ANALYSIS OF AN ASSOCIATIVE MEMORY NEURAL NETWORK FOR PATTERN IDENTIFICATION IN GENE EXPRESSION DATA

Silvio Bicciato
Dipartimento di Processi Chimici dell'Ingegneria, Università di Padova
via Marzolo, 9,
35131, Padova, Italy
silvio.bicciato@unipd.it

Mario Pandin
Dipartimento di Diagnostica Clinica, Ospedale di Cittadella
via Riva Ospedale, 8
35013, Cittadella, Italy
lab_analisis_cit@ulss15.pd.it

Giuseppe Didonè
Dipartimento di Medicina Clinica, Ospedale di Cittadella
via Riva Ospedale, 8
35013, Cittadella, Italy
gdidone@ulss15.pd.it

Carlo Di Bello
Dipartimento di Processi Chimici dell'Ingegneria, Università di Padova
via Marzolo, 9,
35131, Padova, Italy
carlo.dibello@unipd.it

ABSTRACT
DNA microarrays are becoming a standard tool for determining the role of genes in the regulation of any biological process in an organism. The application of this technology for the analysis of gene expression creates enormous opportunities for accelerating the pace towards the understanding of living systems and for the identification of target genes and pathways for drug development. However, equally efficient methods need to be developed for upgrading the information content of the large amounts of data generated by microarray experiments. A procedure for extracting patterns of gene expression through the analysis of the architecture of an associative memory neural network is described. Such patterns contain critical information about the gene-networking relationships observed during changes in cell physiology and the onset of diseases.
The proposed method has been tested on two different microarray data sets, namely DeRisi's experiment on yeast cultures [10] and Golub's analysis of acute human leukemia molecular profiles [17]. Using these data sets, the neural network structure has been examined to extract relationships among different genes involved in major metabolic pathways and to relate specific genes to different classes of leukemia.

Keywords
Gene expression data, *c*DNA microarrays, neural networks, pattern recognition, data mining

1. INTRODUCTION
DNA microarrays and related *lab-on-a-chip* technology are expected to radically boost the understanding of living systems, thus creating enormous opportunities to identify target genes and pathways for drug development and to elucidate networks of genomic regulation by the comparison of the phenotype of cells in different physiological states. However, the increasing pace of genomic data accumulation poses the challenge to develop analysis procedures able to interpret and upgrade the information content of these databases.

Several different methods have been proposed to analyze large amounts of expression profiling data. Among the analytical tools applied to mine microarray data, are visual discovery and interpretation procedures [10, 13], singular value decomposition and projection on principal component planes [2, 20, 28], supervised machine learning techniques [6, 15], Fourier analysis [31], relevance networks [7-8], self-organizing maps [17, 34-35], procedures based on network inference [11], and Shannon entropy calculations [14].

Among all, the most widely used technique is hierarchical agglomerative clustering. As reported in many publications, clustering techniques have been applied to identify groups of genes sharing similar expression profiles and the results obtained so far are extremely valuable [1, 3, 5, 12, 16, 19, 25, 27, 31, 36]. However, it should be pointed out that the metrics normally used in most clustering methodologies allow to capture only a subset of the relationships that potentially exist among the various transcripts [29]. Indeed, the typical clustering algorithm forms groups whose members have a *consistent uniform behavior* thus *shadowing* all the possible associations among the expression profiles. This problem reflects the difficulty in identifying *inverse correlations* among different transcripts, e.g. those genes that are simultaneously up regulated when a specific gene is down regulated. Since transcriptional control is the result of complex networks that interpret a variety of inputs, the development of analytical tools that detect the multivariate nature of complex genetic relationships is essential [22]. Moreover, clustering methodologies represent an example of unsupervised analysis that is not appropriate for the incorporation of prior knowledge about the observations, as for example sample labels (normal or tumor).

As such, cluster analysis may not be a good framework for diagnosis or classification of diseases [38].

The purpose of this work is to present a procedure for determining all the possible relationships and associations among different genes (unsupervised analysis of unlabeled data) and to highlight *dynamic* patterns of expression correlated to peculiar phenotypes (supervised analysis of labeled samples). Specifically, this approach is an attempt to formulate a computational framework that *simultaneously* allows extracting patterns of induction and repression, identifying specific markers of phenotypes, and predicting the class label of a set of previously unseen instances. Since the above issues cannot be resolved mechanistically at the present time due to inadequate understanding of relevant biological mechanisms, the approach proposed here is entirely *data driven* and requires no *a priori* hypothesis. A novel neural network, resembling characteristics of the associative memory scheme and of the multi-layered connected networks, has been designed. Like in multi-layered networks, this particular architecture preserves the ability to learn and recall non-linear associations, but, as in associative memory maps, allows reversing the mechanism of knowledge encoding. Specifically, the strength of each network connection (represented by the positive and negative values of the weight matrix) has been related to the degree of relationship between the different descriptors of the input system, i.e. the measured expression levels and/or the class labels. The objective was to decode and analyze the network parameters (weight matrix) to discover and extract patterns from the gene expression data used to train and cross-validate the network.

This neural network has been applied to two different case studies. In the first one, the network was trained to reconstruct diauxic shift data from DeRisi's experiment on yeast cultures [10]. After completion of the training phase, the weight matrix has been analyzed to identify relationships among different genes involved in some major metabolic pathways. In the second example, namely the data set from Golub's work on cancer classification [17], the network architecture has been inspected to relate specific genes to different classes of leukemia and to predict the phenotype of a previously unseen set of samples.

