Poster: Comprehensive pharmacogenomic pathway screening by data assimilation

Conference Paper · February 2011

DOI: 10.1109/ICCABS.2011.5729899 · Source: DBLP

5 authors, including:

Seiya Imoto
The University of Tokyo
466 PUBLICATIONS **12,086** CITATIONS
SEE PROFILE

Satoru Miyano
The University of Tokyo
1,000 PUBLICATIONS **30,538** CITATIONS
SEE PROFILE

Comprehensive Pharmacogenomic Pathway Screening by Data Assimilation

Takanori Hasegawa, Rui Yamaguchi, Masao Nagasaki, Seiya Imoto, and Satoru Miyano

Human Genome Center, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
{t-hasegw,ruiy,masao,imoto,miyano}@ims.u-tokyo.ac.jp

Abstract. We propose a computational method to comprehensively screen for pharmacogenomic pathway simulation models. A systematic model generation strategy is developed; candidate pharmacogenomic models are automatically generated from some prototype models constructed from existing literature. The parameters in the model are automatically estimated based on time-course observed gene expression data by data assimilation technique. The candidate simulation models are also ranked based on their prediction power measured by Bayesian information criterion. We generated 53 pharmacogenomic simulation models from five prototypes and applied the proposed method to microarray gene expression data of rat liver cells treated with corticosteroid. We found that some extended simulation models have higher prediction power for some genes than the original models.

1 Introduction

Construction and simulation of biological pathways are crucial steps in understanding complex networks of biological elements in cells [4, 7, 8, 9, 13, 15, 16]. To construct simulatable models, structures of networks and chemical reactions are collected from existing literature and the values of parameters in the model are set based on the results of biological experiments or estimated based on observed data by some computational method [9]. However, it is possible that there are some missing relationships or elements in the literature-based networks. Therefore, we need to develop a computational strategy to improve a prototype model and create better ones that can predict biological phenomena.

To propose novel networks of genes, statistical graphical models including Bayesian networks [3] and vector autoregressive models [5, 11] have been applied to gene expression data. An advantage of these methods is that we can find networks with a large number of genes and analyze them by a viewpoint of systems. However, due to the noise and the limited amount of the data, some parts of the networks estimated by these methods are not biologically reasonable and cannot be validated. In this paper, we focus on another strategy. Unlike the

J. Chen, J. Wang, and A. Zelikovsky (Eds.): ISBRA 2011, LNBI 6674, pp. 160–171, 2011.

statistical methods, our method can create a set of extended simulatable models from prototype literature-based models.

There are two key points in our proposed strategy: One is that various structures of candidate simulation models are systematically generated from the prototypes. The other is that, for each created model, the values of parameters are automatically estimated by data assimilation technique [9, 16]; the values of parameters will be determined by maximizing the prediction capability of the model. For each of simulation models, by using data assimilation technique, we can discover that which genes are appropriately predicted their temporal expression patterns by the candidate model. Since we consider pharmacogenomic pathways, these genes are possibly placed on the mode-of-action of target chemical compound. The results obtained by our proposed strategy could be essential to create a larger and more comprehensive simulation model and systems biology driven pharmacology.

To show the effectiveness of the proposed strategy, we analyze time-course microarray data of rat liver cells treated with corticosteroid [2]. In the previous study, differential equation-based simulation models, named fifth generation model [12], were used and predictable expression patterns by this model were discussed for 197 genes selected by clustering analysis [2]. In this paper, we systematically generated 53 simulatable models from five prototypes and determined which 58 models suitably predict expression pattern of each gene. Finally, we show a comprehensive pharmacogenomics pathway screening that elucidates associations between genes and simulation models.

The paper is organized as follows: In Section 2, we elucidate a systematic method to create extended simulation models from prototype ones. The parameter estimation based on data assimilation technique with particle filter [9, 16] and a model selection method [6, 10] are also presented. We apply the proposed pharmacogenomic pathway screening strategy to constructed 58 models and time-course gene expression data of rat liver cells with corticosteroid in Section 3. Discussions are given in Section 4.

