# Accepted Manuscript

Please cite this article as: Yang, E., Maguire, T. J., Yarmush, M. L., Berthiaume, F., & Androulakis, I. P., Identification of Regulatory Mechanisms of the Hepatic Response to Thermal Injury, *Computers & Chemical Engineering* (2007), doi:10.1016/j.compchemeng.2007.02.009

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Identification of Regulatory Mechanisms of the Hepatic Response to Thermal Injury

*E. Yang[1], T.J. Maguire[1], M.L. Yarmush[1], F. Berthiaume[2] and I.P. Androulakis[1≠]*

[1]Biomedical Engineering Department, Rutgers University, Piscataway, NJ, USA
[2]Center for Engineering in Medicine/Surgical Services, Massachusetts General Hospital, Harvard Medical School and the Shriners Hospitals for Children, Boston, MA, USA

**ABSTRACT**

The purpose of this paper is to evaluate the hypothesis that a systems biology approach can be developed such that a set of transcription factors, relevant to burn-induced inflammatory response can be identified and modulated to control the host response. We explore a novel method for identifying coherent and informative expression motifs and we subsequently determine conserved transcription factor binding sites for the sub-sets of co-expressed genes. The responses are rationalized in the context of burn induced inflammation and the putative transcription factors are rationalized in the context of intervention targets for controlling gene expression.

**INTRODUCTION**

Deep thermal injury over greater than 20 percent of the total body surface area is one of the most severe forms of trauma. Following the early acute phase response dealing with the initial injury and shock[1], there exists an equally serious secondary response which include changes in metabolism leading to hypermetabolism and catabolism, decreased function of the immune system, and sepsis[2]. Due to improvements in hospital care, more burn victims survive the acute response to the injury, and are faced with the secondary effects of thermal injury, which have proven to be more difficult to treat and control. In particular, prolonged sepsis and hypermetabolism following severe injury can result in MODS (Multiple Organ Dysfunction Syndrome), currently the most common cause of death in noncoronary intensive care units in the U.S.[3, 4]. A better understanding of the mechanisms by which the early responses to thermal injury predispose to the later hypermetabolic state would make it possible to define points of intervention by which such outcomes can be avoided.

The response to burn injury and trauma results from a complex interplay between inflammation caused by the initial injury and hypermetabolism. The advent of DNA microarrays enables to systematically examine expression changes of a very large number of genes, which provides an opportunity to identify pathways which have not been previously known as important. Given the importance to better understand the progression of the response after injury, temporal expression profiling in which the

---

≠ Corresponding author. E-mail: Yannis@rci.rutgers.edu

dynamics of mRNA expression are measured over time is more useful than simply the measurement of gene expression pre burn and post burn.

Changes in energy expenditure, increased glucose and lipid turnover are hallmarks of the hypermetabolic state [5]. Given that the liver plays a key role in metabolic processes of energy production, gluconeogenesis and lipid synthesis and oxidation, it is likely that the liver is one of the primary organs driving the systemic response to severe thermal injury[6]. Furthermore, a better understanding of the liver response to thermal injury may provide mechanistic insights and suggest new ways in which the hypermetabolism associated with severe thermal injury could be mitigated or even prevented.

To use temporal expression data from microarrays to formulate possible compensatory strategies the following steps must be accomplished:

1. Selection and Classification of Relevant Genes
2. Functional Characterization of extracted genes
3. Generation of Hypothetical Regulatory Networks
4. Associating between Functional Characterization and Regulatory Networks

We will propose an integrative method that combines temporal gene expression profiling, sequence analysis, database mining and network construction in order to formulate the parameters for a future experiment to better understand the response of an organism to severe burn injury. It is our goal to place the bioinformatics approach into an iterative formalism where experimental data guides data analysis which in turn guides the course of future experiments.

**METHODS**

Prominent in engineering analysis is the concept of developing appropriate mathematical representations of the response of the system in order to identify critical components and optimal strategies [7, 8]. Biological systems, much like reaction engineering systems are characterized by input, output, control variables and an underlying dynamic model that propagates disturbances due to input variability across the system. These are manifested in the response of the output variables. The goal of an engineering analysis is to identify the control points that modulare the response. An additional complexity of biological systems is that, by and large, fundamental first principles models that describe these dynamic responses are not available. Therefore, data-based methods are needed to develop appropriate models and dependencies. This is a well accepted paradigm in chemical engineering that has generated significant success. In order to enable the development of such descriptive models a number of critical questions need to be answered. Specifically:
1. we need to identify the inherent dynamic of the system and the informative output variables that best characterize that response
2. we need to identify the controls that modulate the observed responses, and finally

3. we need to determine the functional relationships that dictate how inputs propagate across the systems and how outputs are modulated by the controls.

In the sections that follow we identify these steps and demonstrate how such models can be generated and analyzed to identify and propose intervention strategies that modulate the systemic response.

*Experimental Data*

In a previously published study, male Sprague-Dawley rats were subjected to a cutaneous 3rd degree burn injury consisting of a full skin thickness scald burn of the dorsum, calculated to be ~20% of the rat's total body surface area[9]. Liver samples were obtained at 5 time points (0, 1, 4, 8, and 24h post burn). RNA extracted from the extracted livers was isolated and subsequently hybridized to a U34A GeneChip that had 8,799 probes represented on each chip.  The control for this experiment was obtained at time 0, which was prior to the injury. It has been previously shown that time had no significant effect upon the response of rats to the sham treatment [10].

*Selection and Classification of Relevant Genes*

One of the strengths of microarrays is the fact that they are able to measure the levels of gene expression of thousands of genes at once. Therefore, the researcher is able to measure the expression levels of genes whose role in a given biological phenomenon was previously unknown. However, this strength is problematic for researchers because it then becomes difficult to determine which gene expression levels are important to the organism's response to an external stimulus.

Most automated techniques such as dChip, MAS5, PMMM, ANOVA, t-tests, and RMA attempt to extract genes from microarrays based upon the notion that genes which show statistically significant changes in gene expression levels ought to be informative[11]. The primary drawback with these methods is that they look for genes whose differential expression is statistically significant rather than evaluating whether or not a given gene is relevant to the experiment at hand. The difference between a statistically significant gene and a gene which is relevant to the experiment is a subtle but important one. This difference arises due primarily to the dynamic nature of homeostasis. Even in an unperturbed state there is significant transcriptional activity taking place within the organism regulating events such as feeding, resting, or physical activity. Therefore, even in a control state in which the organism has not been experimentally perturbed, there ought to be gene expression profiles which change at a statistically significant level. However, what researchers should be interested in is the identification of gene expression profiles that are directly responding to the experimental perturbation and not just genes that show a statistically significant dynamic.

Adjunct to this concern is the fact that the currently established filtering techniques are essentially answering the wrong question. All of these techniques were derived in

response to two state microarray experiments in which the primary question was, "Does the gene expression of a given gene change significantly between the two states." However, with temporal expression microarray data the relevant question should be, "Is the shape of the temporal expression profile accurate."

SLINGSHOTS(SeLection of Informative Genes via Symbolic Hashing of Time Series) was developed specifically to isolate genes that could be hypothesized to be important to the underlying response of an organism. The fundamental hypothesis which underlies SLINGSHOTS is that in response to an external perturbation, there ought to be a coordinated set of genes whose expression profiles are highly correlated. In addition to this, this set of genes in aggregate ought to illustrate significant deviations from the baseline distribution of expression levels. By identifying these genes, one ought to be able to obtain the genes whose expression profiles are representative of the underlying changes. The prime innovation of this technique is that the selection criteria uses a global metric to assess the "informativeness" of the isolated genes by evaluating the set in aggregate rather than the local metrics used by other selection techniques which checks the informativeness of each genes individually. The details of the method were recently discussed in [12].

SLINGSHOTS is broken down into two related steps, a hashing step in which the genes are clustered into a large number of highly correlated clusters, and a selection step in which a set of these clusters is evaluated for their ability to represent the experimental perturbation. The behavior of SLINGSHOTS is defined by two primary parameters, $\alpha$ and w, of which $\alpha$ is the size of the alphabet used and w is the number of time points to average together for longer time series. In this evaluation, $\alpha$ was selected to be 4 and w was selected to be 1. Given that the clustering and selection are combined, the results of the selection are already clustered negating the need for a separate clustering step.

The result of the hashing step is a large set of gene clusters of which all of the gene expressions show a correlation coefficient to the average profile above a certain threshold. Unlike in QT clustering where the threshold is set explicitly, SLINGSHOTS sets this cutoff through a combination of w and alpha. In this case, the minimum correlation coefficient for any of the given clusters is greater than .75.

The identification of clusters that comprise up of the hypothetical primary response of the organism to thermal injury then allows for further analysis in terms of the functional role of the genes and the identification of the regulatory mechanisms which give rise to the observed expression profiles. The following steps will allow for the identification of possible mechanism with which to mediate the undesired responses associated with severe thermal injuries.

*Identification of Functional Ontologies*

The purpose behind the identification of functional ontologies is the determination of the underlying biological processes that are related to the phenomenon being investigated. This essentially allows for the interpretation of the biological significance of isolated

genes. This will be conducted by isolating the ontologies which are related to the selected genes via the database present at www.geneontology.org[13]. The primary assumption behind this analysis is that co-expressed genes ought to have related functionality [14]. Therefore, by looking at an aggregate set of clustered genes, it is expected that there should preference of certain biological process ontologies to be localized to a specific cluster.