This approach does not require any mechanistic model or hypothesis of the way in which the genes relate to one another or influence a particular phenotype. Instead, hypotheses can be developed from the synthesis of its results. The trade-off is that the neural network characteristics limit the analysis to small subsets of the entire expression matrix (less than 100 transcripts per analysis). However, the goal is not to produce a complete model of the genetic relationships, but to help focus the identification of unexpected new connections in already characterized processes and the finding of unexpected connections between processes not previously known to be coordinated.

2. METHOD AND IMPLEMENTATION

Associative memory maps normally consist of only two layers of neurons where each neuron is fully connected with all neurons of the other layer thus forming an interconnection or synaptic matrix *W*. Each element w_{ij} of the synaptic matrix is the interconnection weight of the synapse between the neuron *i* of one layer and neuron *j* of the other. The values of matrix *W* are normally determined using the Hebbian learning method during the training phase and store the ability of the map to recall the associations (if any) embedded in the system. Associative memory neural networks can be considered linear associators since the knowledge formation can be reconstructed analytically. On the other hand, autoassociative neural networks are a specific type of multi-layered neural networks trained to generate an identity association in which the network outputs approximate the inputs using non-linear transfer functions [4, 30]. The architecture of these networks is such that they do not learn the identity mapping perfectly. The input to the network is a vector of measurements, so the size of the input layer corresponds to the dimension of the measurement vector. Since the output layer produces a reconstructed version of the inputs it is also of the same dimension as the input. The training process can be considered unsupervised; since no target data (different from the input) is involved in the training process, the weights among the different layers of the network are allowed to evolve until the correct sequence learning has been achieved. A backpropagation algorithm adjusts the network weights so that the reconstructed measurements vector at the output layer matches the input as closely as possible, in a least-squares sense, over the set of the training examples.

Multi-layered neural networks are generally considered black boxes; this aspect limits their applications to all those problems where model interpretation is as important as quality of prediction. Indeed, it is difficult to formulate hypotheses on how to reverse the knowledge encoding process of neural networks. They can approximate any continuous linear or nonlinear function by combining and rotating sigmoid-type transfer functions. The final model can be expressed in the form of a complex relationship:

$$\hat{x}_k = \tilde{g}\left[\sum_{j=0}^{M} w_{kj}^{(2)} g\left(\sum_{i=0}^{d} w_{ji}^{(1)} x_i\right)\right] \qquad (1)$$

where $\hat{x}_k$ is the output of the *k*th output unit, $w_{ji}^{(1)}$ are the weights between the *d* input nodes and the *M* units of the hidden layer, $w_{kj}^{(2)}$ are the weights between hidden and output layer, x_i are the input variables, and $\tilde{g}(x)$ and $g(x)$ are nonlinear transfer functions. In the case of an autoassociative neural network, each *d*-dimensional input vector $\mathbf{x}^n$ is mapped onto itself and the model can be trained by minimizing a sum-of-squares error of the form:

$$E = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{d}\left[\hat{x}_k^n - x_k^n\right]^2 \qquad (2)$$

over the *N* training examples.

Because of their highly interconnected architecture, multi-layered neural networks allow little inference to the user on the relative relationships of the different input variables. Even using the architecture developed by Kramer [22] for non-linear principal component analysis, the full connection between the different layers makes almost impossible to keep track of variable reconstruction at the output layer.

2.1 Associative memory neural network with input-hidden mono-connections

To overcome this bookkeeping problem while preserving the ability to resolve any linear and non-linear association problem, a particular architecture of autoassociative neural network with mono-connections between the input and the hidden layers has been designed (Figure 1).

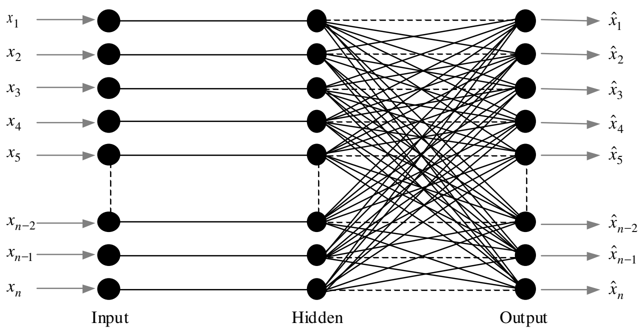

Figure. 1. Architecture of autoassociative neural network with input-hidden mono-connections.

The network consists of three layers, an input layer, a hidden layer, mono-connected to the input layer, and an output layer fully connected to the hidden layer. Thus, input node i is feeding only hidden node i in a manner similar to what the axons do in a real neuronal system. Indeed, the mono-connection between input and hidden layers originated, through a pruning process, from a fully connected network originally used for simulating a biological receptor. Several experiments indicated that the input-hidden mono-connection appeared promising for building a model of biological learning [26]. In spite of its simplicity, the mono-connected architecture preserves the multi-layer perceptron architecture, presenting two layers of adapting weights. As such, this network is capable of approximating any continuous functional mapping [4]. It is worthwhile to stress that the mono-connection between input and hidden layer allows to uniquely identifying a single input variable at a specific hidden node, i.e. node i of the hidden layer represents variable i that has been fed at input node i. The addition of other connections between the input and the hidden layer would result in the bookkeeping problem described above and would not add any benefit in the extraction of the associations. The weights between these two layers are adapted during the training process as in any other multi-layered feed-forward network, with the only difference represented by $w_{ji}^{(1)} = 0$ if $i \neq j$. The full connection between hidden and output layers determines the reconstruction of each input variable as a combination of all the others, i.e. at node k of the output layer there is a reconstructed representation of variable k as a result of all the other n variables.