2 Method

2.1 Corticosteroid Pharmacokinetic and Pharmacogenomics Models

We first introduce a framework of pharmacokinetic and pharmacogenomic models employed in Jin et al. [2]. Under this framework, a pharmacokinetic model that represents a plasma concentration of methylprednisolone (MPL) in nanograms per milliliter, C_{MPL}, is given by

$$C_{\mathrm{MPL}} = C_1 \cdot e^{-\lambda_1 t} + C_2 \cdot e^{-\lambda_2 t}, \tag{1}$$

where C_1, C_2, λ_1 and λ_2 are coefficients for the intercepts and slopes and Jin et al. [2] set by $C_1 = 39,130$ (ng/ml), $C_2 = 12,670$ (ng/ml), $\lambda_1 = 7.54$ (h^{-1}) and $\lambda_2 = 1.20$ (h^{-1}). These values are obtained from other biological experiments than gene expression profilings that we will use for parameter estimation of

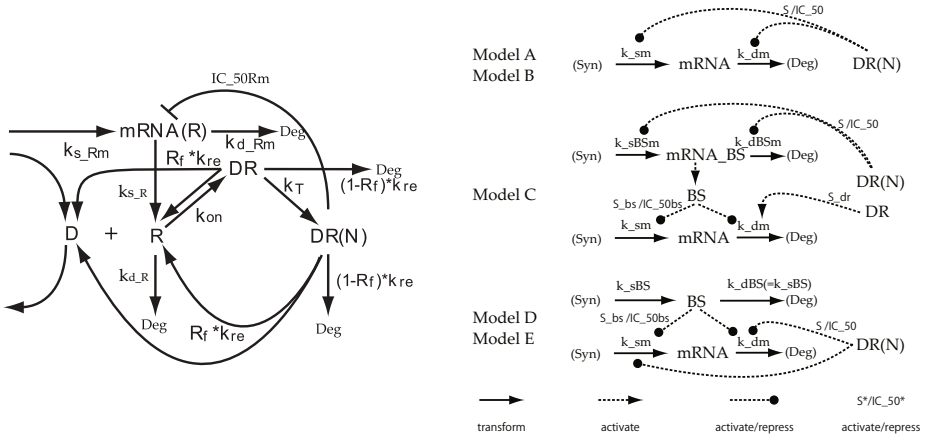

Fig. 1. The left figure and right figure shows core model for corticosteroid pharmacokinetics and prototype pharmacogenomic models with extensions respectively. In the right figure, the dashed lines with circle are the candidate relations to be extended and BS is the intermediate biosignal.

pharmacogenomic models described in the latter section. We thus use these four values for the corticosteroid pharmacokinetics.

In existing literature, corticosteroid pharmacogenomic pathways were investigated [2]. We show the core part of the pathway that includes corticosteroid, represented by D, and its receptor, R, in Figure 1 (left). Here, mRNA(R) denotes the mRNA of the receptor, DR is cytosolic drug-receptor complex and DR(N) is drug-receptor complex in nucleus. The reaction parameters in Figure 1 (left) were set according to Sun et al. [12] and summarized in Table 1 (left). The dynamics of the pathway can be represented by four differential equations given by

$$\frac{d\mathrm{mRNA(R)}}{dt} = k_{s\_Rm} \cdot \left\{1 - \frac{\mathrm{DR(N)}}{\mathrm{IC_{50\_Rm}} + \mathrm{DR(N)}}\right\} - k_{d\_Rm} \cdot \mathrm{mRNA(R)}, \tag{2}$$

$$\frac{d\mathrm{R}}{dt} = k_{s\_R} \cdot \mathrm{mRNA(R)} + R_f \cdot k_{re} \cdot \{\mathrm{DR(N)} + \mathrm{DR}\} - k_{on} \cdot \mathrm{D} \cdot \mathrm{R} - k_{d\_R} \cdot \mathrm{R}, \tag{3}$$

$$\frac{d\mathrm{DR}}{dt} = k_{on} \cdot \mathrm{D} \cdot \mathrm{R} - (k_T + k_{re}) \cdot \mathrm{DR}, \tag{4}$$

$$\frac{d\mathrm{DR(N)}}{dt} = k_T \cdot \mathrm{DR} - k_{re} \cdot \mathrm{DR(N)}. \tag{5}$$

Based on the fundamental model represented in Figure 1 (left), we want to know how DR and DR(N) affect other genes in transcriptional level. As a basic pharmacogenomic model for finding relationship between drug-receptor complex and other genes, we consider extending five pharmacogenomic models [2] shown

Table 1. Parameter Setting for the core model and for the constructed pharmacogenomic models

Fixed Parameter	Value	Unit
k_{s_Rm}	2.90	fmol/g/h
k_{d_Rm}	0.1124	fmol/g/h
IC_{50_Rm}	26.2	fmol/mg
k_{on}	0.00329	l/nmol/h
k_T	0.63	h^{-1}
k_{re}	0.0572	h^{-1}
R_f	0.49	
k_{s_R}	1.2	h^{-1}
k_{d_R}	0.0572	h^{-1}
$mRNA_R^0$	25.8	fmol/g
R^0	540.7	fmol/mg

