Prediction of Gene Expression Levels and the Role of Cis-Acting Elements in Age-Related Cataract by Applying a Promoter-Based Modeling Approach

Jung-Min Lim[†] and Kwang-Hyun Cho*[,‡]

School of Electrical Engineering and Computer Science, Seoul National University, Gwanak P.O. Box 34, Seoul 151-600, Korea, and College of Medicine, Seoul National University, Chongno-Gu, Seoul, 110-799, and Korea Bio-MAX Institute, Seoul National University, Gwanak-Gu, Seoul, 151-818, Korea

Cataract is a dynamical process of lens opacity formation involving many inter- and intracellular regulations, as well as metabolic genes and transcription factors. Using a series of microarray-derived mRNA profiles for human cataractogenesis (Hawse et al. *Mol. Vision* **2003**, *9*, 515−537), we develop a promoter-based system-theoretic modeling to demonstrate model-driven prediction of gene expression levels and to identify the role of critical cis-acting elements. In this study, 14 key mRNA expression data from the structural and pathological molecules of age-related cataract samples are used. The first seven genes consist of structural molecules, and the second half of genes are composed of heat shock proteins, filensin, and glutathione peroxidase 3. The presented result demonstrates that mRNA expression levels of structural proteins such as crystallins can be successfully predicted from 5′ flanking regulatory DNA sequences. In addition, predicted gene expression levels of heat shock protein, β-tubulin, and αA-crystallin accurately estimate the stimulatory or inhibitory role of distributed cis-acting elements, i.e., c-Myc, GATA-1, GR, NE-E, and Pit-1. Although it is difficult to predict the overall gene expression levels in cataract samples, the present study shows the potential use of promoter-based modeling and prediction of the gene expression levels for age-related cataract.

Introduction

Cataract is the most common cause of blindness worldwide. It is in general associated with the breakdown of a lens architecture (*1*). Further, it is well-known as a complex pathological process featuring the abnormal modification of crystallins and extracellular matrices leading to opacity of the lens (*2*, *3*). Although age-related cataract is widely prevalent, most advances have been made in the genetic studies of cataract formation (*4*).

Recently, the development of microarray technology enables us to detect genes that show significant changes in their gene expressions when normal and cataractous classes of samples are being compared. Although the spectrum of genes has been revealed, the detailed relationship between genes and their role for cataract is nevertheless largely unknown (*5*). However, the identification of genes differentially regulated for age-related cataract formation provides an important insight into molecular mechanisms of cataract. For instance, the function of crystallins as a chaperone has been recently reported, whereas the specific function of α-crystallin and β2-crystallin is not yet completely known (*3*, *6*, *7*).

Despite the enormous volumes of experimental microarray mRNA data, their practical utilization is still limited by the repeated confirmation through the conventional cell and molecular biology approaches. However, recent studies have shown that high throughput empirical microarray mRNA data can lead to useful mathematical modeling, and the predicted model can further generate testable hypotheses of biological systems (*8*−*10*). These studies have also shown the implementation of system-theoretic modeling of gene expression levels in conjuction with the microarray data featuring an intracellular and interconnected gene network responsible for the states of complex biological systems. It is now widely accepted that a theoretical modeling approach can provide a useful insight into the massive empirical data by formulating a verifiable hypothesis for biological systems (*11*, *12*).

In this study, a promoter-based system-theoretic approach is employed to predict the gene expression levels in age-related cataract samples. Using the microarray-based mRNA expression profiles reported by Hawse et al. (*13*), we focus on 14 documented genes of lens architecture and pathology, including major structural proteins αA-crystallin and β2-crystallin. The proposed approach can also predict the stimulatory or inhibitory role by the combination of cis-acting elements that have been assumed to most dominantly regulate the mRNA expression levels (*8*, *14*).

Materials and Methods

Definition of mRNA Expression Ratios. The microarray-based mRNA expression profiles of the selected genes, involved in the human age-related cataract, are obtained from the data published by Hawse et al. (*13*). The logarithmic fold changes in mRNA levels are deter-

* To whom correspondence should be addressed. Tel: +82-2-887-2650. Fax: +82-2-887-2692. E-mail: ckh-sb@snu.ac.kr.
[†] School of Electrical Engineering and Computer Science.
[‡] College of Medicine and Korea Bio-MAX Institute.

10.1021/bp050027s CCC: $30.25
Published on Web 06/29/2005

mined for 14 genes, whose 5′-end flanking DNA sequence is identifiable from the GenBank (NIH, Maryland). The positive or negative logarithmic ratio indicates up-regulation or down-regulation compared to the normal lens, respectively.