According to Eq. (1), at the end of the training phase (i.e., given x_i), each hidden-output connection weight can be interpreted as a quantification of the degree of relationship of a variable to another in the presence of all the n descriptors of the system. If the weight of the connection between node j of the hidden layer and node k of the output has a large absolute value, then variable j strongly influences the reconstruction of variable k. In particular:

- a large positive weight indicates a positive contribution to the reconstruction (*positive affinity* or *positive behavior*);
- a large negative weight indicates a negative contribution to the reconstruction (*negative affinity* or *negative behavior*);
- a small absolute weight is to be interpreted as a weak contribution to the reconstruction (*insensitive behavior*).

The network has been trained with a gradient descent network learning mechanism (back-propagation algorithm by Rumelhart *et al.*, 1986), using the generalized delta rule. The learning rate values have been set in the range 0.01÷0.1 to avoid a too slow convergence (learning rate value too small) while preventing network instability (learning rate too large). No particular temperature expression or different momentum parameter for accelerating the learning phase has been used during the training. The logistic function has been utilized as activation function. The stopping criterions for the learning phase were the minimization of the classification mapping error and of the error function derivative defined with respect to the training data set. All the input data (genes) have been individually normalized between 0 and 1, in a preprocessing phase before being presented to the network. The initial weights for the input-hidden and the hidden-output layers were set to fixed values (range 0.01÷0.8). To avoid the creation of trivial relationships between a variable and itself, the weights of the connections between node k of the hidden layer and node k of the output layer can be set to 0 at the network initialization and kept to null values during the entire training phase.

Upon completion of the training phase, the weight matrix can be visualized and inspected using a *Hinton diagram* [4] or a *contribution plot*. In a *contribution plot* (Figure 2) each bar corresponds to the weight, normalized with respect to the maximum value of the weight matrix, of each connection entering the node of the analyzed variable. The height of the bar represents the magnitude of the corresponding parameter, while the color represents the sign of the weight, with gray for positive and white for negative. Multiple bars represent multiple training sessions normally used to check the connection stability. For example, the contribution plot of Figure 2 represents the weights of the connections between hidden layer nodes and output node # 9 in a mono-connected network with 15 nodes per layer. The value of output node # 9 (Var9) receives large positive contributions from nodes 3 and 15 (variables 3 and 15; Var3 and Var15) and negative ones from nodes 10, 11, and 13 (variables 10, 11, and 13; Var10, Var11, and Var13). The remaining nodes (variables) weakly contribute to the value of output node # 9.

3. RESULTS

The proposed method has been tested using two different case studies. The first data set contains the temporal expression profiles of a yeast cell culture during diauxic shift. The data set, originally analyzed by DeRisi *et al.* (1997), can be obtained at Stanford Genomic Resources website http://genome-www.stanford.edu/. The second example comprises data from patients with human acute leukemia. Golub *et al.* (1999) originally published these data (available for download at http://waldo.wi.mit.edu/MPR/data_set_ALL_AML.html).

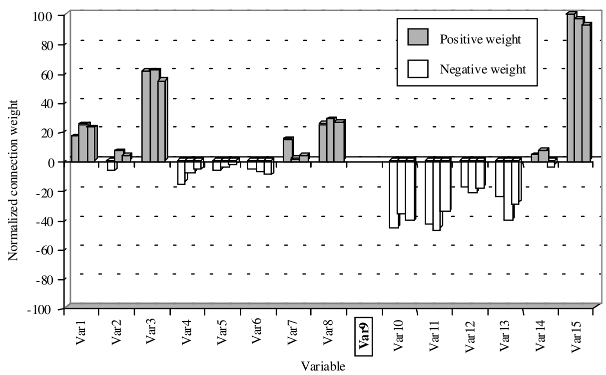

Figure. 2. Contribution plot for an autoassociative neural network with input-hidden mono-connections and 15 input, hidden, and output nodes. The plot refers to a database of fermentation batches described by 15 process variables [21].

3.1 Diauxic shift data set

In DeRisi's experiment DNA microarrays containing virtually every gene of *S. cerevisiae* were used to monitor the temporal changes of gene expression levels accompanying the metabolic shift from fermentation to respiration upon glucose exhaustion [10]. Data consist of 7 time point profiles of logarithmic expression level ratios.

Since regulatory and signal relationships are currently not sufficiently known and because of the size limitation imposed by the relative small number of time points, the analysis was restricted to 77 genes involved in the major metabolic energy pathways (i.e. glycolysis, gluconeogenesis, pentose phosphate pathway, TCA cycle, glyoxylate shunt, fermentation). An autoassociative neural network with input-hidden mono-connections and 77 units per layer has been designed and trained. The training process involved several training sessions over different random selected training sets to verify the stability and the independence of the identified patterns from the network initialization conditions. The random selection of the training sets prevented similar expression profiles from establishing undesired biases in the creation of gene correlations. Upon training completion, the weight matrix has been inspected to identify functional relationships among specific genes. In Figure 3 is shown the contribution plot for the node reconstructing the time series relative to *FBP1*, fructose biphosphate phosphatase. *FBP1* reconstruction receives several positive and negative contemporaneous contributions. Positive and negative contributing genes have been functionally characterized using KEGG metabolic maps of the central carbon metabolism for *S. cerevisiae* (http://www.genome.ad.jp/kegg/kegg.html) and MIPS Yeast Genome Database (http://www.mips.biochem.mpg.de/proj/yeast/).