Estimated Parameter	Model	Unit
k_sm	All	l/nmol/h
k_dm	All	l/nmol/h
S or IC_50	All	l/nmol/h or fmol/mg
k_sBSm	C	l/nmol/h
k_dBSm	C	l/nmol/h
k_sBS	C, DE	l/nmol/h
k_dBS	C	l/nmol/h
S_bs or IC_50bs	C, DE	l/nmol/h or fmol/mg
S_dr	C	l/nmol/h
$mRNA\_BS^0$	C	fmol/mg
BS^0	DE	fmol/mg

in Figure 1 (right). The original five pharmacogenomic pathways [2] have the same elements as the core pharmacokinetic pathway, DR and DR(N), and represent relationships between corticosteroid and its downstream genes. However, more variations can be considered as candidates of pharmacogenomic pathway of corticoid. Therefore, from these five models, we automatically constructed 53 models with the following three rules.

(i) If a regulator, DR(N), DR or BS, activates (represses) the synthesis (degradation) of mRNA, a revised model tests to repress (activate) the degradation (synthesis) of mRNA. However, we do not consider combination effects of them.

(ii) If two regulators regulate the same element, we also consider either two regulator model or one regulator model that is defined by removing one of two edges.

(iii) If two regulators regulate the same element, we consider either independent regulation model that employs additive form or cooperative regulation model with the product of the regulators.

We create these rules for generating simulation models that covers all patterns of regulations when we do not change the number of elements such that mRNAs and proteins in each simulation model.

From Model A: One model with three parameters ("k_sm", "k_dm" and "S or IC_50") was generated by applying the rule (i). These models include only mRNA and can simply represent activation of mRNA expression.

From Model B: One model with three parameters ("k_sm", "k_dm" and "S or IC_50") was generated by applying the rule (i). These models include only mRNA and can simply represent repression of mRNA expression.

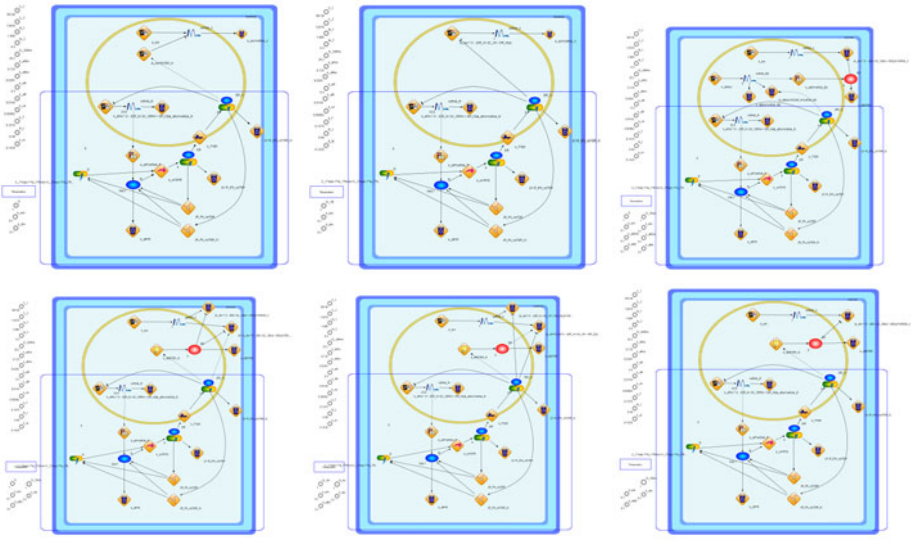

Fig. 2. Six representative pharmacogenomic simulation models (From top left to right, Model A, B, C12, DE10, DE12 and DE20). These models have high predictive power for many of 8799 rat liver genes. These models are described by Cell Illustrator 5.0.

From Model C: First, 15 models with 11 or 10 parameters ("k_sm", "k_dm", "S or IC_50", "k_sBSm", "k_dBSm", "k_sBS", "k_dBS", "S_bs or IC_50bs", "S_dr", and "initial values of mRNA_BS" and "BS") were generated by applying the rule (i) and (ii). These models include mRNA, BS, and mRNA_BS. Since DR is included only in Model C, we evaluate the necessity of the presence of DR by creating models without DR (rule (ii)). Therefore, 16 models that do not have DR were additionally created and finally we have 31 models from Model C.