Samples. The mRNA expression levels of seven structural molecules including crystallin-βB1, -βB2, -βA4, -αA, -βB, β-tubulin, keratin 19, and amyloid-β peptide (APP) are obtained from the study by Hawse et al. The 5′-end upstream regulatory region is used for the modeling. The accession numbers were U09951 (crystallin βB1), U22455 (crystallin βB2), U18260 (crystallin βA4), J00375 (crystallin αA), Z22573 (crystallin γB), AF417567 (β-tubulin), AF089865 (keratin 19), and AF067971 (amyloid-β peptide). Cis-acting elements such as GR, NF-E2, Pit-1, GATA-1, NF-E, Pit-1, c-Myc, et al. are considered.

We also focus on 7 genes such as heat shock proteins, filensin, glutathione peroxidase 3, and serine-threonine kinase. The accession numbers are AB010375 (filensin), X83230 (heat shock protein 90β), AF139178 (heat shock protein 70.2), X04009 (heat shock protein 27), AY552097 (glutathione peroxidase 3), and U01337 (serine-threonine kinase).

Formulation of a Promoter-Based Linear Model. The logarithmic mRNA ratios for the 14 selected genes are estimated using a promoter matrix **H** (n,m) where n is the number of genes, and m is the number of putative cis-acting elements selected from the software SIGSCAN (Advance Biosciences Computing Center, University of Minnesota, MN). The promoter-based linear model has two inputs of "mRNA expression data" and "information on cis-acting DNA regulatory elements". The linear model is designed such that it minimizes the mismatch between the experimental mRNA levels and the predicted mRNA levels. The present mathematical formulation allow us to highlight the merits and limitations of linear approximation in analyzing complex eukaryotic transcriptional regulations.

With the understanding that DNA cis-acting regulatory elements can be regulated with precision, we introduced a linear least-squares model. Based on the 5′-flanking regulatory elements in the promoter, we counted the number of cis-acting elements for each gene. Then, we experimentally modeled the observed mRNA levels using a weighted sum of the frequency of the selected cis-acting elements. Employing the estimation theory in systems science (*14*), the ratio of mRNA expression levels and a state vector can be modeled by

$$z = [h_{ij}]_{n\times m}\begin{bmatrix} x_1^{\,2} \\ x_2^{\,2} \\ \dots \\ x_m^{\,2} \end{bmatrix} = [h_{ij}]_{n\times m}\begin{bmatrix} x_1 0....0 \\ 0x_2 0..0 \\ ..... \\ 0....0x_m \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{bmatrix} = H(n,m)H_A(m,m)x \quad (1)$$

where z denotes the logarithmic ratio of mRNA expression levels, x indicates the state vector representing contribution of the individual cis-acting element to the measured mRNA levels, **H** (n,m) is a promoter matrix with an element h_{ij} representing the number of frequencies of the jth cis-acting element appeared in the 5′-end flanking DNA sequence of the ith gene, and $\mathbf{H_A}$ (m,m) is a $m\times m$ promoter-associated diagonal matrix whose jth diagonal component weighs a contribution of the jth cis-acting element to the measured transcript levels. A positive or negative value in the element of x indicates the a stimulatory or inhibitory effect on z, respectively.

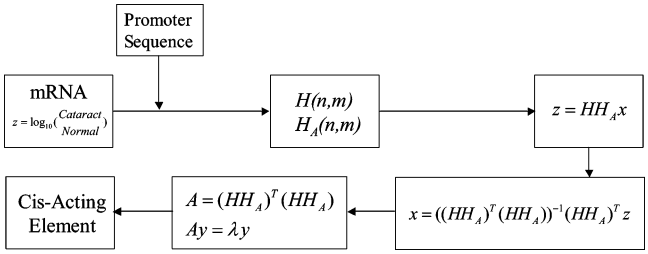

Figure 1. Schematic diagram of the proposed system-theoretic gene expression prediction. System-theoretic prediction of gene expression levels using logarithmic ratios of known mRNA expression levels in human age-related cataract samples.

Table 1. Selected genes for promoter-based modeling

genes	origin	sizes
heat shock protein 90β	*Gallus*	3069
heat shock protein 70.2	*Sus scrofa*	460
heat shock protein 27	*Drosophila melanogaster*	581
β-tubulin	*Crypthecodinium cohnii*	517
keratin 19	*Rattus norvegicus*	1407
amyloid-β peptide	*Macaca mulatta*	5850
filensin	*Mus musculus*	2248
glutathione peroxidase 3	*Homo sapiens*	1340
serine-threonine kinase	human	1380
crystallin βB1	*Gallus*	2350
crystallin βB2	*Mus musculus*	600
crystallin βA4	*Gallus*	2160
crystallin αA	mouse	440
crystallin γB	*Mus musculus*	650

Estimation of the Temporal State of Cis-Acting Elements. The state of x was estimated using an inverse matrix:

$$x = (H_A^T H^T H H_A)^{-1} H_A^T H^T z \quad (2)$$

To evaluate the effectiveness of the selected cis-acting elements on the mRNA expression, the eigenvalue and the eigenvector of the matrix $A = H_A^T H^T H H_A$ were analyzed. For a linear equation, there exists a scalar λ and a vector y such that $Ay = \lambda y$ holds where λ is called an eigenvalue and y is called an eigenvector corresponding to λ.