To determine which of the isolated processes are significant, it is important to determine which of the associated ontologies are enriched at a statistically significant level ( $p <$ .05). Due to the fact that many genes can participate in different biological processes especially those associated with cellular signaling, it is important to determine which of these biological processes are consistent over all of the selected and clustered genes. This is done utilizing the hypergeometric distribution. The hypergeometric distribution however, is inaccurate if the total number of counts is less than five, a more accurate assessment of the count is given by the Fisher distribution[13, 15]. A secondary benefit of such analysis is that it allows for an ad hoc evaluation as to the correctness of selection and clustering. If the genes that were isolated and classified do not show any notable enrichment, then it would suggest that there was a flaw in the either the data or the methodology.

*Regulatory Network Construction*

The dynamic response of biological organisms is governed by a large interconnected network which ties the response of each gene to intercellular conditions or the expression levels of other genes. Part of the rationale behind conducting temporal gene expression experiments is that the measured dynamics can help in the construction of a transcriptional network which gives insights as to how an organism responds to external stimulus. We make the additional assumption that by grouping genes into co-expressed clusters, the network is simplified through the assumption that these co-expressed genes are regulated by the same mechanism. This prunes many of the connections thereby simplifying the overall network. This network can be further simplified by treating it as a bi-partite graph **Figure 1**, in which a set of inputs drives a set of output. It does not explicitly model the existence of feedback loops. However, feedback loops can be handled if an input gene is also present in the output. For example, if a transcription factor such as GATA6 were driving the system, it could also be present in the output. Therefore, despite the fact that a bi-partite graph is a simplification of the overall network interactions, it is possible to retrieve the original DAG (Directed Acyclic Graph) from a bipartite network[16]. The primary benefit of treating our gene regulatory network as a bi-partite graph, is that there are robust methods for estimating the connectivity strengths from gene expression data such as NCA (Network Component Analysis)[17, 18].

Various methods exist for the construction of transcriptional networks from gene expression data. These techniques fall broadly under two primary categories of algorithms, those which seek to identify relationships between the expression level of different genes over multiple conditions[19, 20], and those which utilize outside information to generate regulatory network such as using the set of gene regulators which

have been previously identified, or analyzing the promoter region to look for possible links between genes and their regulators[21-23]. Given the form of our data which consists of a single condition and five time points, the construction of the network requires external data in the form of predicted transcription factor binding sites. Due to the relative dearth of data, it is imperative for us to localize our analysis to a small portion of the regulatory network, namely the genes and regulators that are directly responding to the input perturbation, and hence make the gene selection process discussed previously an integral part of the analysis.

The network associated with temporal gene expression data is the gene regulatory network in which the primary links are the transcription factors whose activity mediates the production of the genes. Transcription factors are proteins that bind to the promoter region of a given gene and through that activity can either up or down-regulate the gene expression of that gene. Given that transcription factors themselves are genes, they are also regulated by other transcription factors and sometimes by a protein product further down in the signaling cascade which they initiate[24]. In yeast, many of the transcription factors have been experimentally identified. However, in more complex mammalian systems such as rat, there is a limited number of experimental information forcing the use of algorithms which predict transcription factor binding sites. These transcription factors binding site prediction algorithms fall under two general categories, algorithms such as AlignACE[21] which work by looking for over-expressed motifs (n-mers) within the promoter region of a cluster of genes, and those like CONSITE[25], which base their predictions off transcription factor position weight matrices which are obtained experimentally through procedures such as SELEX[26]. The latter method was chosen for this analysis due to the relative difficulty in associating the over-expressed motifs to transcription factors whereas the second method is based off of known transcription factors. The specific tool used for the determination of possible transcription factor binding sites was CORG(COmparative Regulatory Genomic)[27], which is an online tool that is able to extract the promoter sequence from homologous genes between two organisms and obtain the associated transcription factor binding sites. The one piece of information which we utilized was simply whether or not a possible connection existed between a transcription factor an its associated binding site.

**RESULTS**

Out of the original 8799 probes, the algorithm has identified 281 probes corresponding to 208 known genes located in four clusters **Appendix 2** of which the z-score normalized expression profiles are given in **Figure 2**. The transcriptional state of these genes over the experimental time course is given in **Figure 3**, and what is clearly evident is that these genes illustrate a two-wave response to the initial burn injury. At Hour 1, a large disjoint in the transcriptional levels can be seen in where there is a large amount of activity with all of the genes either significantly up-regulated or down-regulated. At Hours 4 and 8, the expression profiles are fairly close to the pre-injury profile, suggesting a return to the initial homeostatic state. Then, at Hour 24, an even greater disjoint representing a major shift in the cellular transcriptional state is visible, which is associated with an abundant over-expression of the inflammation marker and acute phase protein A2M (Alpha 2

Macroglobulin). Elucidating the mechanism that gives rise to this more delayed response would have significant implications in the treatment of severe thermal injury.

In the transformation from gene expression profile to hash value by SLINGSHOTS[28], the z-score normalization was used Therefore scale information is discounted when forming the clusters. Commonly used selection metrics such as n-fold changes or t-test require the preservation of scale information within the individual expression profiles. Since this information is eliminated during the transformation, the selection of incorrect genes was a concern. If for the most part, the selected genes are part of co-regulated processes and the clusters that were selected, there ought to be a distinct separation in the functional ontologies between the different clusters. The ontology localization is evident in **Figure 4**. It is clearly evident that the ontologies are indeed localized to the cluster, and despite the fact that Clusters 2 and 3 seem to differ by only one time point after normalization, they do encapsulate two different sets of functions.

Taking the set of genes as a whole instead of as four clusters, we find that the most significant biological processes revolve around metabolism, inflammation, protein production, and signaling which is in agreement with the macroscopic observations of pathophysiology after severe thermal injury. The localization of functional ontologies to each cluster suggests to us that our clustering is indeed correct and the selection of significant ontologies related to metabolism, inflammation and protein catabolism and synthesis suggests that the selection process was likewise successful. However, while significant processes and their underlying dynamics have been identified, there still remains an open question as to what the best way to mediate the response of the system.

The most prevalent transcription factors amongst all of the clusters are STAT5/STAT6 associated with the JAK-STAT pathway, and TEF1 (Translation Elongation Factor). While there are genes which are part of the JAK-STAT pathway expressed in Cluster 4 such as erythropoietin, the JAK-STAT transcription regulation pathway is not particularly informative due to its widespread use in cellular signaling. In fact, looking at a random selection of genes, the JAK-STAT pathway is found to also be highly prevalent.

The only clusters with a consistent set of transcription factors binding sites that were not STAT5, STAT6, and TEF1 amongst the genes of that cluster (greater than 95% of identified genes containing a transcription factor binding site) were Clusters 2 and 4. It was somewhat surprising that we weren't able to find a small set of transcription factors that regulated all of the clusters. However, it was noticed that the clusters that contained a significant set of genes that coded for transcription factors had a consistent set of activators, while the clusters of genes that did not contain transcription factors (Cluster 1 and 3) were not regulated by transcription factors other than STAT5 and STAT6. The associated transcription factors are given in **Table 1**.

The primary activators that we found were AP2 Alpha, GATA 6, and CIZ. AP2 Alpha was localized to Cluster 2, while GATA 6 and CIZ were co-localized to Cluster 4. While Cluster 1 and Cluster 3 did not have a set of consistent transcriptional regulators, they did

however, have a large fraction of genes that were regulated by the Octamer Binding Family (OCT) of transcription factors and genes that respond to myogenin (MYOD) both of which are transcription factors that were present in Cluster 2 and 4. This suggests that there are some elements of a signaling cascade within the transcriptional regulatory network. A gross overview of the transcriptional network is given in **Figure 5.**

**Discussion**

From the results of the clustering, we find that the two wave phenomenon can be rationalized by the expression dynamics of Cluster 1 and Cluster 4, in which Cluster 1 remains constant up until Hour 24 in which there is a large increase in the level of gene expression. Cluster 4 is indicative of the early response showing significant up-regulation during this period, and a relaxation afterwards with significant overshoot. Clusters 2 and 3 are similar in terms of their response except at Hour 1 with Cluster 1 remaining constant up until that time and Cluster 3 being down regulated at Hour 1.

Characteristic of the genes that show activity early on such as Cluster 3 and 4, we find that there is a significant over-representation of genes that code for transcription factors, regulate metabolism, and protein production within these clusters **Appendix 2**. We hypothesize that the initial thermal injury causes the change in the expression levels of Clusters 3 and 4, which affect the dynamics of Cluster 1 and 2, either directly through the production of transcription factors or indirectly by altering the levels of circulating metabolites  through changes in metabolism of macromolecules such as proteins, carbohydrates and fatty acids.