From Model DE: 20 models with 5 or 6 parameters ("k_sm", "k_dm", "S or IC_50", "k_sBS", "S_bs" or "IC50_bs", and "initial value of BS") were generated by applying the rules (i), (ii) and (iii). These models include mRNA and BS. We unified the notation of Model D and E, because these two models are similar and the extended models are hard to be separated. We constructed 16 models, 4 models and 2 models according to rule (i), (ii) and (iii) respectively. In these simulation models, the parameters, "k_sBSm", "k_sBS", "k_sm", "BS^0 (initial concentration of BS)" and "$mRNA^0_{BS}$ (initial concentration of $mRNA_{BS}$)" were fixed in the original work [2], but we estimate these five parameters together with the other parameters.

For these 53 and original 5 pharmacogenomic models, we estimate the values of parameters by using time-course microarray gene expression data from liver cells of rats received glucocorticoid. We also evaluate which models can predict the expression profiles of each gene; it enables us to find better pharmacogenomic models for each gene. For this purpose, a mathematical technique called data assimilation for parameter estimation and model selection is described in the next section.

2.2 Data Assimilation for Parameter Estimation and Model Selection

To perform simulations by the pharmacogenomic models described in the previous section, we implemented them using Cell Illustrator [8], a software for biological pathway simulation based on hybrid functional Petri net with extensions. Six representative models in Cell Illustrator are shown in Figure 2.

Let $y_j[t]$ be the expression value of jth gene at time t and let $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\theta})$ be a simulation model, where $\boldsymbol{x}$ is a vector of variables in the simulation model and $\boldsymbol{\theta}$ is a parameter vector described in the previous section. For example, $\boldsymbol{x}$ includes the concentration of drug-recepter complex, DR. The simulation variable $\boldsymbol{x}$ will be updated by a system model:

$$\boldsymbol{x}_t = \boldsymbol{f}(\boldsymbol{x}_{t-1}, \boldsymbol{\theta}) + \boldsymbol{v}_t, \quad t \in \mathcal{N}, \tag{6}$$

where $\boldsymbol{x}_t$ is the vector of values for the simulation variables at time t, $\boldsymbol{v}_t$ represents innovation noise and N is the set of simulation time points and set $N = \{1, ..., T\}$. To connect the simulation model with the observed data, we formulate an observation model:

$$y_j[t] = h(\boldsymbol{x}_t) + w_t, \quad t \in \mathcal{N}_{obs}, \tag{7}$$

where h is a function that maps simulation variables to the observation and w_t is an observation noise. Here, N_{obs} is the set of time points that we measured gene expression data. We should note that N_{obs} is a subset of N. In our case, since $\boldsymbol{x}_t$ contains a variable representing the abundance of mRNA of the gene, i.e., the jth gene in Eq. (7), the function h takes out the element of $\boldsymbol{x}_t$ corresponding to $y_j[t]$. The model constructed by combining Eq.s (6) and (7) is called a nonlinear state space model. To simplify the notation, we assume $N_{obs} = N$, however, it is easy to generalize the theory described below to the case of $N_{obs} \subset N$.

The parameter vector $\boldsymbol{\theta}$ is estimated by the maximum likelihood method that chooses the values of $\boldsymbol{\theta}$ that maximize the likelihood

$$L(\boldsymbol{\theta}|Y_{jT}) = \int p(\boldsymbol{x}_0) \prod_{t=1}^{T} p(y_j[t]|\boldsymbol{x}_t) p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{\theta}) d\boldsymbol{x}_1 \cdots d\boldsymbol{x}_T,$$

where $Y_{jT} = (y_j[1], ..., y_j[T])$. For the computation of the likelihood, we use the particle filter algorithm [9]. For details of the particle filter algorithm for biological pathway model, we refer Nagasaki et al. [8] and Koh et al. [4]. In the parameter estimation, we restricted the values of parameters so that they take positive and not so large from a biological point of view.

For the comparison of multiple simulation models $\boldsymbol{f}_1, ..., \boldsymbol{f}_M$, we employ Bayesian information criterion (BIC) [10]. For the mth model, $\boldsymbol{f}_m$, BIC is defined by

$$\mathrm{BIC}(\boldsymbol{f}_m) = -2 \log L(\hat{\boldsymbol{\theta}}_m|Y_{jT}) + \nu_m \log T,$$

where $\hat{\boldsymbol{\theta}}_m$ is the maximum likelihood estimate of the vector of parameters in $\boldsymbol{f}_m$ and ν_m is the dimension of $\boldsymbol{\theta}_m$. Therefore, for the jth gene, the optimal simulation model, $\boldsymbol{f}^*$, can be obtained by

$$\boldsymbol{f}^* = \arg\min_{f_m} \mathrm{BIC}(\boldsymbol{f}_m).$$

The model ranking for a gene can also be determined by the values of BIC.