Regulatory Network with the Predicted Cis-Acting Elements. The stimulatory or inhibitory role of the model-predicted cis-acting elements is illustrated in a regulatory network for the 14 selected genes as a genomic cis-regulatory logic. A linkage map between the known cis-acting elements and the predicted cis-acting elements are drawn based on a sequence similarity. Matrix operations such as multiplication, transposition, and inversion as well as the computation of eigenvectors and eigenvalues were conducted using MATLAB (version 6, The Math Works Inc.)

Results

Selection of Genes. To examine the promoter-based modeling of gene expression levels, 14 mRNA microarray data of logarithmic fold changes in age-related cataract samples are used. Those are mainly composed of crystallins, a major structural molecule of the lens, intracellular signaling molecules, matrix proteins, and others. The selected genes are listed in Table 1.

Determination of the Promoter Matrix H. A schematic procedure of promoter-based modeling is illustrated in Figure 1. The candidates for potential cis-acting elements were identified using the software SIGSCAN (University of Minnesota, MN), as described in Materials and Methods. To determine the promoter-based

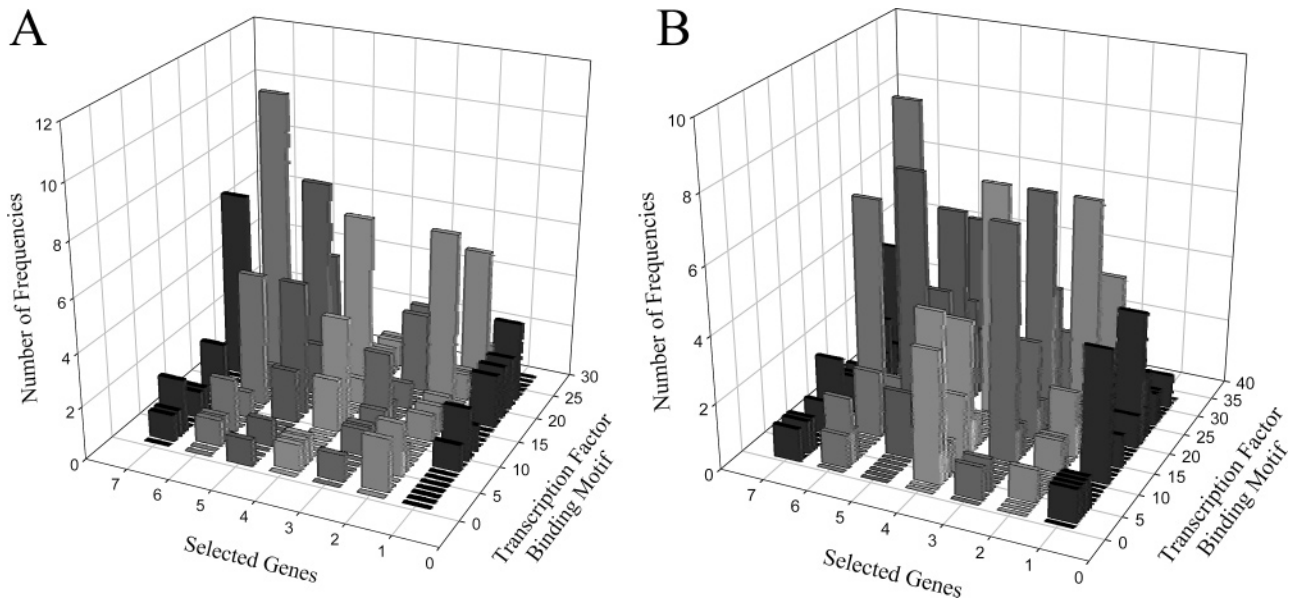

Figure 2. Classification of genes according to the transcription factor binding motifs of the selected genes for 250-bp promoters. Using the putative cis-acting elements on 250-bp upstream DNA sequences, the data sets are divided into the transcript levels of structural molecules and intracellular proteins.