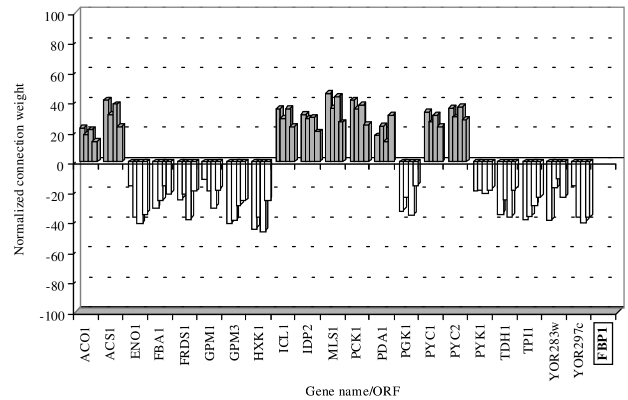

Figure. 3. Contribution plot for the node reconstructing the time series relative to *FBP1*,. Multiple bars represent multiple training sessions. For sake of simplicity, weight values between -15 and +15 have been omitted.

The contribution plot and the functional annotations show that *FBP1* is positively related to genes that are involved in anaplerotic pathways, which convert the products of fermentation in acetyl-CoA. Acetyl-CoA is in turn used to fuel the TCA cycle, the glyoxylate shunt, and gluconeogenesis. At the same time, *FBP1* has opposite relationships with genes coding for glycolytic enzymes. Indeed, *FBP1* activity is almost absent in cells grown on glucose while is greatly increased when cells are grown on media different from glucose [32-33].

A confirmation of the metabolic changes upon glucose exhaustion can be derived also from the weight contributions to phosphoenolpyruvate carboxykinase (*PCK1*) reconstruction (Figure 4).

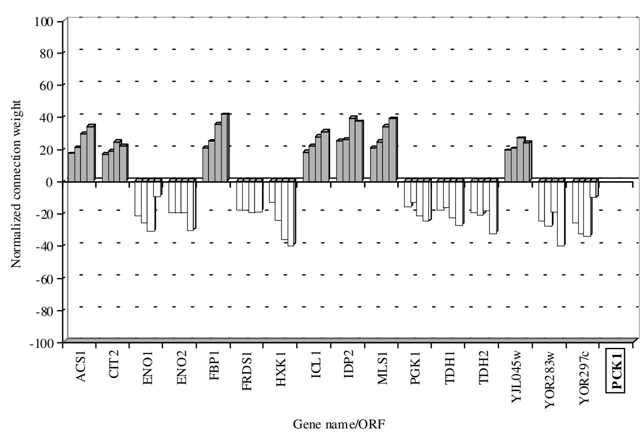

Figure. 4. Contribution plot for the node reconstructing the time series relative to *PCK1*, phosphoenolpyruvate carboxykinase. Multiple bars represent multiple training sessions. For sake of simplicity, weight values between -15 and +15 have been omitted.

From Figure 4, it can be seen how positive and negative weights account for different specific metabolic pathways. Specifically, it is to note the positive affinity of *PCK1* to *FBP1*. These two pivotal genes, through the glyoxylate shunt, reverse the flow of metabolites along the glycolytic pathway whose genes have indeed an opposite affinity to *PCK1*. Moreover, all the genes having a positive influence on *PCK1* reconstruction (with the only exception of *IDP2*) are known to be glucose-repressed and to share a common upstream activating sequence, the carbon source response element (CSRE).

Similarly to correlation coefficients, positive and negative contributions to the reconstruction of a specific gene (positive and negative weights) simply reflect the kind of relationship exiting among different transcripts. Thus, the sign of the network weights do not represent by itself up or down regulation. An example of the fact that positive and negative weights are only an indication of the type of correlation and do not directly mean induction or repression is given by the different relationships established by the two acetyl-CoA synthases, *ACS1* and *ACS2*, in the yeast Acetyl-CoA synthase system. *ACS1* is induced by glucose exhaustion and, as such, receives positive contribution by induced genes (i.e., genes of gluconeogenesis and glyoxylate shunt), as shown by the weights of Figure 5a. Instead *ACS2*, which is weakly repressed during diauxic shift, receives positive contribution by repressed genes of glycolysis and negative ones by the induced genes of TCA cycle (Figure 5b).