3 Pharmacogenomic Pathway Screening for Corticosteroid 58 Models

3.1 Time-Course Gene Expressions

We analyze microarray time-course gene expression data of rat liver cells [2]. The microarray data were downloaded from GEO database (GSE487). The time-course gene expressions were measured at 0, 0.25, 0.5, 0.75, 1, 2, 4, 5, 5.5, 7, 8, 12, 18, 30, 48 and 72 hours (16 time-points) after receiving glucocorticoid. The data at time 0 hour are control (non-treated). The number of replicated observations is 2, 3 or 4 at a time point.

3.2 Results of Pathway Screening with Data Assimilation

First, we focused on 197 genes that were identified by the previous work [2] as the drug-affected genes by the clustering analysis. For the genes in each cluster, we explored which simulation models have better prediction power and the results are summarized in Figure 3. According to the results obtained previously [2, 12], the genes in the clusters 1, 2, 3, 4, 5 and 6 were reported to be well predicted by the Models "A", "A", "C", "D or E", "cell-cell interaction model"and "B, D or E", respectively. This result indicated that the genes in the cluster 1, 2 have almost same expression profiles. We should note that the cell-cell interaction model is not included in the five prototype models.

Figure 3 shows the results for each cluster and the gene expression profiles. We can summarize the results as follows:

Cluster 1: The previous research [2] suggested that these genes are well predicted by Model A. However, interestingly, in our results, Model A was selected few times. On the other hand, Models D and E and their extended models were selected many times. We presume the reason is that, particularly in the first part, the profiles of these genes are not so simple.

Cluster 2: These genes are also suggested to be suitably predicted with Model A. Like cluster 1, similar results, however, were obtained; for these genes, Model A was not selected in many times.

Cluster 3: The previous research [2] suggested that these genes fitted to Model C. However, in our results, not so many genes in cluster 3 are well predicted by Model C, but they fit to Models D and E and their extended models. We guess the reason is that Model C has more parameters than necessary. Therefore, in BIC, the second term, i.e., penalty for the number of parameters, takes

large value and BIC cannot be small, so Model C and its extended versions were not selected. The same things can be said from the other works [1, 14].

Cluster 4: These genes were suggested to be fit with Models D or E. In our results, Model B and its extension and extension of Model A fit well, and Model E is especially fit, but Model D is not selected much. Instead, some extended versions of Models D and E fit well. The genes in cluster 4, we can see that some expression profiles do not vary widely. Such genes are well fit to Models A, B and its extensions, because of these simplicity. On the other hand, Models D and E and their extended models can follow complex behaviors and were selected in many times for other genes.

Cluster 5: Since these genes were judged to be fitted with the cell-cell interaction model that is not included in the five prototypes, these genes are not covered by our prepared models. However, in practice, the extended models of Model DE showed high predictive power for these genes. The expression profiles of these genes show sudden increasing patterns. Actually, our models can represent such dynamic patterns of gene expression profiles.

Cluster 6: These genes were suggested to be fit with Models B, D and E, but most genes were selected as the extended models of Models D and E. We presume the reason is that Models D and E are flexible and can follow various types of complex expression patterns.

We next illustrate the results of pharmacogenomic pathway screening for whole 8799 rat liver genes. Figure 4 shows the results with heatmap of the selected top 5 models for each gene and time-course expression profiles of genes that are specific for Models C6, C12, DE10 and DE12. For each gene, we test the significance of the top ranked simulation model by using Smirnov–Grubbs test. If the expression profile of a gene was predicted very well by several simulation models, we cannot find pharmacogenomic mechanism specific for the gene. However, if only one model could predict the behavior of a gene, the model is a strong candidate that represents corticosteroid's mode-of-action for the gene. In such a case, we say the gene is specific for the above model.