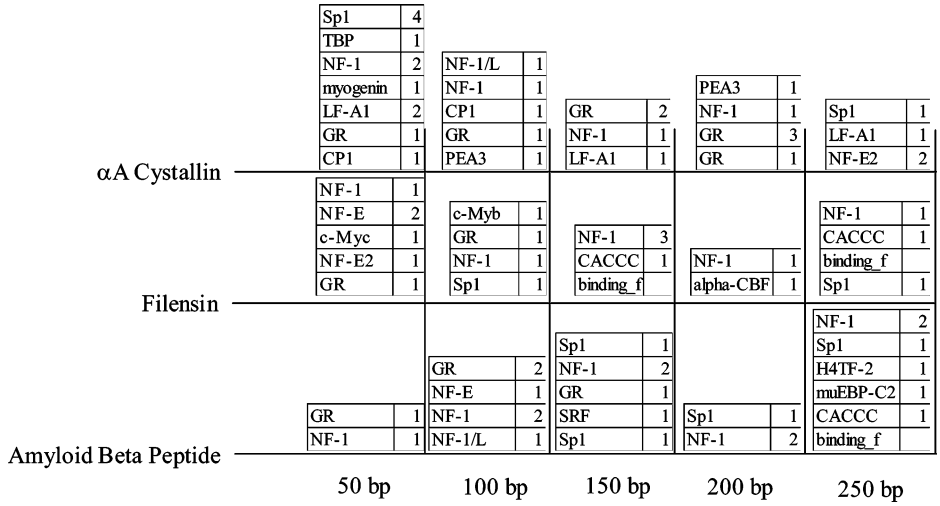

Figure 3. Map of cis-acting elements for αA-crystallin, filensin, and amyloid-β peptide. Distribution of cis-acting motifs on 250-bp upstream sequences where GR, NF-1, c-Myc, Sp1, LF-A1, etc. were represented, respectively.

matrix **H**, we mainly focused on 5′-end flanking DNA sequences of selected genes. The selected cis-acting elements from 5-bp DNA sequences were 5′-TGACTCA-3′ (AP1), 5′-CCACC-3′ (CAC-binding_pro), 5′-CAGAG-3′ (GR), 5′-GCCA-3′ (NF-1), 5′-TGCAC-3′ (SP1), and 5′-GATTTC-3′ (H4TF-1). The number of putative cis-acting elements in the regulatory DNA region of the selected genes is shown in Figure 2. In addition, the selected promoter matrix **H** featuring Figure 4A is shown in Table 2.

Modeling of mRNA Gene Expression Levels. Using the putative cis-acting elements on 5′ flanking DNA sequences (Figure 3), the expression level of each gene pair (i.e., reference vs unknown) is derived from the measured expression level in logarithmic mRNA expression ratios. In both measured and predicted gene expression levels, the positive or negative value suggests up-regulation or down-regulation with respect to the expression levels of normal lenses, respectively (Figure 4).

In promoter-based modeling, two major modeling factors implemented are the size of regulatory DNA regions and the number of genes. We predict the mRNA levels of selected genes from the mRNA levels of counterpart genes. The predicted patterns of the mRNA expression levels are illustrated for the selected genes (Figure 4). Nine out of 10 cases are properly predicted, whereas the case of crystalline beta B2 is slightly off. Interestingly, the size of a promoter does not necessarily guarantee successful prediction of the gene expression level (Figure 5A). However, the number of genes used for promoter matrix **H** is well fitted by successful prediction of gene expression levels (Figure 5B).

State Estimation of the Predicted Cis-Acting Elements. The temporal alteration of the state variable x, which represents the contribution of the predicted cis-acting elements, is plotted in Figure 6. The level of zero indicates no contribution, whereas the positive or negative values suggest the stimulatory or inhibitory roles, respectively. The temporal levels change significantly from the normal to the cataractous lenses. For crystallin αA, six out of 18 cis-acting elements (GR, LF-A1, NF-1, NF-E, NF-E2, and SP1) are predicted to play the stimulatory role at all time points, while three cis-acting elements (CP1, myogenin, TBP) are predicted to have no

