approximately balanced. In other words, the number of genes underexpressed almost equals the number of genes overexpressed.

Almost the same behavior is manifested in the two clusters analyzed with background correction (Fig. 3). When genes are normalized with log-transformation (Fig. 3(a)), it seems that the majority of the remainder genes in the *0vs1* experiment are overexpressed, while the opposite happens in the *0vs3* experiment. In addition, genes normalized with the LOWESS algorithm (Fig. 3(b)) manifest a balance between over- and under-regulated genes.



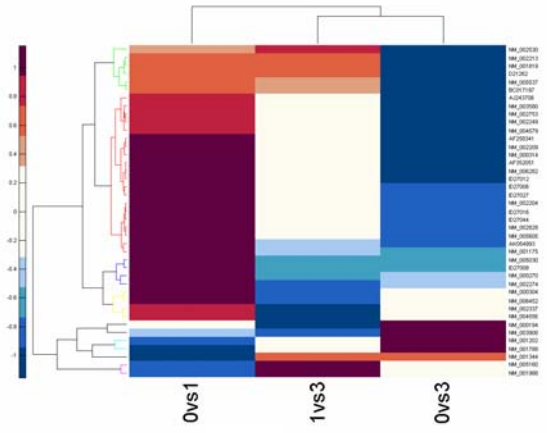
(a)



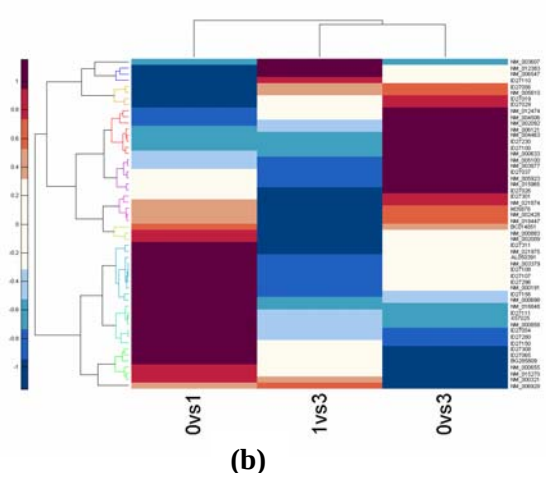
(b)

Fig. 2. Clustered display of the samples handled with no background correction. (a) Average-linkage clustering of the three experiments with log-transformation applied, and (b) average-linkage clustering of the three experiments processed using the LOWESS algorithm for normalization. Expression levels for each gene are standardized across the samples so that the mean is 0, and the standard deviation is 1 [24]. Samples with standardized log ratios of 0 are colored white, increasingly positive log ratios are colored with reds of increasing intensity, and increasingly negative log ratios are colored with blues of increasing intensity.

By column comparison of the expression matrices, the expression profiles of the experiments can be compared for all pre-processing methods. The cluster's pattern of all pre-processing methods reveals the same antithesis in expression profiles between the experiments *0vs1* and *0vs3* (Fig. 2 and Fig. 3). More specifically, genes underexpressed in the *0vs1* experiment appear to be overexpressed in the *0vs3* experiment, and vice versa.



(a)



(b)

Fig. 3. Clustered display of the samples being background corrected. (a) Average-linkage clustering of the three experiments with log-transformation applied, and (b) average-linkage clustering of the three experiments processed using the LOWESS algorithm for normalization. Expression levels for each gene are standardized across the samples so that the mean is 0, and the standard deviation is 1 [24]. Samples with standardized log ratios of 0 are colored white, increasingly positive log ratios are colored with reds of increasing intensity, and increasingly negative log ratios are colored with blues of increasing intensity.

IV. Discussion & Conclusions

the present study we have been concerned with the analysis of the microarray data obtained from a glucocorticoid-resistant cell line treated with two concentrations of prednisolone. Microarray analysis has been further used to assign interpretations to the question whether cells possess an intrinsic mechanism of glucocorticoid resistance or this acquired through the treatment itself.

The higher prednisolone dosage used was 701 μM. Sensitive leukemic cells become apoptotic already at 20nM of dexamethasone treatment [25]. It is reasonable to hypothesize that intrinsic resistance properties would leave genes unaffected at higher prednisolone dosages. Keeping in mind that the set of genes studied consists of apoptosis, cell cycle and tumor suppression genes, the absence of differential expression between untreated and treated with high prednisolone dose would be an evidence for the inherent character of resistance.

If microarray data are pre-processed with the log-transformation only then the conclusion extracted is that the set of genes under study is overexpressed in their majority in the *0vs3* experiment (Fig. 2(a)), which means that prednisolone triggers the resistance mechanism probably through glucocorticoid receptor-regulated target genes. On the other hand, if data are pre-processed with no background correction and LOWESS normalization the same conclusion is extracted, but with the difference that the number of genes underexpressed almost equals the number of genes overexpressed (Fig. 2(b)). This behavior was expected as all genes were utilized in the local regression normalization process.