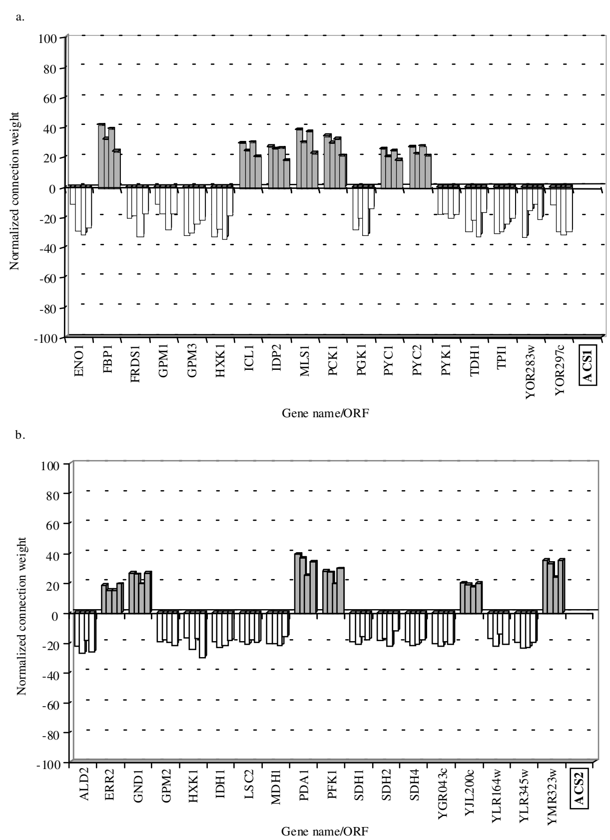

Figure. 5. Contribution plots for nodes reconstructing the time series relative to yeast Acetyl-CoA synthase system. a. *ACS1*; b. *ACS2*. Multiple bars represent multiple training sessions. For sake of simplicity, weight values between -15 and +15 have been omitted.

3.2 Leukemia data set

The second case study is Golub's data set on cancer classification in which 72 samples coming from patients affected by acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) have been hybridized on an Affymetrix chip with 7129 human genes [17].

Before training the network, a mean hypothesis testing procedure has been applied to focus on a key number of genes most discriminating between the two classes of leukemia. Mean Hypothesis Testing (MHT) is a statistical method to identify variables that are discriminating among different classes of system behavior [21]. Utilizing the mean as a measure of the class's behavior, MHT seeks those measurements that violate the null hypothesis that the means of two classes are similar. In this paper MHT has been adapted to identify those gene expression measurements that are able to discriminate between ALL and AML. The decision to reject or to accept the null hypothesis has been made using a *t*-statistic [24]. The applicability of the statistical test has been confirmed by a check for normality of the data. The two means and a pooled standard deviation have been used to construct the *t*-statistic, which has been then compared to a tabulated (threshold) *t*-value, with a selected significance level of 95%. The quantity *t-ratio*, representing the ratio of the *t*-statistic to the threshold *t*-value for each gene of the data set, and the confidence of the comparison between the means, has been calculated. The null hypothesis has been rejected or accepted whether or not the *t*-ratio exceeded 1 (details of the entire hypothesis test can be found in Mardia *et al.*, (1979) and Kamimura *et al.*, (2000)). Finally, the 50 genes with the highest *t*-ratios and the lowest the probability of observing the given result by chance have been selected for the network analysis (25 over-expressed in ALL, i.e. having a mean expression level larger in ALL than in AML, and 25 over-expressed in AML).

We designed and trained a mono-connected autoassociative neural network with 52 nodes per layer, 50 nodes representing the 50 selected genes plus 2 nodes coding the class label (ALL or AML) of the 72 samples. Following the experimental setup of Golub *et al.*, the data has been split into a training set consisting of 38 samples (27 ALL and 11 AML samples) and a test set of 34 samples (20 ALL and 14 AML). ALL class has been coded as class 1, AML being class 0. After training, the weight matrix has been inspected to identify specific expression attributes related to the two distinct phenotypes. The transcripts positively contributing to the reconstruction of ALL and AML classes are listed in Table 1 and 2, respectively. Each weight value is calculated as the mean weight value over 100 different training sessions. The standard deviation of the weights over the 100 training sessions gives a measure of the stability of the association found by the neural network. In Table 1 and 2 are reported only those variables (genes) that presented a positive weight value larger than 2-times the relative standard deviation over the 100 sessions. The ranking of each gene in each of the two classes according to the MHT procedure is in the last column of the table. Thus, *ALL 1* means that the given gene showed the highest t-ratio and confidence values among the 25 genes that are over-expressed in the ALL samples.

Table 1. Genes with positive contributions to the reconstruction of ALL class.

Acc. #	Symbol	Description	Weight value*	Weight std*	MHT ranking
M92287	CCND3	Cyclin D3	97.41	0.403	ALL 11
U05259	MB-1	MB-1 gene	53.98	5.448	ALL 16
X59417	PSMA6	Proteasome iota chain	28.17	3.583	ALL 3
M31303	LAP18	Oncoprotein 18 (Op18) gene	23.22	2.754	ALL 8
U22376	MYB	C-myb gene extracted from Human (c-myb) gene	21.65	2.894	ALL 1
L47738	PIR121	Inducible protein mRNA	13.01	1.456	ALL 21
D26156	SMARCA4	Transcriptional activator hSNF2b	6.15	0.735	ALL 5
M28170	CD19	CD19 antigen	3.74	0.270	ALL 9
S50223	ZNF22	Zinc finger protein 22 (KOX 15)	3.44	0.256	ALL 10
M65214	TCF3	Transcription factor 3 (E2A immunoglobulin enhancer binding factor)	3.38	0.342	ALL 25

*Average value over 100 different training sessions.

Table 2. Genes with positive contributions to the reconstruction of AML class.