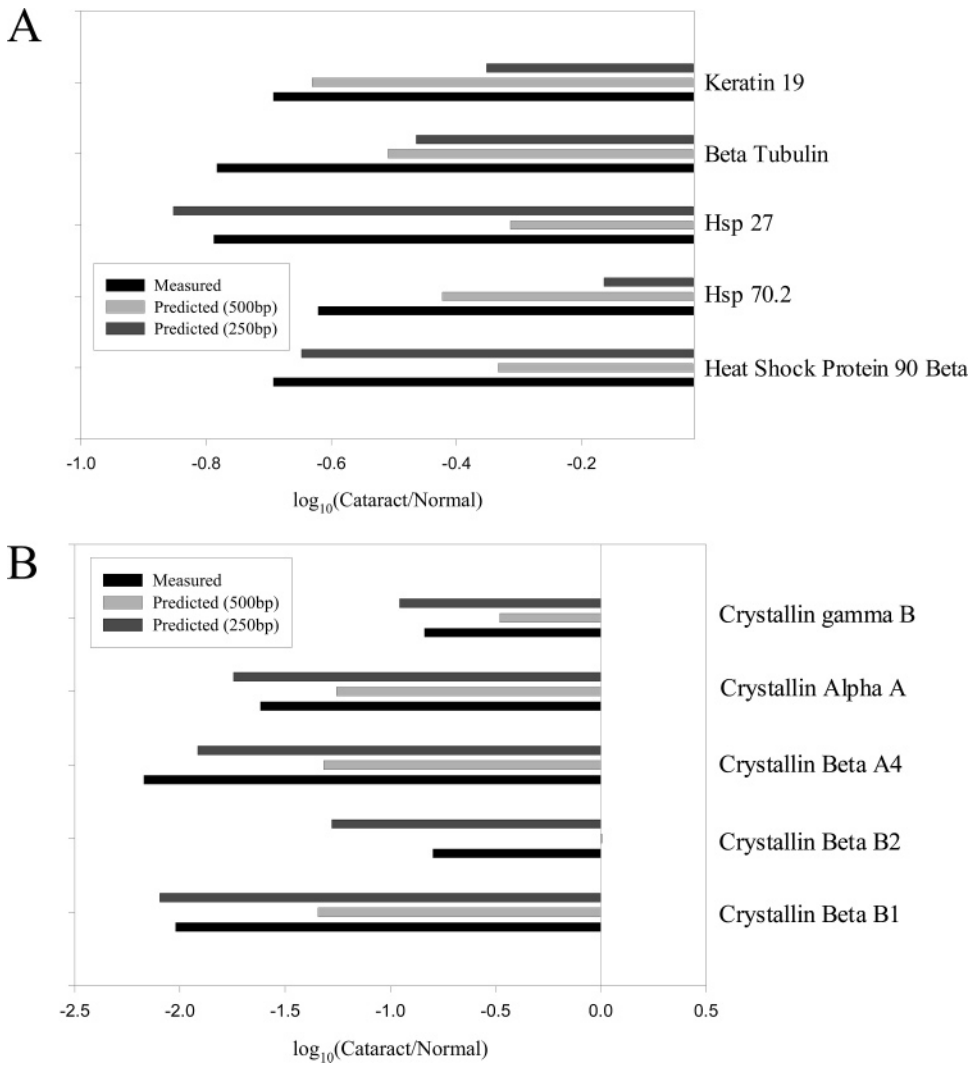

Figure 4. Comparison of the measured mRNA and the predicted mRNA levels. Ten genes are grouped into two categories. The measured and the predicted expression levels of each gene are shown by the gray colored lines for predicted expression whereas the bottom black indicates the measured expression. In addition, the positive or negative values illustrate the alteration of mRNA expression levels.

Table 2. Promoter matrix H for Figure 4A

	cis-acting element													
gene	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
Hsp90β	0	0	0	0	2	1	0	2	2	1	0	3	0	0
Hsp70.2	2	0	0	0	0	1	1	7	0	1	1	0	5	1
Hsp27	1	0	0	1	0	3	0	1	0	3	0	3	2	0
β-tubulin	1	1	1	0	2	4	1	6	0	0	0	1	1	0
keratin 19	0	0	0	0	2	5	2	8	5	0	0	1	0	0

contribution at all. Similarly, for Hsp 27, 10 out of 17 cis-acting elements (α-CBF, CAC-binding_pro, c-Myb, c-Myc, F2F, GR, NF-1, NF-1/L, NF-E2, Pit-1, and Sp1) are predicted to play a negative role at all time points, while two cis-acting elements (NF-2, TGF3) were modeled to have a stimulatory role in the gene expression level.

Discussion

The purpose of microarray technology is to detect genes that show significant changes in their expression levels when two classes of samples are being compared. In this paper, we have presented a 5′-flanking DNA sequence based model with parameters for the promoter matrix, the overall level of gene expression, and the change of expression levels across the human cataractous samples. We have presented and applied our modeling in the context of gene expressions measured by microarray technology (*13*, *15*, *16*). We have successfully dealt with the vast amount of data generated logistically to predict the gene expression profiling associated with the particular ocular disease cataract and to show the potential benefits in ophthalmic research (*16*).

A pool of 5′-flanking DNA sequences of structural and molecular genes have been used as modeling library to predict mRNA expression levels of age-related cataract and to analyze the role of cis-acting elements corresponding to its regulated genes. The presented results have demonstrated the feasibility of predicting the mRNA expression levels of structural genes from molecular genes or vice versa and identifying the role of putative cis-acting elements from the estimated eigenvalue and eigenvector of 5′-flanking pooled DNA sequences (Figure 6).