Coupled with the transcriptional driving forces found within these genes, are associated metabolic processes associated with catecholamines which are important in the hypermetabolic response seen after thermal injury[5, 29].  Although the hypermetabolic response is known to occur 3 days or more after injury, the presence of these metabolic genes in **Appendix 2** suggests that there is a significant metabolic component early on in the response of the burn injury. It is currently unclear as to whether these early changes in liver metabolism and corresponding changes in levels of metabolites in the circulation play a role in the subsequent more systemic and chronic changes in metabolism and the inflammatory response, events that lie outside of the experimental time frame analyzed herein. There is evidence that the sustained inflammation leads to inhibition of transporters leading to abnormal levels of circulating metabolites such as lipids and glucose[30, 31], which may affect the other organs in the organism.

As part of the early response to thermal injury, a large portion of the genes appears to be either metabolic in nature or involved in transcriptional signaling. Cluster 4, which has an immediate response at Hour 1, contains genes that are responsible for the metabolism of fatty acids. This is paired with the activity in Cluster 2 in which a major component of its activity is a corresponding down-regulation in fatty acid transport. It has been noted in previous work that gene expression levels often do not correlate well with the levels of their corresponding metabolites in circulation[9]. This opposing dynamics of the fatty

acid transporters and the fatty acid enzymes suggests that the levels of fatty acids are controlled by competing transport and metabolic processes. Coupled with the changes in metabolism are a series of transcriptional changes. What we find is that within these early response clusters, there are changes in expression of murinoglobulin 1, and complement component 5. These regulate key inflammatory and acute phase responses and work to dampen the long term inflammatory response of the genes found in Clusters 1 and 2[32, 33].

The longer term responses shown by Clusters 1 and 2 seem to revolve around a transcriptional signaling component and a protein catabolism component. Given the localization of protein catabolism along with inflammation in Cluster 1, we believe that the increased protein turnover rate may be one of the primary driving forces leading to the hypermetabolic state. So, while there are undoubtedly changes in the energetics of an organism after thermal injury, we believe that long-term changes in energetics manifest themselves primarily in the levels of protein turnover rate. Given the severe fall in ATP levels[9] post burn, we believe the organism is making up for a significant shortfall in available energy through the catabolism of protein. This counter-productive process may be one of the significant barriers to recovery from thermal injury.

While there is a significant metabolic component to the burn response, there is also a significant role played by transcriptional signaling pathways. The regulatory network which we were able to infer from a combination of gene expression data and promoter region analysis has suggested 3 possible initiators of the burn response GATA 6, AP2 Alpha, and CIZ. We believe that the identification of these initiatory transcription factors are not wholly unreasonable. These three transcription factors have been cited in functions related to inflammation[34-36]. Unlike other factors such as the ubiquitous JAK-STAT pathway, these transcription factors offer the best hope of altering individual burn injury-induced responses independently.

While Clusters 2 and 4 have a possible set of regulators with which their responses can be perturbed, Clusters 1 and 3 do not have a readily apparent set of regulators. It is possible that the co-expression of the genes within Clusters 1 and 3 is due primarily to the co-expression of their regulators, meaning that the genes in Clusters 1 and 3 may be co-expressed not because they have precisely the same regulators, but because the activity of their individual regulators is similar. Looking at Clusters 2 and 4, we find a set of regulators which may not bind to the same recognition sequence but may have very similar responses, of which the most notable ones are MYOG and the POU family of transcription factors which are known to be important regulators in liver. MYOG is present in more than 35% of the genes found in Clusters 1,2,3. This is notable given the relative long length of the MYOG recognition sequence which is 29 base pairs long. Finding it in such a large number of promoter regions is highly statistically significant given the fact that the expected hit rate of MYOG is 1 out of $5.21 \times 10^{12}$ bases. What is even more notable is that MYOG is found in many of the genes which are responsible for metabolism, including catecholamine metabolism. The transcription factor family POU seems to be more biased towards the genes which code for calcium and potassium ion channels. Given that the transcription factor MYOG is highly significant in terms of the

number of matches, and the fact that it appears to regulate genes related to metabolism rather than the genes associated with inflammation, we hypothesize that by altering the activity of MYOG, we may be able to affect the severity of the hypermetabolism while having a minimal impact upon the inflammatory respose.

**CONCLUSIONS**

The role of bioinformatics is not to look for answers independently of experimentation, but rather to look for the basis of new experiments. By integrating the experimental gene expression data with genomic data and the results of SELEX experiments, we were able construct a rough network which gave hints as to possible points of intervention. Seeking to prevent the induction of the system into hypermetabolism, we have identified the myogenin transcription factor as perhaps one of the critical signals which drives the system from its acute response to burn injury to the longer term systemic hypermetabolic state.

Identifying a regulatory layer and the core nodes of that layer provides a mechanism to elucidate intervention points to attenuate the inflammatory process. Intervention utilizing these TF proteins could theoretically take one of three forms: 1) inhibition of TF production using knockout or silencing techniques; 2) blocking TF activity through competitive inhibition; 3) blocking TF activity through suicide inhibition. Promising approaches for silencing focus on the use of siRNA techniques [37]. In this approach double-stranded RNA (dsRNA) is digested by the dsRNA-specific RNase III enzyme dicer into small interfering RNAs (siRNAs). The siRNAs then assemble with a multiprotein nuclease complex, RNA-induced silencing complex [38], which unwinds the dsRNAs and degrades target mRNAs homologous to the single stranded siRNA in a sequence-specific manner. The result of this process is the degradation of mRNA needed as a template for protein production, thereby inhibiting the production process, and depleting pools of proteins needed for specific enzymatic reactions. One specific example of siRNA utilized for intervention in inflammatory response is the application of siRNA techniques to inhibition the production of STAT-3 in order to elucidate key signaling molecules in the inflammatory response pathway [39]. Therefore, a key advantage of the methodology discussed in this work is the systematic identification of several putative regulatory proteins of the inflammatory response, thus enabling the rational selection of multiple targets and design of combination therapies for the modulation of the inflammatory response.

**BIBLIOGRAPHY**

1. Carter EA, Tompkins RG, Yarmush ML, Walker WA, Burke JF: **Redistribution of blood flow after thermal injury and hemorrhagic shock.** *J Appl Physiol* 1988, **65:**1782-1788.
2. Strock LL, Singh H, Abdullah A, Miller JA, Herndon DN: **The effect of insulin-like growth factor I on postburn hypermetabolism.** *Surgery* 1990, **108:**161-164.
3. Tredget EE, Yu YM: **The metabolic effects of thermal injury.** *World J Surg* 1992, **16:**68-79.
4. Cooper AB, Cohn SM, Zhang HS, Hanna K, Stewart TE, Slutsky AS: **Five percent albumin for adult burn shock resuscitation: lack of effect on daily multiple organ dysfunction score.** *Transfusion* 2006, **46:**80-89.
5. Wilmore DW, Long JM, Mason AD, Jr., Skreen RW, Pruitt BA, Jr.: **Catecholamines: mediator of the hypermetabolic response to thermal injury.** *Ann Surg* 1974, **180:**653-669.
6. Mager DR, Wykes LJ, Roberts EA, Ball RO, Pencharz PB: **Mild-to-moderate chronic cholestatic liver disease increases leucine oxidation in children.** *J Nutr* 2006, **136:**965-970.
7. Biegler LT, Grossmann IE, Westerberg AW: *Systematic Methods of Chemical Process Design.* Prentice Hall; 1997.
8. Aris R: *Mathematical Modeling : A Chemical Engineer's Perspective (Process Systems Engineering).* Academic Press; 1999.
9. Vemula M, Berthiaume F, Jayaraman A, Yarmush ML: **Expression profiling analysis of the metabolic and inflammatory changes following burn injury in rats.** *Physiol Genomics* 2004, **18:**87-98.
10. Lee K, Berthiaume F, Stephanopoulos GN, Yarmush ML: **Profiling of dynamic changes in hypermetabolic livers.** *Biotechnol Bioeng* 2003, **83:**400-415.
11. Millenaar FF, Okyere J, May ST, van Zanten M, Voesenek LA, Peeters AJ: **How to decide? Different methods of calculating gene expression from short oligonucleotide array data will give different results.** *BMC Bioinformatics* 2006, **7:**137.
12. Yang E, Maguire T, Yarmush ML, Berthiaume F, Androulakis IP: **Bioinformatics analysis of the early inflammatory response in a rat thermal injury model.** *BMC Bioinformatics* 2007, **8:**10.
13. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25:**25-29.
14. Wolfe CJ, Kohane IS, Butte AJ: **Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks.** *BMC Bioinformatics* 2005, **6:**227.
15. Draghici S, Khatri P, Martins RP, Ostermeier GC, Krawetz SA: **Global functional profiling of gene expression.** *Genomics* 2003, **81:**98-104.
16. Liu B, Lai H-J: *Matrices in combinatorics and graph theory.* Dordrecht ; Boston: Kluwer Academic Publishers; 2000.

17.   Liao JC, Boscolo R, Yang YL, Tran LM, Sabatti C, Roychowdhury VP: **Network component analysis: reconstruction of regulatory signals in biological systems.** *Proc Natl Acad Sci U S A* 2003, **100:**15522-15527.

18.   Boulesteix AL, Strimmer K: **Predicting transcription factor activities from combined analysis of microarray and ChIP data: a partial least squares approach.** *Theor Biol Med Model* 2005, **2:**23.