Acc. #	Symbol	Description	Weight value*	Weight std*	MHT ranking
M27891	CST3	Cystatin C (amyloid angiopathy and cerebral hemorrhage)	72.49	4.474	AML 17
M84526	DF	D component of complement (adipsin)	60.05	4.962	AML11
M63138	CTSD	Cathepsin D (lysosomal aspartyl protease)	26.31	3.373	AML 18
X85116	EPB72	Epb72 gene exon 1	22.03	2.541	AML 21
X17042	PRG1	Proteoglycan 1, secretory granule	20.64	3.477	AML 10
X95735	ZYX	Zyxin	12.10	1.561	AML 2
M28130	IL8	Interleukin 8 (IL8) gene	9.61	2.404	AML 19
M23197	CD33	CD33 antigen (differentiation antigen)	7.68	0.446	AML 6
M55150	FAH	Fumarylacetoacetate	6.51	0.381	AML 3
Y12670	LEPR	Leptin receptor	6.27	0.558	AML 5
U82759	HoxA9	Homeodomain protein HoxA9 mRNA	5.84	0.345	AML 7
U50136	LTC4S	Leukotriene C4 synthase (LTC4S) gene	3.77	0.586	AML 1
D49950	IL18	Interleukin 18 (interferon-gamma-inducing factor)	2.95	0.301	AML 9
M80254	PPIF	Peptidyl-prolyl-cis-trans isomerase, mithocondrial precursor	2.18	0.343	AML 12

*Average value over 100 different training sessions.

Each class is characterized by a peculiar set of genes defining the specific phenotype and the results of Table 1 and Table 2 represent *class fingerprints* of gene expression correlation. This class fingerprint has to be interpreted as a hypothesis to help focusing the definition of a set of markers for the specific phenotype or a group of targets for studying the interactions in genetic networks. A strong class marker is a gene that has a strong positive connection with one class and, at the same time, a large negative connection weight to the other. These molecules that discriminate between different phenotypes are good candidates for subsequent in-depth experimental studies as well as developing decision support systems. For example, gene encoding for cyclin D3 (*CCND3*) can be considered a strong marker of lymphoblastic subtype (connection weight of 97.41 with ALL and of –100 with AML, data not shown). This result is confirmed by Chow and Golub findings [9, 17]. Even if the role of this late-G1/S phase gene in ALL setup is not completely clear, its chromosomal assignment raises the hypothesis that it may be involved in the pathogenesis of human neoplasms with abnormalities of chromosome 6. Similarly, the CD79A antigen (immunoglobulin-associated alpha), or *MB-1* gene, is highly correlated to lymphoblastic phenotype (connection weight of 53.98) and non-correlated to myeloid subtype (connection weight of –54.35, data not shown). *MB-1* gene encodes the Ig-alpha protein of the B-cell antigen component but may have other functions in addition to its role in signal transduction in B lineage cells [18]. The correlation between ALL and *MB-1* expression is confirmed by the expression studies of the NCI/NIH Developmental Therapeutics Program (http://dtp.nci.nih.gov/) and by Ha *et al.* [18]. Ha reports that *MB-1* transcripts could be detected in pre-B cell lines and fetal bone marrow, in normal, mitogen activated- and transformed B cells but not in myeloma plasma cells. Furthermore, *MB-1* is located in the 19q13 chromosomal region known to be a site of recurrent abnormalities in ALL. On the contrary, cystatin C, adipsin, and cathepsin D are identified by the associative memory as relevant features of the myeloid subtype. Indeed, they are characterized by large positive connection weights with the node reconstructing AML class (72.5, 60.1, and 26.3, respectively) and large negative ones with the ALL node (-84.1, -64.3, and –32.9, respectively). Cystatins are endogenous protein inhibitors of cathepsins and these protease-inhibitor pairs, reported in myeloid cell lines with altered development, might be important in the etiology of AML. Adipsin is contained in the locus 19p13.3 whose chromatin reorganization is known to be associated with myeloid cell differentiation [37].

Similarly, the weight values of gene *EPB72* (22 toward AML and –24.5 with ALL) suggest that this transcript is a myeloid subtype marker. Indeed, *EPB72* is located in the region 9q34.1 which is proximal to the breakpoint that creates the Philadelphia chromosome of chronic myeloid leukemia and, therefore, to the Abelson myeloid leukemia oncogene.

It is worthwhile noting that the feature relevance determined by the network does not resemble completely the t-test ranking. Indeed, the associative memory is influenced also by the magnitude of the signals and thus it ranks the transcripts according to a combination of their discriminating power and absolute expression levels. In this sense the top ranking genes found inspecting the network weights are to be considered markers with a stronger quantitative relevance than those identified using the MHT procedure.

Once completed the training phase, the values of the two nodes coding for the class label have been analyzed to verify the network performance in class prediction. In Table 3 are reported the actual (target) and predicted (output) classification during testing. No sample has been misclassified during the 100 training sessions, while the network could never classify correctly samples 57 and 66 during the testing phase and samples 58 and 61 present a *borderline* classification. The generalization performance of Table 3, i.e. 88% of correct estimates, is slightly better than the best classifications obtained using Chow and Golub methods [9, 17].