Although the presented system-theoretic modeling approach has shown an interesting similarity to the work of Qian et al. (*10*), there is one fundamental difference between the current work and Qian's work regarding the formulation of a promoter matrix for the estimation of the measurement variable (i.e., mRNA expression level). A bias-free simple sequential determination of the 5′-flanking DNA sequences of the promoter matrix seems

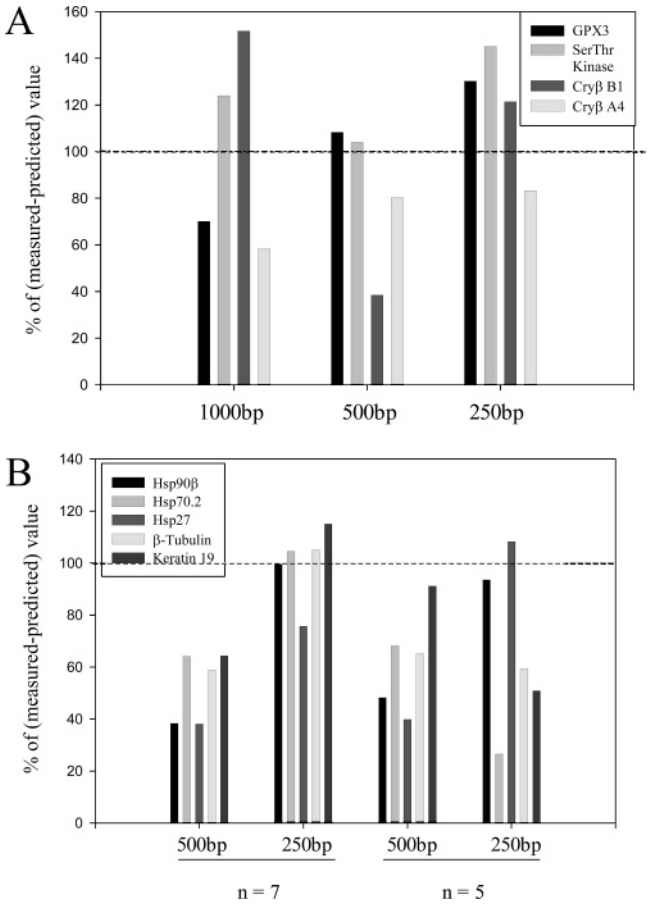

Figure 5. Effect of a promoter size and the number of genes for predicted mRNA expression levels. 5′-Flanking regulatory regions of 250-, 500-, and 1000-bp length are evaluated. The vertical line indicates the percentage of the predicted vs measured value for the selected genes. A transcript level of "*n*" genes is plotted on the *x*-axis, and the percentage of predicted/measured is plotted on the *y*-axis.

to surpass the performance of Qian's complex linear square estimated modeling. However, it is also noteworthy to point out the interspecies variation and correlation among the genes for the successful prediction of mRNA expression levels (Figure 5).

In this study, the selected genes are equally distributed among the gene classes of the oxidoreductase, structural molecule, chaperone, heat shock protein, and the ligand binding. All of the selected genes are known to have a significant role in biological and molecular functioning during cataractogenesis (*17*). Furthermore, the selected genes can be the representative of the overall expression in the lens since they constitute the major gene classes of the lens (*13*, *17*).

As previously demonstrated in Materials and Methods, the combination of cis-acting elements (i.e., eigenvectors) and the associated values (i.e., eigenvalues) results in the predicted role of cis-acting elements. In particular, the positive values of GR, NF-1, and SP1 in the primary eigenvector for crystallin αA suggest a stimulatory effect of transcriptions as shown in Figure 6B. Moreover, a stimulatory role of the glucocorticoid receptor (GR), which is theoretically identified by the 250-bp model, experimentally well accords with the recent study as demonstrated by Jobling et al. (*7*).

In general, temporal or transient alteration of gene expressions is governed by the limited number of transcription factors closely related to a given cis-element of a specific gene. It is intriguing to notice that TBP does