19.   Kauffman S, Peterson C, Samuelsson B, Troein C: **Random Boolean network models and the yeast transcriptional network.** *Proc Natl Acad Sci U S A* 2003, **100:**14796-14799.

20.   Dojer N, Gambin A, Mizera A, Wilczynski B, Tiuryn J: **Applying dynamic Bayesian networks to perturbed gene expression data.** *BMC Bioinformatics* 2006, **7:**249.

21.   Manson McGuire A, Church GM: **Predicting regulons and their cis-regulatory motifs by comparative genomics.** *Nucleic Acids Res* 2000, **28:**4523-4530.

22.   Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM: **Systematic determination of genetic network architecture.** *Nat Genet* 1999, **22:**281-285.

23.   Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N: **Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data.** *Nat Genet* 2003, **34:**166-176.

24.   Ayte J, Schweitzer C, Zarzov P, Nurse P, DeCaprio JA: **Feedback regulation of the MBF transcription factor by cyclin Cig2.** *Nat Cell Biol* 2001, **3:**1043-1050.

25.   Sandelin A, Wasserman WW, Lenhard B: **ConSite: web-based prediction of regulatory elements using cross-species comparison.** *Nucleic Acids Res* 2004, **32:**W249-252.

26.   Klug SJ, Famulok M: **All you wanted to know about SELEX.** *Mol Biol Rep* 1994, **20:**97-107.

27.   Dieterich C, Wang H, Rateitschak K, Luz H, Vingron M: **CORG: a database for COmparative Regulatory Genomics.** *Nucleic Acids Res* 2003, **31:**55-57.

28.   Yang E, Berthiamume F, Yarmush ML, Androulakis IP: **An Integrative Systems Biology Approach for Analyzing Liver Hypermetabolism.** *Proceedings of the Joint 9th International Symposium, Processing Systems Engineering and 16th European Symposium* 2006.

29.   Herndon DN, Hart DW, Wolf SE, Chinkes DL, Wolfe RR: **Reversal of catabolism by beta-blockade after severe burns.** *N Engl J Med* 2001, **345:**1223-1229.

30.   Khovidhunkit W, Moser AH, Shigenaga JK, Grunfeld C, Feingold KR: **Endotoxin down-regulates ABCG5 and ABCG8 in mouse liver and ABCA1 and ABCG1 in J774 murine macrophages: differential role of LXR.** *J Lipid Res* 2003, **44:**1728-1736.

31.   Savage DB, Petersen KF, Shulman GI: **Mechanisms of insulin resistance in humans and possible links with inflammation.** *Hypertension* 2005, **45:**828-833.

32.   Northemann W, Shiels BR, Braciak TA, Fey GH: **Structure and negative transcriptional regulation by glucocorticoids of the acute-phase rat alpha 1-inhibitor III gene.** *Biochemistry* 1989, **28:**84-95.

33.   Arumugam TV, Woodruff TM, Stocks SZ, Proctor LM, Pollitt S, Shiels IA, Reid RC, Fairlie DP, Taylor SM: **Protective effect of a human C5a receptor**

antagonist against hepatic ischaemia-reperfusion injury in rats.** *J Hepatol* 2004, **40:**934-941.

34.     Lin L, Spoor MS, Gerth AJ, Brody SL, Peng SL: **Modulation of Th1 activation and inflammation by the NF-kappaB repressor Foxj1.** *Science* 2004, **303:**1017-1020.

35.     Gille J, Swerlick RA, Caughman SW: **Transforming growth factor-alpha-induced transcriptional activation of the vascular permeability factor (VPF/VEGF) gene requires AP-2-dependent DNA binding and transactivation.** *Embo J* 1997, **16:**750-759.

36.     Torrungruang K, Alvarez M, Shah R, Onyia JE, Rhodes SJ, Bidwell JP: **DNA binding and gene activation properties of the Nmp4 nuclear matrix transcription factors.** *J Biol Chem* 2002, **277:**16153-16159.

37.     Lee LK, Roth CM: **Antisense technology in molecular and cellular bioengineering.** *Curr Opin Biotechnol* 2003, **14:**505-511.

38.     Wu MS, Kwon S, Driscoll JF, Faeth GM: **Preferential diffusion effects on the surface structure of turbulent premixed hydrogen/air flames.** *Combust Sci Tech* 1991, **78:**69-96.