At the same time, the opposite behavior seems to be consistently present for data processed with background correction and without or with LOWESS normalization. If microarray data are pre-processed with the log-transformation only then the conclusion extracted is that the set of genes under study is underexpressed in their majority in the *0vs3* experiment (Fig. 3(a)). On the other hand, if data are pre-processed with background correction and LOWESS normalization the same conclusion is extracted, but with the difference that the number of genes underexpressed almost equals the number of genes overexpressed (Fig. 3(b)), which was also an expected behavior as local regression applied to the total number of the remainder genes.

In addition, in all algorithmic approaches, cells treated with low dose of prednisolone (*0vs1*) manifest a differential expression profile which seems to be inverted when looking at the *0vs3* experiment, depicting cells treated with high dose of prednisolone. This phenomenon was more or less expected since there is an enormous difference in prednisolone concentrations utilized in these two experimental set-ups. It is also consistent with the known fact that glucocorticoid action (in this case prednisolone) is dose-dependent.

Summarizing, it appeared that, although the different analyses of the microarray data manifested different results one phenomenon has been consistent for all analysis methods performed: the gene expression profile between the low prednisolone concentration versus the untreated cells and the high prednisolone concentration versus the untreated cells manifested an opposite pattern of expression. It has been prementioned that this was an expected result since it is known that prednisolone action is dose-dependent.

Also, it is known that glucocorticoids activate the glucocorticoid receptor which in its turn represses or activates specific genes. The high prednisolone dose was expected to be cytotoxic in a smaller extent as in sensitive cells since the present cell line is resistant. It was, however, unknown whether the high prednisolone dose would trigger gene regulation (through the glucocorticoid receptor) at an early time point such the 4h treatment done in the present study. No differential expression in cells treated with high prednisolone dose would probably mean that cells have an intrinsic mechanism of resistance, while differential expression under the same conditions would mean that prednisolone activates the glucocorticoid receptor which in its turn activates or represses genes responsible for resistance: in other words cells respond to treatment instantly "producing" the resistance.

However, this is a highly controversial subject on which much debate is active. Other analyses performed (data not shown) have manifested that genes differentially expressed by the low prednisolone dose remain unaffected by the high prednisolone dose.

We believe that gene expression patterns revealed by computerized analysis of microarray data from the T-cell leukemia cell line (CCRF-CEM), treated with different

prednisolone concentrations can lead to important understanding regarding whether resistance to glucocorticoids is inherent or acquired in this type of cells. As mentioned before, resistance to glucocorticoids is of crucial importance for leukemia prognosis. The knowledge acquired on the glucocorticoid-induced early gene expression pattern would lead to more effective therapies. In other words, it is of crucial importance to know whether leukemic cells are resistant to glucocorticoids or not, before the later are administered. This approach allows the identification of genes emerging from expression patterns that may constitute molecular targets for drugs in a combination therapy with glucocorticoids. These drugs may affect the potential of glucocorticoids to inhibit growth of resistant leukemia cells. Further evaluation of these methods and further application of other methods is mandatory to evaluate its efficacy and reliability in cancer diagnosis and therapy.

REFERENCES

[1] J. Quackenbush, "Computational analysis of microarray data", *Nat.Rev.Genet.*, vol. 2, pp. 418-427, Jun 2001.

[2] A. Butte, "The use and analysis of microarray data", *Nat.Rev.Drug Discov.*, vol. 1, pp. 951-960, Dec 2002.

[3] M.T. Lee, *Analysis of microarray gene expression data*, Boston: Kluwer Academic, 2004, pp. 371.

[4] D. Venet, "MatArray: a Matlab toolbox for microarray data", *Bioinformatics*, vol. 19, pp. 659-660, Mar 2003.

[5] N. Knowlton, I.M. Dozmorov and M. Centola, "Microarray Data Analysis Toolbox (MDAT): for normalization, adjustment and analysis of gene expression data", *Bioinformatics*, vol. 20, pp. 3687-3690, Dec 2004.

[6] J. Wang, V. Nygaard, B. Smith-Sorensen, E. Hovig and O. Myklebost, "MArray: analysing single, replicated or reversed microarray experiments", *Bioinformatics*, vol. 18, pp. 1139-1140, Aug 2002.

[7] S. Schmidt, J. Rainer, S. Riml, C. Ploner, S. Jesacher, C. Achmuller, E. Presul, S. Skvortsov, R. Crazzolara, M. Fiegl, T. Raivio, O.A. Janne, S. Geley, B. Meister and R. Kofler, "Identification of glucocorticoid-response genes in children with acute lymphoblastic leukemia", *Blood*, vol. 107, pp. 2061-2069, Mar 2006.