Unlike the genes from the clustering analysis, two prototype models, Models A and B, were selected as top 5 in many times. We presume the reason is that, in the whole gene, there are some genes whose expression patterns are somewhat flat (not show clear dynamic patterns) and Models A and B can follow them with a small number of parameters. Although the prototype D and E models were not selected many times, their extended models were frequently selected as top 5. This suggests that Models D and E can work well as the seed models for generating other simulation models with higher predictive power. The amount of genes obtained by this test varied widely depending on the models. From ModelA1, B1, C6, C12, C16, DE2, DE10, DE12 and DE20, we can obtained some specific genes. Interestingly, the number of genes fitting to Model C is relatively low, but many specific genes are obtained by Model C. It suggest that there are some expression profiles that can be represented by only the one of Model C. We then perform a functional analysis in order to reveal enriched gene

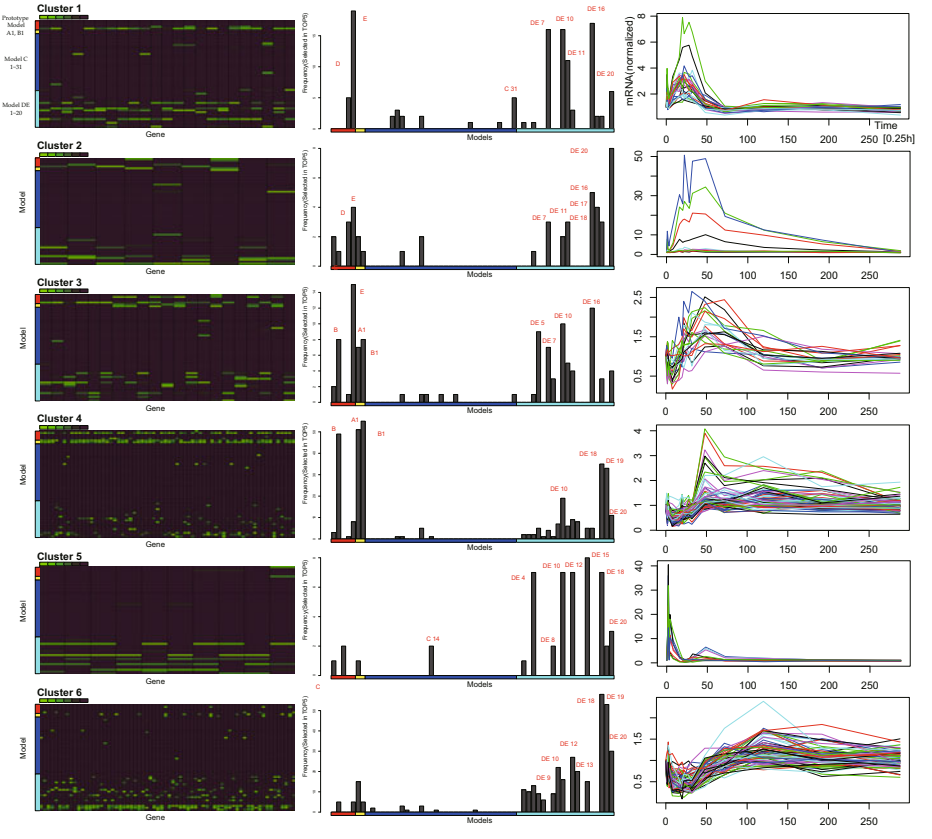

Fig. 3. Top 5 simulation models for each gene in a cluster defined by Jin et al. [2] are represented by a heat map. The green elements means that the model well fits to the gene expression profiles. The histograms of the frequencies of the models selected as top 5 are shown in the middle panels, and gene expression profiles are also shown in the right panels.

functions for each set of Model-specific genes. For the functional analysis, we used Ingenuity and the results can be summarized as follows:

ModelC_6: These genes have function of "Cellular Assembly and Organization" and "RNA Post–Transcriptional Modification" and relate to "Protein Ubiquitination Pathway". **ModelC_12:** These genes are most interesting genes. These have "Amino acid Metabolism", "Nucleic Acid Metabolizm", "Cell Death", "Cellular Grows and Proliferation", "Drug Metabolism" and "Lipid Metabolism" and so on. Additionally, these genes relate to "Aldosterone Signaling Epithelial Cells" and "Glucocorticoid Recepter Signaling". Beneficial effects of Corticosteroid is inhibition of immune system and adverse effect is numerous metabolic side effects, including osteoporosis, muscle wasting, steroid diabetes, and others. Therefore, these result in ModelC_12 is biologically significant because these genes may have a function concerning metabolic side effects. **ModelDE_10:** These genes are