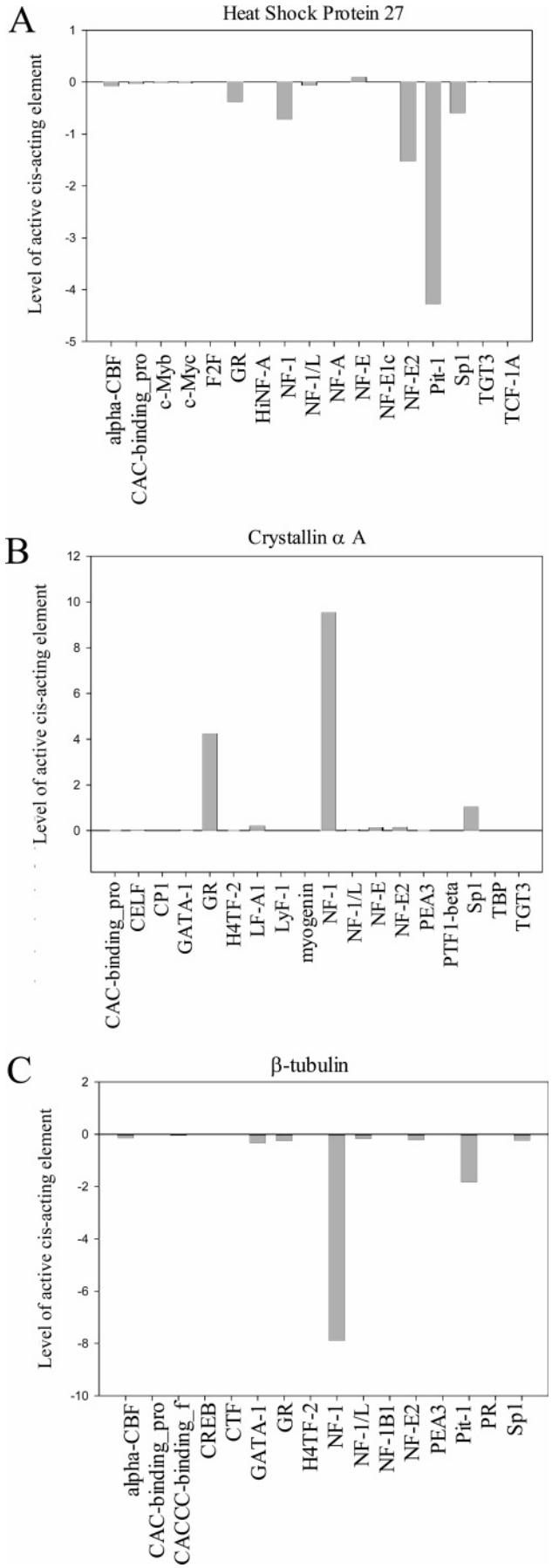

Figure 6. Estimated role of cis-acting elements. The predicted role of active cis-elements for heat shock protein 27, crystalline αA, and *β*-tubulin is demonstrated. The vertical axis ranges from −10 (strong inhibitory) to +10 (strong stimulatory).

not contribute to the αA-crystallin gene expression (Figure 6B) since TBP has been usually considered as one of the essential and general transcription factors. However, previous studies by Sax et al. have also demonstrated that the ability of the αA-crystallin promoter to function in the absence of TBP is mainly due to the role of other proteins such as Sp1, which bound on the αA-crystallin promoter (*18*). Furthermore, the role of transcription factor SP1, which has been previously reported to play a critical role in lens differentiation and cataractogenesis, is in accord with the predicted role for hsp 27, *β*-tubulin, and αA-crystallin (Figure 6B) (*19*−*21*).

Although the accurate identification of 5′ flanking DNA regulatory sequences still remains as a future challenge, we have shown the significance of systems biological approach in terms of predicting gene expression levels (*10*, *12*, *22*). The previous result of expressed sequence tag analysis (*23*) of the human lens showed that the most abundant transcripts in the unnormalized adult human lens library is dominated by the chaperone, structural molecule, and oxidation related transcripts, which is in accord with the result of the present study. However, further experimental confirmation of the model-driven prediction using such as electrophoretic mobility shift assay (EMSA) and reverse transcriptase-polymerase chain reaction (RT-PCR) remains as an important ingredient to strengthen the system-theoretic approach as proposed in this study. Moreover, the comparison of the proposed predictive model with other research groups remains as a further study.

In conclusion, the proposed promoter-based system-theoretic approach has shown the successful prediction of the gene expression levels and the role of cis-acting elements in age-related cataract. This system-theoretic approach is further applicable to normal lens developmental processes based on an extensive interplay between the dry-lab modeling and the wet-lab high throughput microarray mRNA data.

Acknowledgment

This work is supported by a grant from the Korea Ministry of Science and Technology (Korean System Biology Research Grant, M10503010001-05N030100111), Korea Research Foundation Grant (2004-041-DD00280), and the 21C Frontier Microbial Genomics and Application Center Program, Ministry of Science and Technology (Grant MG05-0204-3-0), Republic of Korea.

References and Notes

(1) Francis, P. J.; Berry, V.; Moore, A. T.; Bhattacharya, S. Lens biology: development and human cataractogenesis. *Trends Genet.* **1999**, *15*, 191−196.

(2) Lim, J. M.; Kim, J. A.; Lee, J. H.; Joo, C. K. Downregulated expression of integrin alpha6 by transforming growth factor-beta(1) on lens epithelial cells in vitro. *Biochem. Biophys. Res. Commun.* **2001**, *284*, 33−41.

(3) Hwang, K. H.; Lee, E. H.; Jho, E. H.; Kim, J. H.; Lee do, H.; Chung, S. K.; Kim, E. K.; Joo, C. K. Accumulation and aberrant modifications of alpha-crystallins in anterior polar cataracts. *Yonsei Med. J.* **2004**, *45*, 73−80.

(4) Hejtmancik, J. F.; Kantorow, M. Molecular genetics of age-related cataract. *Exp. Eye Res.* **2004**, *79*, 3−9.