[8] G. Cario, A. Fetz, C. Bretscher, A. Moricke, A. Schrauder, M. Stanulla and M. Schrappe, "Initial leukemic gene expression profiles of patients with poor in vivo prednisone response are similar to those of blasts persisting under prednisone treatment in childhood acute lymphoblastic leukemia", *Ann.Hematol.*, Jun 2008.

[9] A.L. Miller, S. Komak, M.S. Webb, E.H. Leiter and E.B. Thompson, "Gene expression profiling of leukemic cells and primary thymocytes predicts a signature for apoptotic sensitivity to glucocorticoids", *Cancer.Cell.Int.*, vol. 7, pp. 18, Nov 2007.

[10] T. Hongo, S. Yajima, M. Sakurai, Y. Horikoshi and R. Hanada, "In vitro drug sensitivity testing can predict induction failure and early relapse of childhood acute lymphoblastic leukemia", *Blood*, vol. 89, pp. 2959-2965, Apr 1997.

[11] M.L. Den Boer, D.O. Harms, R. Pieters, K.M. Kazemier, U. Gobel, D. Korholz, U. Graubner, R.J. Haas, N. Jorch, H.J. Spaar, G.J. Kaspers, W.A. Kamps, Van der Does-Van den Berg,A., E.R. Van Wering, A.J. Veerman and G.E. Janka-Schaub, "Patient stratification based on prednisolone-vincristine-asparaginase resistance profiles in children with acute lymphoblastic leukemia", *J.Clin.Oncol.*, vol. 21, pp. 3262-3268, Sep 2003.

[12] M.E. Lippman, S. Perry and E.B. Thompson, "Cytoplasmic glucocorticoid-binding proteins in glucocorticoid-unresponsive human and mouse leukemic cell lines", *Cancer Res.*, vol. 34, pp. 1572-1576, Jul 1974.

[13] M.R. Norman and E.B. Thompson, "Characterization of a glucocorticoid-sensitive human lymphoid cell line", *Cancer Res.*, vol. 37, pp. 3785-3791, Oct 1977.

[14] E. Laane, T. Panaretakis, K. Pokrovskaja, E. Buentke, M. Corcoran, S. Soderhall, M. Heyman, J. Mazur, B. Zhivotovsky, A. Porwit and D. Grander, "Dexamethasone-induced apoptosis in acute lymphoblastic leukemia involves differential regulation of Bcl-2 family members", *Haematologica*, vol. 92, pp. 1460-1469, Nov 2007.

[15] The MathWorks, Inc.. (2008, June 1). *MATLAB - The Language of Technical Computing* [Online documentation]. Available: http://www.mathworks.com/access/helpdesk/help/techdoc/

[16] M.L. Samuels and J.A. Witmer, *Statistics for the life sciences*, Upper Saddle River, N.J.: Prentice Hall, Pearson Education, 2003, pp. 724.

[17] H.C. Causton, J. Quackenbush and A. Brazma, *Microarray gene expression data analysis : a beginner's guide*, Malden, MA: Blackwell Pub., 2003, pp. 160.

[18] J. Quackenbush, "Microarray data normalization and transformation", *Nat.Genet.*, vol. 32 Suppl, pp. 496-501, Dec 2002.

[19] Y.H. Yang, S. Dudoit, P. Luu, D.M. Lin, V. Peng, J. Ngai and T.P. Speed, "Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation", *Nucleic Acids Res.*, vol. 30, pp. e15, Feb 2002.

[20] T. Park, S.G. Yi, S.H. Kang, S. Lee, Y.S. Lee and R. Simon, "Evaluation of normalization methods for microarray data", *BMC Bioinformatics*, vol. 4, pp. 33, Sep 2003.

[21] W.S. Cleveland, "Robust Locally Weighted Regression and Smoothing Scatterplots", *Journal of the American Statistical Association*, vol. 74, pp. 829, 1979.

[22] M.B. Eisen, P.T. Spellman, P.O. Brown and D. Botstein, "Cluster analysis and display of genome-wide expression patterns", *Proc.Natl.Acad.Sci.U.S.A.*, vol. 95, pp. 14863-14868, Dec 1998.

[23] Z. Bar-Joseph, D.K. Gifford and T.S. Jaakkola, "Fast optimal leaf ordering for hierarchical clustering", *Bioinformatics*, vol. 17 Suppl 1, pp. S22-9, 2001.

[24] T.R. Golub, D.K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J.P. Mesirov, H. Coller, M.L. Loh, J.R. Downing, M.A. Caligiuri, C.D. Bloomfield and E.S. Lander, "Molecular classification of cancer: class discovery and class prediction by gene expression monitoring", *Science*, vol. 286, pp. 531-537, Oct 1999.

[25] R.D. Medh, M.F. Saeed, B.H. Johnson and E.B. Thompson, "Resistance of human leukemic CEM-C1 cells is overcome by synergism between glucocorticoid and protein kinase A pathways: correlation with c-Myc suppression", *Cancer Res.*, vol. 58, pp. 3684-3693, Aug 1998.