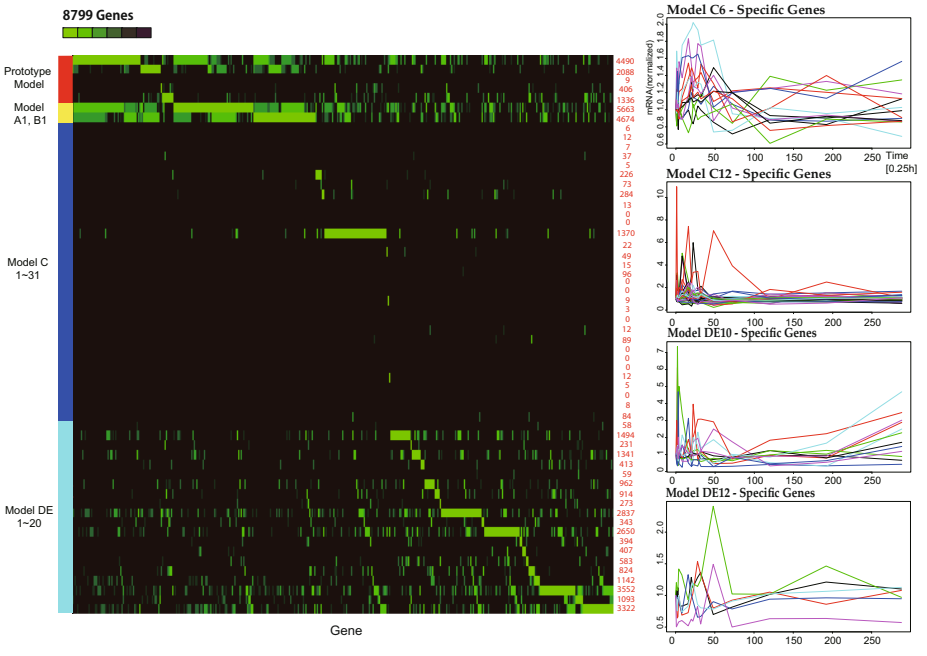

Fig. 4. The result of comprehensive pharmacogenomic pathway simulation model screening. Heat map for top 5 models is shown from 58 simulation models for 8799 rat liver genes. Time-course expression profiles are shown for genes that are specific for Models C6, C12, DE10 and DE12.

also interesting. The functions are "Neurological Disease", "Organismal Injury and Abnormalities" and "Immunological Disease", and are affected by "Graft–versus–Host Disease", "Autoimmune Thyroid Disease, "T Helper Cell Differentiation" and so on. Because of the above therapeutic and adverse effects of CS, the function of these genes are also significant concerning immune system function. **ModelDE_12:** The functions of these genes are "Cellular Development", "Cardiovascular Disease", and "Hematological Disease". These are also affected by "EIF2 signaling".

We consider that such genes are important among 8799 genes, because these were estimated to have a similar pathway and it may be difficult to collect these genes by clustering analysis simply using the gene expression profile.

4 Discussion

In this paper, we proposed a computational strategy for automatic generation of pharmacogenomic pathway simulation models from the prototype simulation models that are built based on literature information. The parameters in the constructed simulation models were estimated based on the observed time-course

gene expression data measured by dosing some chemical compound to the target cells. We constructed totally 58 pharmacogenomic simulation models on a pathway simulation software, Cell Illustrator, and used data assimilation technique for parameter estimation. For pathway screening, we introduce Bayesian information criterion for pathway model selection in the framework of data assimilation. We performed comprehensive pathway screening for constructed 58 pharmacogenoimc simulation models with gene expression data of rat liver cells treated with glucocorticoid.

The prototype five models fit to somewhat large number of genes well. However, there are more extended models that can predict the dynamic patterns of gene expressions better than the prototypes. This suggests that, from the prototype simulation models, we can automatically construct various extended simulation models and some of them could have higher prediction ability than the originals. Also, we performed a functional analysis to the sets of Model-specific genes identified by the Smirnov-Grubbs test. As shown above, some meaningful functions were found. We would like to discuss the relationship between Model-specific genes and enriched function in future paper with biological evidences.

We consider the followings as our future research topics. We simply use the pharmacokinetic model described in Section 2. However, we can generated many candidates and may construct true model from observed data by data assimilation technique. Also, we may combine multiple simulation models to create bigger one. As we mentioned before, data analysis based on statistical methods like Bayesian networks can produce network information that would be affected by a chemical compound. It should be useful if we combine the results from statistical data analysis with pharmacogenomic pathway simulations.

Acknowledgments. The computational resource was provided by the Super Computer System, Human Genome Center, Institute of Medical Science, University of Tokyo.