39.     Verga Falzacappa MV, Vujic Spasic M, Kessler R, Stolte J, Hentze MW, Muckenthaler MU: **STAT-3 mediates hepatic hepcidin expression and its inflammatory stimulation.** *Blood* 2006.
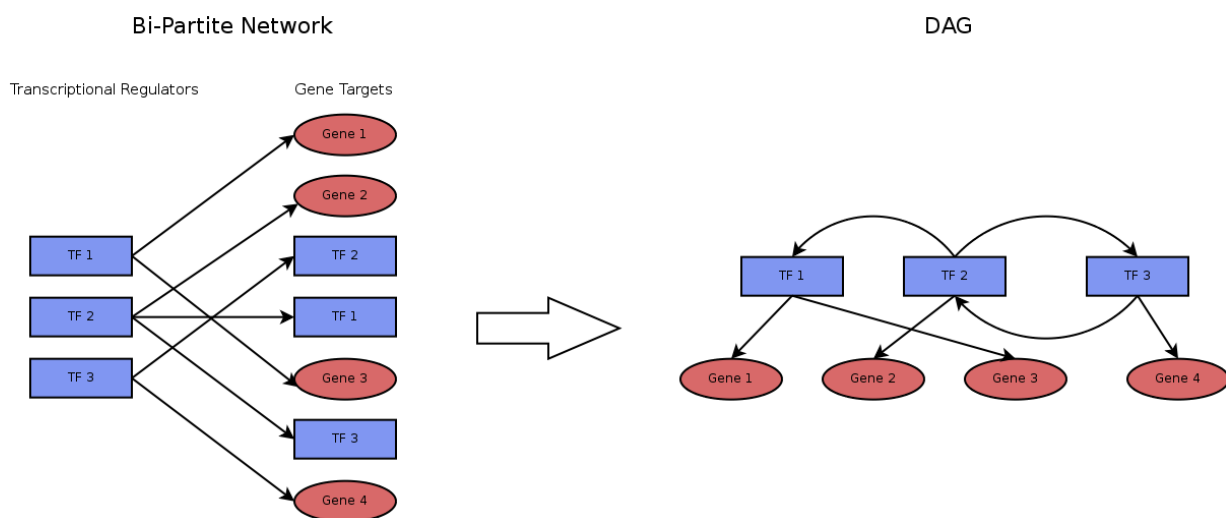
**Figure 1:** A bi-partite representation of a transcriptional network and its associated DAG. There is no loss in generality in terms of the possible networks that can be represented. The representation as a bi-partite network however allows for efficient quantification of the network through various algorithms such as NCA and PLS.
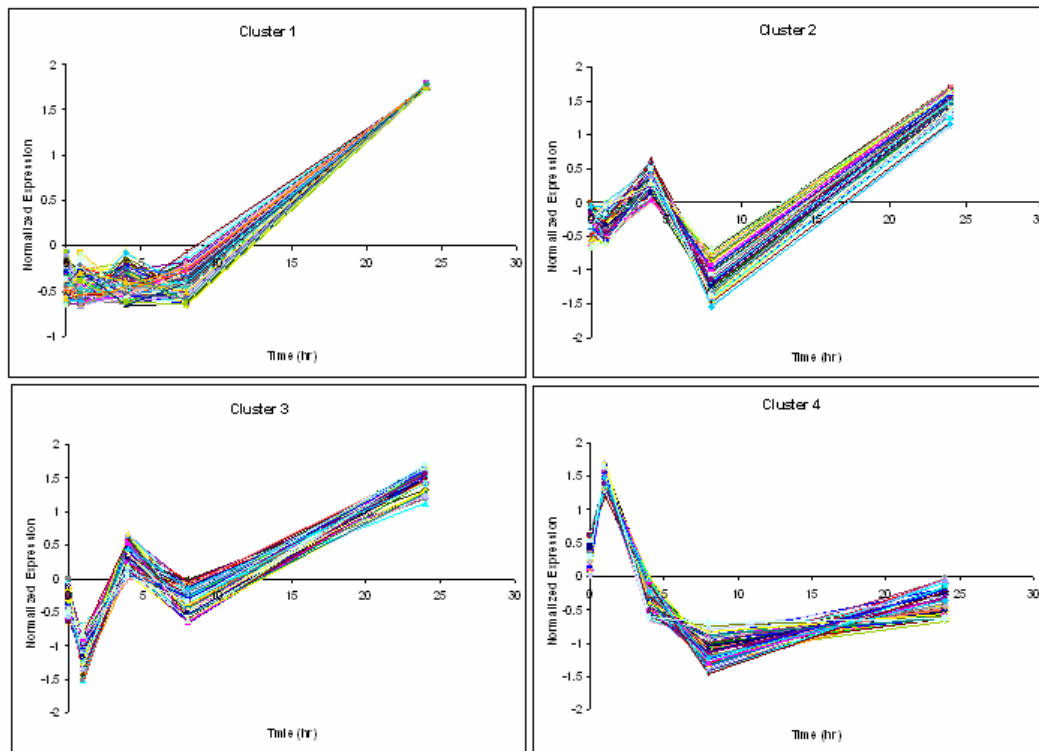
**Figure 2:** The expression profiles of the selected genes. Clusters 3 and 4 have an early phase response, while Cluster 1 and 2 primarily have a late term response
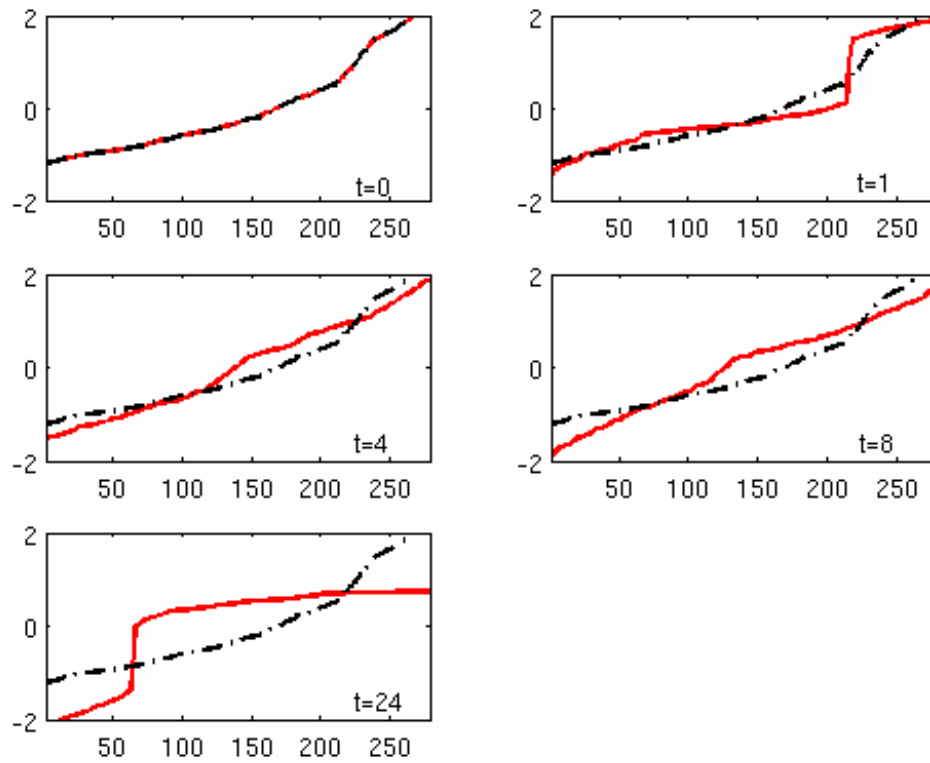
**Figure 3:** The transcriptional state of the selected genes. At t = 1 and t = 24, we have evidence of a 2 wave effect in which significant transcriptional processes are being altered. The response at t=1 is evidence of the short term compensatory mechanism, while the response at t = 24 represents a, potentially irreversible, state change into the chronic inflammation and hypermetabolic state.
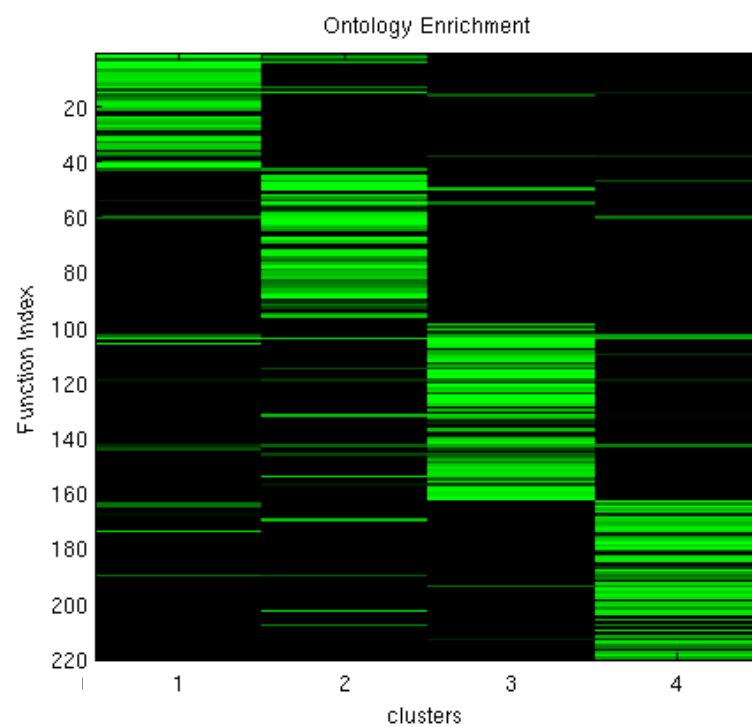
**Figure 4:** The enrichment of ontologies associated with the clustered and selected genes. The diagonally dominant nature of the graph suggests that our clustering has indeed separated out genes with related functionalities despite superficial similarities in their overall shape.
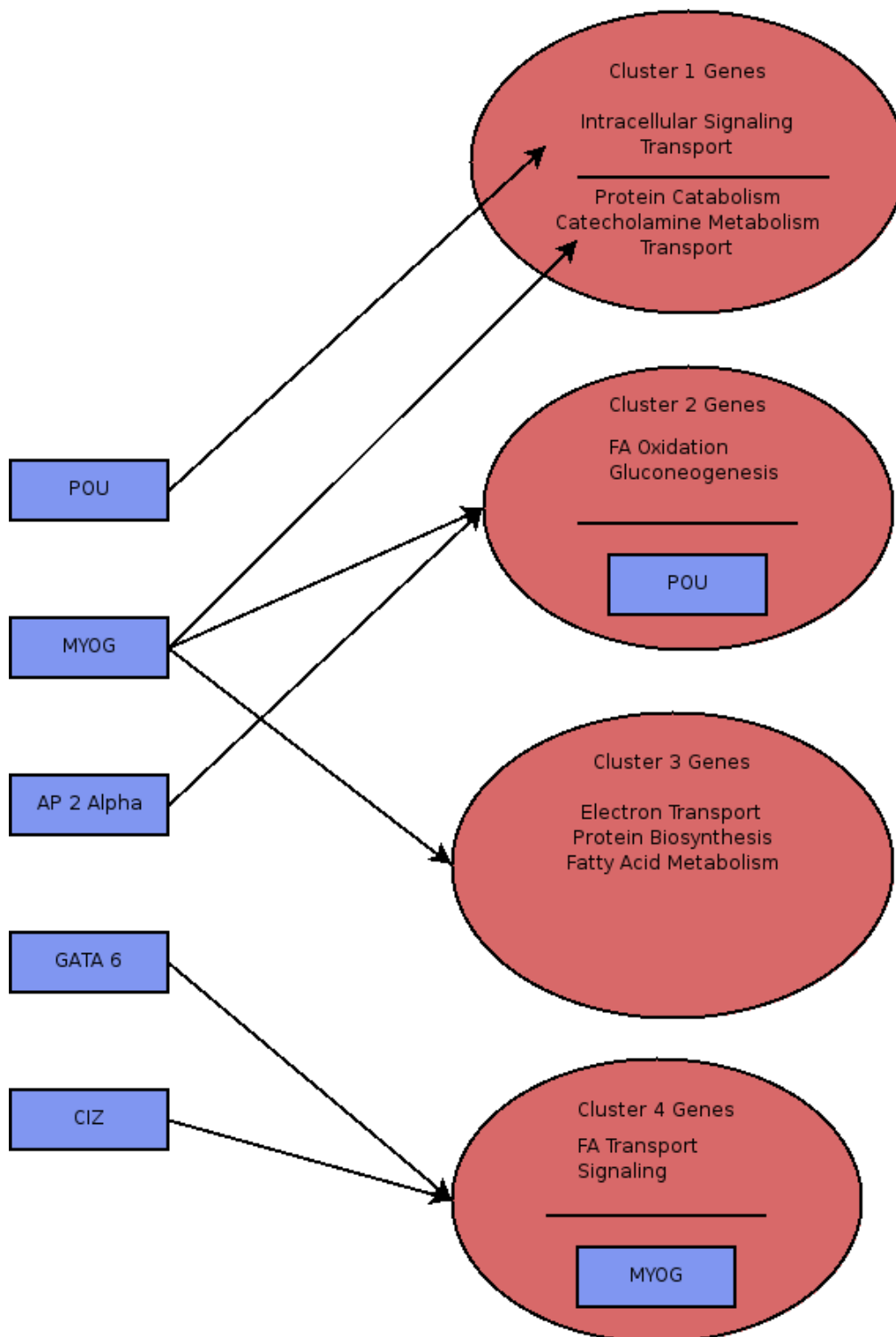
**Figure 5:** A gross schematic of the predicted response mechanisms. The notable thing is that MYOG which is up-regulated early is a regulator in a significant number of genes in Cluster 1 and 2, and may drive the secondary response.

| Cluster | Associated Transcription Factors |
|---------|----------------------------------|
| 1 | STAT5, STAT6, TEF1 |
| 2 | AP2-Alpha |
| 3 | STAT5, STAT6 |
| 4 | TEF1, STAT5, STAT6, CIZ, CDXA, GATA6, AP2-Alpha |

**Table 1: Associated transcription factors. Those highlighted in orange are transcription factors that are highly conserved, but not found**