4. CONCLUSIONS

A procedure for extracting patterns of gene expression through the analysis of the architecture of an associative memory with mono-connection between the input and the hidden layer has been presented. The particular network connection scheme (mono-connection between the input and the hidden layers) allows keeping track of variable reconstruction at the output layer, while preserving the multi-layer perceptron architecture. This feature has been applied for analyzing the parameters of knowledge encoding (weight matrix) to discover patterns of gene expression data and relate them to the gene-networking representing changes in cell physiology and the onset of diseases. Upon training completion, the network has been able to identify relationships among different genes involved in yeast major metabolic pathways as well as to extract markers responsible for distinct subtypes of leukemia (unsupervised learning). The generalization performance of the network has also been checked during the supervised analysis of labeled samples. The associative memory has been able to estimate the label (ALL or AML) of previously unseen samples in the 88% of the cases, thus representing one of the more robust and performing models for cancer classification.

Research is presently focusing on the application of multiple parallel networks, since the size limitation imposed by the use of neural nets still represents a major unsolved drawback for the analysis of large genomic databases.

5. ACKNOWLEDGMENTS

R.T. Kamimura is acknowledged for valuable comments and revision of the manuscript.

Table 3. Actual (target) and predicted (output) classification during testing phase. Output values (output) and standard deviations (out_std) are averaged over 100 training and testing sessions.

Sample #	Class	Target	Output	Out_std	Call
39	ALL	1	1.01	0.02	Correct
40	ALL	1	0.83	0.12	Correct
41	ALL	1	1.01	0.01	Correct
42	ALL	1	0.79	0.11	Correct
43	ALL	1	0.89	0.04	Correct
44	ALL	1	1.03	0.01	Correct
45	ALL	1	0.99	0.04	Correct
46	ALL	1	1.02	0.01	Correct
47	ALL	1	0.84	0.07	Correct
48	ALL	1	1.03	0.01	Correct
49	ALL	1	1.02	0.02	Correct
50	AML	0	-0.04	0.00	Correct
51	AML	0	0.11	0.05	Correct
52	AML	0	0.37	0.11	Correct
53	AML	0	0.08	0.09	Correct
54	AML	0	0.09	0.07	Correct
55	ALL	1	0.94	0.04	Correct
56	ALL	1	0.99	0.03	Correct
57	**AML**	**0**	**0.64**	**0.14**	**Wrong**
58	**AML**	**0**	**0.44**	**0.08**	**Border**
59	ALL	1	0.97	0.04	Correct
60	AML	0	0.34	0.08	Correct
61	**AML**	**0**	**0.55**	**0.11**	**Border**
62	AML	0	0.21	0.10	Correct
63	AML	0	0.10	0.06	Correct
64	AML	0	0.02	0.04	Correct
65	AML	0	0.33	0.11	Correct
66	**AML**	**0**	**0.73**	**0.04**	**Wrong**
67	ALL	1	0.75	0.10	Correct
68	ALL	1	1.00	0.01	Correct
69	ALL	1	0.98	0.03	Correct
70	ALL	1	1.00	0.02	Correct
71	ALL	1	0.96	0.03	Correct
72	ALL	1	0.89	0.09	Correct

6. REFERENCES

[1] Alon, U., Barkai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., and Levine, A.J. (1999) Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc Natl Acad Sci USA.*, **96**(12), 6745-6750.

[2] Alter, O., Brown, P.O., and Botstein, D. (2000) Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci USA.*, **97**(18), 10101-10106.

[3] Ben-Dor, A., Shamir, R., and Yakhini, Z. (1999) Clustering Gene Expression Patterns. *Journal of Computational Biology*, **6**(3/4), 281-297.

[4] Bishop, C.M. (1995) *Neural networks for pattern recognition*. Oxford University Press, Oxford.

[5] Brazma, A., and Vilo, J. (2000) Gene expression data analysis. *FEBS Letters*, **480**, 17-24.

[6] Brown, M.P., Grundy, W.N., Lin, D., Sugnet, C., Ares, M., and Haussler, D. (2000) Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc Natl Acad Sci USA.*, **97**, 262-267.

[7] Butte, A.J., and Kohane, I.S. (2000) Mutual information relevance networks: functional genomics clustering using pairwise entropy measurements. *Proc. Pacific Symposium on Biocomputing*, **5**, 415-426.

[8] Butte, A.J., Tamayo, P., Slonim, D., Golub, T.R., and Kohane, I.S. (2000) Discovering Functional Relationships Between RNA Expression and Chemotherapeutic Susceptibility Using Relevance Networks. *Proc Natl Acad Sci USA.*, **97**(22), 12182-12186.

[9] Chow, M.L., Moler, E.J., Mian, I.S. (2001) Identifying marker genes in transcription profiling data using a mixture of feature relevance experts. *Physiol Genomics*, **5**, 99-111.

[10] DeRisi, J.L., Iyer, V.R., and Brown, P.O. (1997) Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. *Science*, **278**(5338), 680-686.

[11] D'haeseleer, P., Liang, S., and Somogyi, R. (2000) Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics*, **16**(8), 707-726.

[12] Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns, *Proc Natl Acad Sci USA.*, **95**(25), 14863-14868.

[13] Ermolaeva, O., Rastogi, M., Pruitt, K.D., Schuler, G.D., Bittner, M.L., Chen, Y., Simon, R., Meltzer, P., Trent, J.M., and Boguski, M.S. (1998) Data management and analysis for gene expression arrays. *Nature Genetics*, **20**(1), 19-23.

[14] Fuhrman, S., Cunningham, M.J., Wen, X., Zweiger, G., Seihamer, J.J., and Somogyi, R. (2000) The application of Shannon entropy in the identification of putative drug targets. *Biosystem*, **55**, 5-14.