References

1. Hazra, A., Dubois, C.D., Almon, R.R., Snyder, H.G., Jusko, J.W.: Pharmacodynamic Modeling of Acute and Chronic Effects of Methylprednisolone on Hepatic Urea Cycle Genes in Rats. Gene Regulation and System Biology 2, 1–19 (2008)
2. Jin, Y.J., Almon, R.R., Dubois, D.C., Jusko, W.J.: Modeling of corticosteroid pharmacogenomics in rat liver using gene microarrays. The Journal of Pharmacology and Experimental Therapeutics 307(1), 93–107 (2003)
3. Kim, S., Imoto, S., Miyano, S.: Dynamic Bayesian network and nonparametric regression for nonlinear modeling of gene networks from time series gene expression data. Biosystems 75(1-3), 57–65 (2004)
4. Koh, C.H., Nagasaki, M., Saito, A., Wong, L., Miyano, S.: DA1.0: Parameter estimation of biological pathways using data assimilation approach. Bioinformatics 26(14), 1794–1796 (2010)
5. Kojima, K., Yamaguchi, R., Imoto, S., Yamauchi, M., Nagasaki, M., Yoshida, R., Shimamura, T., Ueno, K., Higuchi, T., Gotoh, N., Miyano, S.: A state space representative of VAR models with sparse learning for dynamic gene networks. Genome Informatics 22, 59–68 (2009)

6. Konishi, S., Ando, T., Imoto, S.: Bayesian information criteria and smoothing parameter selection in radial basis function networks. Biometrika 91(1), 27–43 (2004)
7. Matsuno, H., Inoue, S., Okitsu, Y., Fujii, Y.: A new regulatory interaction suggested by simulations for circadian genetic control mechanism in mammals. Journal of Bioinformatics and Computational Biology 4(1), 139–153 (2006)
8. Nagasaki, M., Yamaguchi, R., Yoshida, R., Imoto, S., Doi, A., Tamada, Y., Matsuno, H., Miyano, S., Higuchi, T.: Genomic data assimilation for estimating hybrid functional petri net from time-course gene expression data. Genome Informatics 17(1), 46–61 (2006)
9. Nakamura, K., Yoshida, R., Nagasaki, M., Miyano, S., Higuchi, T.: Parameter estimation of *In Silico* biological pathways with particle filtering toward a petascale computing. In: Pacific Symposium on Biocomputing, vol. 14, pp. 227–238 (2009)
10. Schwarz, G.: Estimating the dimension of a model. Ann. Statist. 6, 461–464 (1978)
11. Shimamura, T., Imoto, S., Yamaguchi, R., Fujita, A., Nagasaki, M., Miyano, S.: Recursive elastic net for inferring large-scale gene networks from time course microarray data. BMC Systems Biology 3, 41 (2009)
12. Sun, Y., Dubois, D.C., Almon, R.R., Jusko, W.J.: Fourth-generation model for corticosteroid pharmacodynamics: A model for methylprednisolone effects on receptor/gene-mediated glucocorticoid receptor down-regulation and tyrosine aminotransferase induction in rat liver. Journal of Pharmacokinetics and Biopharmaceutics 26(3), 289–317 (1998)
13. Tasaki, S., Nagasaki, M., Oyama, M., Hata, H., Ueno, K., Yoshida, R., Higuchi, T., Sugano, S., Miyano, S.: Modeling and estimation of dynamic EGFR pathway by data assimilation approach using time series protemic data. Genome Informatics 17(2), 226–238 (2006)
14. Yao, Z., Hoffman, P.E., Ghimbovschi, S., Dubois, C.D., Almon, R.R., Jusko, W.J.: Mathematical Modeling of Corticosteroid Pharmacogenomics in Rat Muscle following Acute and Chronic Methylprednisolone Dosing. Molecular Pharmaceutics 5(2), 328–339 (2007)
15. Yamaguchi, R., Imoto, S., Yamauchi, M., Nagasaki, M., Yoshida, R., Shimamura, T., Hatanaka, Y., Ueno, K., Higuchi, T., Gotoh, N., Miyano, S.: Predicting difference in gene regulatory systems by state space models. Genome Informatics 21, 101–113 (2008)
16. Yoshida, R., Nagasaki, M., Yamaguchi, R., Imoto, S., Miyano, S., Higuchi, T.: Bayesian learning of biological pathways on genomic data assimilation. Bioinformatics 24(22), 2592–2601 (2008)