**APPENDIX 1**

1. Normalization of the gene expression profile to N(0,1) via the z-score transform.

2. If the sequences are longer than 10 time points, piecewise averaging is conducted, i.e. averaging together sets of n time points to reduce the exponential expansion of the search space. In the case of our data, the 17 time points are interpolated to 18 time points, and the time series are broken down into sets of 2 to be piecewise averaged

3. These piecewise averaged points are then converted into symbols through the use of Gaussian breakpoints. Gaussian breakpoints are divisions in the Gaussian distribution such that the cumulative probability of each section are equivalent. These can be obtained through the use of CDF tables found in statistics text books or by solving the following equation for b:

$$\frac{i}{k-1} = \frac{1}{2}\left[1 + erf\left(\frac{b}{\sqrt{2}}\right)\right];$$

$i = 1,..,k; k$ = number of breakpoints; b = breakpoint value

The overall process of assigning a letter to each piecewise averaged point is illustrated in below:
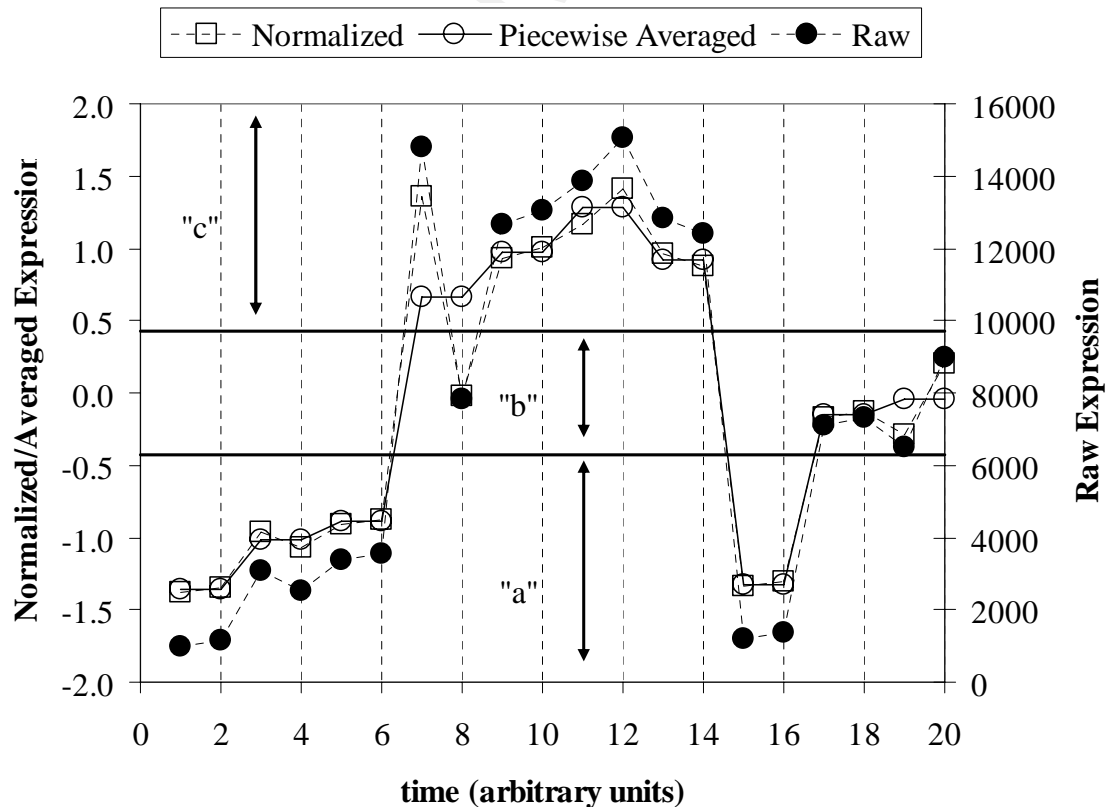
4. After the symbolic transformation, the series of symbols is converted into a single integer via the formula:

$$hash(c, w, a) = 1 + \sum_{j=1}^{w} [ord(c_j) - 1] \times a^{j-1}$$

Where c is the letter assigned to each piecewise averaged point, a is the size of the alphabet(3), and w is the total length of the expression profile divided by the number of points per piecewise average (2). The parameters of the alphabet were selected to so that the population distribution of motifs is non-exponential, to reflect the non-random distribution of expression profiles present in the data. w was chosen to preserve as much of the high frequency component of the signal as possible.

**APPENDIX 2**

Informative genes, class membership and ontology annotation.

| Gene ID | Gene | Cluster | Function |
|---|---|---|---|
| AI030286 | brain derived neurotrophic factor | 1 | Apoptosis |
| AA859878 | ret proto-oncogene | 1 | Cell Cycle Regulation |
| AI234604 | heat shock protein 8 | 1 | Cell Cycle Regulation |
| U60416 | myosin 5B | 1 | Cytoskeleton Organization |
| AI171243 | replication protein A3 (predicted) | 1 | DNA Repair |
| Z50084 | Ameloblastin | 1 | ECM Organization |
| AA874941 | Adipose differentiation-related protein | 1 | Fatty Acid Transport |
| AA819776 | heat shock 90kDa protein 1, alpha-like 3 (predicted) | 1 | Heat Shock |
| M23566 | alpha-2-macroglobulin | 1 | Inflammation |
| AA900582 | alpha-2-macroglobulin | 1 | Inflammation |
| X13983 | alpha-2-macroglobulin | 1 | Inflammation |
| M22670 | alpha-2-macroglobulin | 1 | Inflammation |
| M22670 | alpha-2-macroglobulin | 1 | Inflammation |
| M83209 | parotid secretory protein | 1 | L-serine biosynthesis |
| D10233 | renin binding protein | 1 | Metabolism |
| L12407 | dopamine beta hydroxylase | 1 | Metabolism |
| L26043 | perilipin | 1 | Metabolism |
| U24071 | unc-13 homolog B (C. elegans) | 1 | Neurotransmitter secretion |
| D64061 | eukaryotic translation initiation factor 5B | 1 | Protein Biosynthesis |
| AA892680 | peptidylprolyl isomerase (cyclophilin)-like 3 | 1 | Protein Biosynthesis |
| D50696 | peptidase (prosome, macropain) 26S subunit, ATPase 1 | 1 | Protein Catabolism |
| U50194 | tripeptidyl peptidase II | 1 | Protein Catabolism |
| D10699 | ubiquitin carboxy-terminal hydrolase L1 | 1 | Protein Catabolism |
| D45247 | proteasome (prosome, macropain) subunit, beta type 5 | 1 | Protein Catabolism |
| D21799 | proteasome (prosome, macropain) subunit, beta type 2 | 1 | Protein Catabolism |
| D90265 | proteasome (prosome, macropain) subunit, alpha type 1 | 1 | Protein Catabolism |
| AF054270 | prolactin induced protein | 1 | Protein Processing |
| X93352 | ribosomal protein L10A | 1 | Protein Synthesis |
| AA891713 | ribosomal protein L13A | 1 | Protein Synthesis |
| M17419 | ribosomal protein L5 | 1 | Protein Synthesis |
| AI170685 | DnaJ (Hsp40) homolog, subfamily A, member 2 | 1 | Protein Synthesis |
| AI103238 | regulatory subunit B (PR 52), beta isoform | 1 | Reproduction |
| AB011068 | deiodinase, iodothyronine, type II | 1 | Response to Cold |
| X68400 | protein kinase C, eta | 1 | Signaling |
| D45412 | protein tyrosine phosphatase, receptor type, O | 1 | Signaling |
| M17526 | guanine nucleotide binding protein, alpha o | 1 | Signaling |
| AF064706 | G protein-coupled receptor 6 | 1 | Signaling |
| U66274 | neuropeptide Y receptor Y5 | 1 | Signaling |