(5) Reddy, M. A.; Francis, P. J.; Berry, V.; Bhattacharya, S. S.; Moore, A. T. Molecular genetic basis of inherited cataract and associated phenotypes. *Surv. Ophthalmol.* **2004**, *49*, 300−315.

(6) Fujii, N.; Awakura, M.; Takemoto, L.; Inomata, M.; Takata, T.; Saito, T. Characterization of alphaA-crystallin from high molecular weight aggregates in the normal human lens. *Mol. Vision* **2003**, *9*, 315−322.

(7) Jobling, A. I.; Stevens, A.; Augusteyn, R. C. Binding of dexamethasone by alpha-crystallin. *Invest. Ophthalmol. Vision Sci*. **2001**, *42*, 1829−1832.

(8) Gao, F.; Foat, B. C.; Bussemaker, H. J. Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. *BMC Bioinformatics* **2004**, *5*, 31.

(9) Conlon, E. M.; Liu, X. S.; Lieb, J. D.; Liu, J. S. Integrating regulatory motif discovery and genome-wide expression analysis. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3339−3344.

(10) Qian, L.; Liu, Y.; Sun, H. B.; Yokota, H. Systems analysis of matrix metalloproteinase mRNA expression in skeletal tissues. *Front. Biosci*. **2002**, *7*, a126−134.

(11) Cho, K. H.; Wolkenhauer, O. Analysis and modelling of signal transduction pathways in systems biology. *Biochem. Soc. Trans.* **2003**, *31*, 1503−1509.

(12) Wolkenhauer, O.; Ullah, M.; Kolch, W.; Cho, K. H. Modeling and simulation of intracellular dynamics: choosing an appropriate framework. *IEEE Trans. Nanobiosci.* **2004**, *3*, 200−207.

(13) Hawse, J. R.; Hejtmancik, J. F.; Huang, Q.; Sheets, N. L.; Hosack, D. A.; Lempicki, R. A.; Horwitz, J.; Kantorow, M. Identification and functional clustering of global gene expression differences between human age-related cataract and clear lenses. *Mol. Vision* **2003**, *9*, 515−537.

(14) Sun, H. B.; Liu, Y.; Qian, L.; Yokota, H. Model-based analysis of matrix metalloproteinase expression under mechanical shear. *Ann. Biomed. Eng.* **2003**, *31*, 171−180.

(15) Kauermann, G.; Eilers, P. Modeling microarray data using a threshold mixture model. *Biometrics* **2004**, *60*, 376−387.

(16) Wilson, A. S.; Hobbs, B. G.; Speed, T. P.; Rakoczy, P. E. The microarray: potential applications for ophthalmic research. *Mol. Vision* **2002**, *8*, 259−270.

(17) Ruotolo, R.; Grassi, F.; Percudani, R.; Rivetti, C.; Martorana, D.; Maraini, G.; Ottonello, S. Gene expression profiling in human age-related nuclear cataract. *Mol. Vision* **2003**, *9*, 538−548.

(18) Sax, C. M.; Cvekl, A.; Kantorow, M.; Gopal-Srivastava, R.; Ilagan, J. G.; Ambulos, N. P., Jr.; Piatigorsky, J. Lens-specific activity of the mouse alpha A-crystallin promoter in the absence of a TATA box: functional and protein binding analysis of the mouse alpha A-crystallin PE1 region. *Nucleic Acids Res.* **1995**, *23*, 442−451.

(19) Alemany, J.; Klement, J. F.; Borras, T.; De Pablo, F. DNA binding factors which interact with the Sp1 site of the chicken delta 1-crystallin promoter are developmentally regulated. *Biochem. Biophys. Res. Commun.* **1992**, *183*, 659−665.

(20) Brunekreef, G. A.; van Genesen, S. T.; Lubsen, N. H. Sp1- and octamer-consensus sequence binding proteins during lens fibre differentiation. *Exp. Eye Res*. **1997**, *64*, 295−299.

(21) Ohtaka-Maruyama, C.; Wang, X.; Ge, H.; Chepelinsky, A. B. Overlapping Sp1 and AP2 binding sites in a promoter element of the lens-specific MIP gene. *Nucleic Acids Res.* **1998**, *26*, 407−414.

(22) Yeung, K. Y.; Medvedovic, M.; Bumgarner, R. E. From coexpression to co-regulation: how many microarray experiments do we need? *Genome Biol*. **2004**, *5*, R48.

(23) Wistow, G.; Bernstein, S. L.; Wyatt, M. K.; Behal, A.; Touchman, J. W.; Bouffard, G.; Smith, D.; Peterson, K. Expressed sequence tag analysis of adult human lens for the NEIBank Project: over 2000 non-redundant transcripts, novel genes and splice variants. *Mol. Vision* **2002**, *8*, 171−184.

Accepted for publication June 6, 2005.

BP050027S