[15] Furey, T.S., Cristiannini, N., Duffy, N., Bednarski, D.W., Schummer, M., and Haussler, D. (2000) Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics*, **16**(10), 906-914.

[16] Getz, G., Levine, E., and Domany, E. (2000). Coupled two-way clustering analysis of gene microarray data. *Proc Natl Acad Sci USA.*, **97**(22), 12079-12084.

[17] Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., and Lander, E.S. (1999) Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring, *Science*, **286**(5439), 531-537.

[18] Ha HJ, Kubagawa H, Burrows PD Molecular cloning and expression pattern of a human gene homologous to the murine mb-1 gene. *J Immunol* 1992 Mar 1;148(5):1526-31

[19] Herwig, R., Poustka, A.J., Muller, C., Bull, C., Lehrach, H., and O'Brien, J. (1999) Large-scale clustering of cDNA-fingerprinting data. *Genome Research*, **9**, 1093-1105.

[20] Holter, N.S., Mitra, M., Maritan, A., Cieplak, M., Banavar, J.R., and Fedoroff, N.V. (2000) Fundamental patterns underlying gene expression profiles: simplicity from complexity. *Proc Natl Acad Sci USA.*, **97**(15), 8409-8414.

[21] Kamimura, R.T., Bicciato, S., Shimizu, H., Alford, J., and Stephanopoulos, G.N. (2000) Mining of Biological Data I: Identifying Discriminating Features via Mean Hypothesis Testing. *Metabolic Engineering*, **2**(3), 218-227.

[22] Kim, S., Dougherty, E.R., Chen, Y., Sivakumar, K., Meltzer, P., Trent, J.M., and Bittner, M. (2000) Multivariate Measurement of Gene Expression Relationships. *Genomics*, **67**(2), 201-209.

[23] Kramer, M.A. (1991) Nonlinear Principal Component Analysis Using Autoassociative Neural Networks. *AIChE J.*, **37**(2), 233-243.

[24] Mardia, K.V., Kent, J.T., and Bibby, J.M. (1979) *Multivariate Analysis*. Academic Press, San Diego.

[25] Michaels, G.S., Carr, D.B., Askenazi, M., Fuhrman, S., Wen, X., and Somogyi, R. (1998) Cluster analysis and data visualization of large-scale gene expression data. *Proc. Pacific Symposium on Biocomputing* **3**, 42-53.

[26] Pandin, M., and Didonè, G. (1996) Chaos in Information Processing: Simulation of a Biological Learning Process by Time Evolution in an Unsupervised Neural Network. *Int. J. of Bifurcation and Chaos*, **6**, 203-210.

[27] Perou, C.M., Jeffrey, S.S., van de Rijn, M., Rees, C.A., Eisen, M.B., Ross, D.T., Pergamenschikov, A., Williams, C.F., Zhu, S.X., Lee, J.C., Lashkari, D., Shalon, D., Brown, P.O., and Botstein, D. (1999) Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci USA.*, **96**(16), 9212-9217.

[28] Raychaudhuri, S., Stuart, J.M., and Altman, R.B. (2000) Principal components analysis to summarize microarray experiments: application to sporulation time series. *Proc. Pacific Symposium on Biocomputing*, **5**, 452-463.

[29] Rigoutsos, I., Floratos, A., Parida, L., Gao, Y., and Platt, D. (2000) The emergence of pattern discovery techniques in computational biology. *Metab. Eng.*, **2**(3), 159-77.

[30] Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986) Learning internal representations by error propagation. In Rumelhart, D.E., McClelland, J.L., and the PDP Research Group (Eds), *Parallel distributed processing: explorations in the microstructure of cognition, Volume I: Foundations*. MIT Press, Cambridge, MA, pp. 318-362.

[31] Spellman, P.T., Sherlock, G., Iyer, V.R., Zhang, M., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. (1998) Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast Saccharomyces cerevisiae by Microarray Hybridization. *Molecular Biology of the Cell*, **9**, 3273-3297.

[32] Stephanopoulos, G.N., Aristidou, A.A., and Nielsen, J. (1998) *Metabolic engineering – Principles and methodologies*. Academic Press, London.

[33] Stryer, L. (1988) *Biochemistry*. W. H. Freeman & C., New York.

[34] Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., and Golub, T.R. (1999) Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc Natl Acad Sci USA.*, **96**(6), 2907-2912.

[35] Toronen, P., Kolehmainen, M., Wong, G., and Castren, E. (1999) Analysis of gene expression data using self-organizing maps. *FEBS Letters*, **451**(2), 142-146.

[36] Wen, X., Fuhrmanm, S., Michaels, G.S., Carr, D.B., Smith, S., Barker, J.L., and Somogyi, R. (1998), Large-scale temporal gene expression mapping of CNS development. *Proc Natl Acad Sci USA.*, **95**(1), 334-339.

[37] Wong, E.T., Jenne, D.E., Zimmer, M., Porter, S.D., Gilks, C.B. (1999), Changes in the chromatin organization at the neutrophil elastase locus associated with myeloid cell differentiation, *Blood*, **94**, 3730-3736.

[38] Zhang, H., Yu, C.Y., Singer, B., Xiong, M. (2001), Recursive partitioning for tumor classification with gene expression microarray data. *Proc Natl Acad Sci USA.*, **98**(12), 6730-6735.