| | | | |
|---|---|---|---|
| D15069 | adrenomedullin | 1 | Signaling |
| Z35654 | mcf.2 transforming sequence-like | 1 | Signaling |
| AA859752 | noggin | 1 | Skeletal Development |
| AF053101 | paired box gene 4 | 1 | Transcription |
| AI639353 | pleiotropic regulator 1  homolog (Arabidopsis) | 1 | Transcription |
| AF062594 | nucleosome assembly protein 1-like 1 | 1 | Transcription |
| AA866472 | Nucleosome assembly protein 1-like 1 | 1 | Transcription |
| M99221 | calcium channel, voltage-dependent, L type, alpha 1D subunit | 1 | Transport |
| X78997 | cadherin 17 | 1 | Transport |
| AF019043 | dynamin 1-like | 1 | Transport |
| X81448 | Similar to cytokeratin | 1 | No Identified Ontologies |
| AI104388 | aurora-A kinase interacting protein | 1 | No Identified Ontologies |
| AA891829 | WD40 protein Ciao1 (predicted) | 1 | No Identified Ontologies |
| X99338 | stromal cell derived factor receptor 1 | 1 | No Identified Ontologies |
| X77815 | variable coding sequence A2 | 1 | No Identified Ontologies |
| AI072634 | --- | 1 | No Identified Ontologies |
| AA859804 | --- | 1 | No Identified Ontologies |
| AA892310 | --- | 1 | No Identified Ontologies |
| AA800017 | Similar to hypothetical protein BC011833 (predicted) | 1 | No Identified Ontologies |
| AA893307 | Nuclear cap binding protein subunit 2 (predicted) | 1 | No Identified Ontologies |
| H31648 | Transcribed locus | 1 | No Identified Ontologies |
| AA874849 | --- | 1 | No Identified Ontologies |
| AA892369 | --- | 1 | No Identified Ontologies |
| AA894054 | CDNA clone IMAGE:7326015 | 1 | No Identified Ontologies |
| AA956941 | --- | 1 | No Identified Ontologies |
| AA892818 | Transcribed locus | 1 | No Identified Ontologies |
| M13949 | --- | 1 | No Identified Ontologies |
| AA800275 | Transcribed locus | 1 | No Identified Ontologies |
| AA799865 | Transcribed locus | 1 | No Identified Ontologies |
| AI639039 | --- | 1 | No Identified Ontologies |
| AA875554 | Transcribed locus | 1 | No Identified Ontologies |
| H33467 | Transcribed locus | 1 | No Identified Ontologies |
| H31753 | Transcribed locus | 1 | No Identified Ontologies |
| AI230789 | --- | 1 | No Identified Ontologies |
| AI639464 | --- | 1 | No Identified Ontologies |
| AI639289 | --- | 1 | No Identified Ontologies |
| AI639459 | --- | 1 | No Identified Ontologies |
| AI639033 | --- | 1 | No Identified Ontologies |
| AA800948 | similar to Tubulin alpha-4 chain (Alpha-tubulin 4) | 1 | No Identified Ontologies |
| L03386 | Olf-1/EBF associated Zn finger protein Roaz | 1 | No Identified Ontologies |
| U66707 | densin-180 | 1 | No Identified Ontologies |
| AF053987 | putative pheromone receptor V2R1 | 1 | No Identified Ontologies |
| AF079873 | zinc finger protein 162 | 2 | Apoptosis |
| AI070295 | --- | 2 | Apoptosis |
| AA945608 | serum amyloid P-component | 2 | Cell Adhesion |
| AA859869 | proteasome (prosome, macropain) 26S subunit, non-ATPase, 1 | 2 | Cell Cycle Regulation |
| X67805 | synaptonemal complex protein 1 | 2 | Cytokinesis |

| | | |
|---|---|---|
| AF017757 | axin2 | 2 Development |
| AA893280 | adipose differentiation-related protein | 2 Fatty Acid Transport |
| AA858520 | follistatin | 2 Gametogenesis |
| AB014722 | squamous cell carcinoma antigen recognized by T-cells 1 | 2 Immune Response |
| J05446 | glycogen synthase 2 | 2 Metabolism |
| AA891286 | thioredoxin reductase 1 | 2 Metabolism |
| M27440 | apolipoprotein B | 2 Metabolism |
| S78217 | protein phosphatase 1, catalytic subunit, gamma isoform | 2 Metabolism |
| S78218 | protein phosphatase 1, catalytic subunit, beta isoform | 2 Metabolism |
| J02810 | glutathione S-transferase, mu 1 | 2 Metabolism |
| U91847 | mitogen activated protein kinase 14 | 2 Metabolism |
| AA859920 | Nucleosome assembly protein 1-like 1 | 2 Nucleosome Assembly |
| U95052 | Eukaryotic translation initiation factor 4 gamma, 2 | 2 Protein Biosynthesis |
| AA875205 | eukaryotic translation initiation factor 3, subunit 9 (eta) (predicted) | 2 Protein Biosynthesis |
| U13176 | ubiquitin-conjugating enzyme E2D 2 | 2 Protein Catabolism |
| L38615 | glutathione synthetase | 2 Protein Catabolism |
| U48592 | interleukin 1 receptor accessory protein | 2 Signaling |
| M23591 | protein phosphatase 2a, catalytic subunit, beta isoform | 2 Signaling |
| X16044 | protein phosphatase 2a, catalytic subunit, beta isoform | 2 Signaling |
| X89704 | olfactory receptor 1283 | 2 Signaling |
| AF055291 | signal transducer and activator of transcription 4 | 2 Signaling |
| AF037199 | RE1-silencing transcription factor | 2 Transcription |
| AB017044 | forkhead box A3 | 2 Transcription |
| AB017044 | forkhead box A3 | 2 Transcription |
| AI177751 | transcription elongation factor B (SIII), polypeptide 1 | 2 Transcription |
| AI104524 | heterogeneous nuclear ribonucleoprotein A/B | 2 Transcription |
| X66022 | neuronal d4 domain family member | 2 Transcription |
| AJ001641 | POU domain, class 3, transcription factor 3 | 2 Transcription |
| M18416 | early growth response 1 | 2 Transcription |
| D17711 | heterogeneous nuclear ribonucleoprotein K | 2 Transcription |
| S59893 | Sjogren syndrome antigen B | 2 Transcription |
| AF044910 | survival of motor neuron 1, telomeric | 2 Transcription |
| X62145 | ribosomal protein L8 | 2 Transcription |
| AF083341 | potassium large conductance calcium-activated channel | 2 Transport |
| J05510 | inositol 1,4,5-triphosphate receptor 1 | 2 Transport |
| AA799276 | ATPase, Ca++ transporting, cardiac muscle, slow twitch 2 | 2 Transport |
| J03969 | nucleophosmin 1 | 2 Transport |
| U12402 | ADP-ribosylation factor-like 1 | 2 Transport |
| AA875099 | nuclear pore associated protein | 2 Transport |
| H31747 | --- | 2 Transport |
| H32977 | actin related protein 2/3 complex, subunit 5-like (predicted) | 2 No Identified Ontologies |
| D88035 | beta-1,3-glucuronyltransferase 1 (glucuronosyltransferase P) | 2 No Identified Ontologies |
| AF016702 | Glycoprotein hormones, alpha subunit | 2 No Identified Ontologies |

| AF090692 | cystatin 8 | 2 | No Identified Ontologies |
|---|---|---|---|
| AFFX-TrpnX-M | --- | 2 | No Identified Ontologies |
| AI178828 | --- | 2 | No Identified Ontologies |
| AB013454 | Similar to Ac2-210 | 2 | No Identified Ontologies |
| AA893603 | Solute carrier family 35, member E1 (predicted) | 2 | No Identified Ontologies |
| AA891842 | Tumor necrosis factor receptor superfamily, member 21 (predicted) | 2 | No Identified Ontologies |
| AI639257 | --- | 2 | No Identified Ontologies |
| AI639476 | --- | 2 | No Identified Ontologies |
| AI639474 | --- | 2 | No Identified Ontologies |
| AA892010 | Similar to Hypothetical protein CGI-128 homolog (predicted) | 2 | No Identified Ontologies |
| AA893422 | Transcribed locus | 2 | No Identified Ontologies |
| AI071399 | --- | 2 | No Identified Ontologies |
| AI639486 | --- | 2 | No Identified Ontologies |
| AA893180 | Transcribed locus | 2 | No Identified Ontologies |
| AA900850 | --- | 2 | No Identified Ontologies |
| AI639120 | similar to RIKEN cDNA 1700088E04; | 2 | No Identified Ontologies |
| S68589 | protein kinase N3 (predicted) | 2 | No Identified Ontologies |
| AA891838 | similar to ribosomal protein P0-like protein | 2 | No Identified Ontologies |
| D10854 | aldo-keto reductase family 1, member A1 | 3 | Aldehyde Catabolism |
| Z75029 | Heat shock 70kD protein 1A | 3 | Apoptosis |
| X53428 | glycogen synthase kinase 3 beta | 3 | Cell Cycle |
| X59859 | decorin | 3 | ECM Organization |
| AF054618 | cortactin isoform B | 3 | Endocytosis |
| S45392 | heat shock 90kDa protein 1, beta | 3 | Heat Shock |
| M81225 | farnesyltransferase, CAAX box, alpha | 3 | Metabolism |
| AA799466 | adenylate kinase 2 | 3 | Metabolism |
| D86215 | NADH dehydrogenase (ubiquinone) 1 alpha subcomplex  5 | 3 | Metabolism |
| J02752 | acyl-Coenzyme A oxidase 1, palmitoyl | 3 | Metabolism |
| J05470 | carnitine palmitoyltransferase 2 | 3 | Metabolism |
| U17901 | phospholipase A2, activating protein | 3 | Metabolism |
| AA946040 | cytochrome c oxidase, subunit VIb (predicted) | 3 | Metabolism |
| AF059530 | heterogeneous nuclear ribonucleoprotein methyltransferase-like 3 | 3 | Methylation |
| M37568 | homeo box C8 | 3 | Neuron Differentiation |
| AA875069 | H3 histone, family 3B | 3 | Nucleosome Assembly |
| X62166 | ribosomal protein L3 | 3 | Protein Biosynthesis |
| AA875327 | eukaryotic translation initiation factor 4H | 3 | Protein Biosynthesis |
| AA859719 | --- | 3 | Protein Biosynthesis |
| X62146 | ribosomal protein L11 (predicted) | 3 | Protein Biosynthesis |
| X62146 | ribosomal protein L11 (predicted) | 3 | Protein Biosynthesis |
| AA892367 | ribosomal protein L3 | 3 | Protein Biosythesis |
| D10755 | proteasome (prosome, macropain) subunit, alpha type 6 | 3 | Protein Catabolism |
| X51536 | ribosomal protein S3 | 3 | Protein Synthesis |
| AA800054 | ribosomal protein L19 | 3 | Protein Synthesis |
| X51536 | ribosomal protein S3 | 3 | Protein Synthesis |
| AI178750 | eukaryotic translation elongation factor 2 | 3 | Protein Synthesis |

| X58465 | ribosomal protein S5 | 3 | Protein Synthesis |
|--------|---------------------|---|-------------------|
| X53378 | ribosomal protein L13 | 3 | Protein Synthesis |
| X78327 | ribosomal protein L13 | 3 | Protein Synthesis |
| M89646 | Ribosomal protein S24 | 3 | Protein Synthesis |
| AI176546 | heat shock protein 1, alpha | 3 | Protein Synthesis |
| AA944397 | heat shock protein 1, alpha | 3 | Protein Synthesis |
| AA800211 | pyridoxine 5'-phosphate oxidase | 3 | Pyridoxine Biosynthesis |
| AF036335 | NonO/p54nrb homolog | 3 | RNA splicing |
| AF036335 | NonO/p54nrb homolog | 3 | RNA splicing |
| U87960 | protein tyrosine phosphatase, receptor type, F | 3 | Signaling |
| U10303 | endothelial differentiation sphingolipid G-protein-coupled receptor 1 | 3 | Signaling |
| U14409 | melatonin receptor 1A | 3 | Signaling |
| U50949 | olfactory receptor 1641 | 3 | Signaling |
| AB007688 | homer homolog 1 (Drosophila) | 3 | Signaling |
| M36317 | thyrotropin releasing hormone | 3 | Signaling |
| X01454 | thyroid stimulating hormone, beta subunit | 3 | Signaling |
| S63167 | hydroxysteroid dehydrogenase-1, delta<5>-3-beta (predicted) | 3 | Steroid Hormone Biosynthesis |
| AI009098 | hypoxia up-regulated 1 | 3 | Stress |
| M96630 | SEC61, alpha subunit (S. cerevisiae) | 3 | Transport |
| M30312 | potassium voltage-gated channel, shaker-related subfamily, member 3 | 3 | Transport |
| AI230914 | farnesyltransferase, CAAX box, beta | 3 | Wound healing |
| U82591 | chromosome 6 open reading frame 108 | 3 | No Identified Ontologies |
| AF051155 | G protein beta subunit-like | 3 | No Identified Ontologies |
| X52815 | --- | 3 | No Identified Ontologies |
| AA891742 | Similar to cDNA sequence BC019806 (predicted) | 3 | No Identified Ontologies |
| AA892863 | --- | 3 | No Identified Ontologies |
| AI104513 | --- | 3 | No Identified Ontologies |
| S76758 | --- | 3 | No Identified Ontologies |
| AI012942 | --- | 3 | No Identified Ontologies |
| AF027188 | Similar to RIKEN cDNA 4933424N09 (predicted) | 3 | No Identified Ontologies |
| AI639409 | --- | 3 | No Identified Ontologies |
| AI233591 | --- | 3 | No Identified Ontologies |
| AA859835 | Transcribed locus | 3 | No Identified Ontologies |
| AF053097 | --- | 3 | No Identified Ontologies |
| AF034753 | Similar to hypothetical protein FLJ22490 (predicted) | 3 | No Identified Ontologies |
| U47311 | --- | 3 | No Identified Ontologies |
| H33253 | Similar to tubulin-specific chaperone d | 3 | No Identified Ontologies |
| AI113046 | --- | 3 | No Identified Ontologies |
| AA900850 | --- | 3 | No Identified Ontologies |
| AA875265 | --- | 3 | No Identified Ontologies |
| D17349 | similar to cytochrome P450 2B15 | 3 | No Identified Ontologies |
| AI007820 | heat shock 90kDa protein 1, beta | 3 | No Identified Ontologies |
| H31907 | embryo-related protein | 3 | No Identified Ontologies |
| H33725 | associated molecule with the SH3 domain of STAM | 4 | Apoptosis |
| AA800206 | actinin alpha 2 (predicted) | 4 | Apoptosis |
| AB010436 | cadherin 8 | 4 | Cell Adhesion |

| | | |
|---|---|---|
| X95990 | complement component 5, receptor 1 | 4 Chemotaxis |
| M33605 | perforin 1 (pore forming protein) | 4 Cytolysis |
| D90219 | natriuretic peptide precursor type C | 4 DNA Metabolism |
| X84210 | nuclear factor I/A | 4 DNA Replication |
| AA892798 | uterine sensitization-associated gene 1 protein | 4 Embryo Implantation |
| M22360 | Murinoglobulin 1 homolog (mouse) | 4 Inflammation |
| M22993 | Murinoglobulin 1 homolog (mouse) | 4 Inflammation |
| M94548 | cytochrome P450, family 4, subfamily F, polypeptide 2 | 4 Metabolism |
| AF008554 | implantation-associated protein | 4 Metabolism |
| D00688 | monoamine oxidase A | 4 Metabolism |
| D00729 | Dodecenoyl-coenzyme A delta isomerase | 4 Metabolism |
| AF036761 | stearoyl-Coenzyme A desaturase 2 | 4 Metabolism |
| M83196 | microtubule-associated protein 1 A | 4 Microtubule depolymerization |
| M91652 | glutamine synthetase 1 | 4 Nitrogen Fixation |
| AF061726 | calpain 3 | 4 Protein Catabolism |
| L34262 | palmitoyl-protein thioesterase | 4 Protein Production |
| X13905 | similar to Ras-related protein Rab-1B | 4 Protein Transport |
| AF073891 | potassium voltage-gated channel, subfamily H (eag-related), member 5 | 4 Signaling |
| U57500 | protein tyrosine phosphatase, receptor type, A | 4 Signaling |
| L19112 | fibroblast growth factor receptor 2 | 4 Signaling |
| L35921 | guanine nucleotide binding protein (G protein), gamma 8 subunit | 4 Signaling |
| D10763 | erythropoietin | 4 Stress |
| AB015432 | tumor-associated protein 1 | 4 Tansport |
| M24393 | myogenin | 4 Transciprtion |
| L13206 | forkhead box D4 | 4 Transcirption |
| U01146 | nuclear receptor subfamily 4, group A, member 2 | 4 Transcription |
| AI145177 | early growth response 4 | 4 Transcription |
| AF059273 | glucocorticoid modulatory element binding protein 2 | 4 Transcription |
| AJ006519 | amiloride-sensitive cation channel 2, neuronal | 4 Transport |
| AF104399 | Cbp/p300-interacting transactivator | 4 Transport |
| J02844 | carnitine O-octanoyltransferase | 4 Transport |
| D12573 | hippocalcin | 4 Transport |
| U02096 | fatty acid binding protein 7, brain | 4 Transport |
| U09211 | solute carrier family 18 (vesicular acetylcholine), member 3 | 4 Transport |
| AI102031 | bridging integrator 1 | 4 Transport |
| X63744 | solute carrier family 1 | 4 Transport |
| AF090692 | cystatin 8 | 4 No Identified Ontologies |
| D26492 | dynein, axonemal, heavy polypeptide 1 | 4 No Identified Ontologies |
| AI639159 | solute carrier family 23 | 4 No Identified Ontologies |
| X01115 | seminal vesicle secretion 5 | 4 No Identified Ontologies |
| AA875001 | tripartite motif protein 8 (predicted) | 4 No Identified Ontologies |
| U77626 | WW domain binding protein 4 (formin binding protein 21) | 4 No Identified Ontologies |
| AI639076 | --- | 4 No Identified Ontologies |
| AA892394 | CUG triplet repeat, RNA binding protein 1 (predicted) | 4 No Identified Ontologies |
| AI639179 | --- | 4 No Identified Ontologies |

| AI231445 | --- | 4No Identified Ontologies |
| AA866264 | --- | 4No Identified Ontologies |
| AI638972 | --- | 4No Identified Ontologies |
| AA892228 | CUG triplet repeat, RNA binding protein 1 (predicted) | 4No Identified Ontologies |
| AA866293 | Transcribed locus | 4No Identified Ontologies |
| AA875089 | --- | 4No Identified Ontologies |
| AI639500 | --- | 4No Identified Ontologies |
| AI009658 | --- | 4No Identified Ontologies |
| H31550 | Transcribed locus | 4No Identified Ontologies |
| AI137538 | --- | 4No Identified Ontologies |
| AA859992 | --- | 4No Identified Ontologies |
| AI639012 | similar to cDNA sequence BC019776 | 4No Identified Ontologies |
| AA800803 | similar to RIKEN cDNA 6720485C15 | 4No Identified Ontologies |
| AI169372 | RAS-like family 11 member A | 4No Identified Ontologies |
| S65091 | similar to protein phosphatase 1, regulatory (inhibitory) subunit 1C; | 4No Identified Ontologies |
| AA858621 | CaM-kinase II inhibitor alpha | 4No Identified Ontologies |
| AI171848 | apical early endosomal glycoprotein | 4No Identified Ontologies |