Apoptosis (2009) 14:236–245
DOI 10.1007/s10495-008-0288-4

ORIGINAL PAPER

Nonlinear cooperation of p53-ING1-induced bax expression and protein *S*-nitrosylation in GSNO-induced thymocyte apoptosis: a quantitative approach with cross-platform validation

Shaojin Duan · Lin Wan · Wenjiang J. Fu · Hong Pan · Qi Ding · Chang Chen · Peiwei Han · Xiaoyan Zhu · Liying Du · Hongxiao Liu · Yuxia Chen · Ximing Liu · Xiting Yan · Minghua Deng · Minping Qian

Published online: 12 December 2008

Abstract Increasing evidence has been gathered for p53-dependent apoptosis, but it is still unclear how p53 initiates apoptosis by employing its transcriptional program. Pairwise interactions of p53 with expression of other genes fail to predict p53 levels or rate of apoptosis. A more sophisticated approach, using neural networks, permits prediction of interaction among three or more genes (p53, bax, and ING1). These interactions are decidedly nonlinear. Careful measurements and advanced mathematical treatments will permit us not only to understand how expression of pro- and anti-apoptotic genes is regulated, but also to integrate cross-platform and cross-experimental data for the validation of predicted interactions.

Keywords Thymocyte apoptosis · p53 · Bax · ING1 · Neural network · Normalization

Abbreviations

NO	Nitric oxide
GSNO	*S*-Nitrosoglutathione
Dex	Dexamethasone

Introduction

Thymocyte positive/negative selection is a crucial stage in development of thymocyte, and negative selection plays a key role in T cell developing process via thymocyte apoptosis pathway within thymus. Thymocyte apoptosis is an intricate process coupled with negative selection depending on integrated diverse endogenous and exogenous signals to sustain homeostasis in the immune system. Nitric oxide (NO) and its donors *S*-nitrosoglutathione (GSNO) are able to induce immune cell apoptosis, such as macrophages, thymocytes, lymphocytes and endothelial cells via various signal pathways, especially via *S*-nitrosylation/denitrosylation as a reversible redox switch [1]. GSNO-derived NO concentration in immune cells plays a significant role in determining thymocyte fate: whether GSNO-induced thymocyte apoptosis or the inhibition of apoptosis implying immature thymocyte survival to develop T lymphocyte. The mechanisms on GSNO-

Shaojin Duan and Lin Wan contributed equally to this work.

S. Duan · X. Zhu · H. Liu · Y. Chen · X. Liu
Guang An Men Hospital, China Academy of Chinese Medicine Sciences, 100053 Beijing, People's Republic of China

L. Wan · X. Yan · M. Deng · M. Qian (✉)
School of Mathematical Sciences and Center for Theoretical Biology, Peking University, 100871 Beijing, People's Republic of China
e-mail: qianmp@math.pku.edu.cn

W. J. Fu · Q. Ding
Department of Epidemiology, Michigan State University, East Lansing, MI 48824, USA

H. Pan
Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

C. Chen · P. Han
National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 100101 Beijing, People's Republic of China

L. Du
College of Life Sciences, Peking University, 100871 Beijing, People's Republic of China

initiated mouse thymocyte apoptosis involve the intricate regulations of various genes, among which tumor suppressor gene p53 plays a vital role during the processes of apoptosis [2].

It is indicated that when the apoptosis pathway is triggered, p53 undergoes numerous modifications that result in its stabilization and accumulation in the cell [3]. p53 not only regulates the expression of numerous downstream pro-apoptotic genes (e.g., bax, noxa, puma, bid and cd95) in the nucleus, but also accumulates in the cytoplasm and directly activates the pro-apoptotic protein BAX [4]. The gene bax is a direct transcriptional target of transcription factor p53, and bax deficiency decreases apoptosis and accelerates oncogenesis [4–6]. The tumor suppressor gene ING1 has also been studied extensively. It shares many biological functions with p53, such as growth arrest and apoptosis [7, 8]. As a component of the p53 signaling pathway, ING1 can cooperate in parallel with p53 in cell cycle and apoptosis control [9, 10], and positively regulate the expression levels of the downstream gene bax [11, 12].

Although growing evidence has been gathered for p53-dependent apoptosis, it still remains largely unclear how p53 initiates apoptosis by employing its transcriptional program through a complex interdependent regulatory network [13]. It is a common practice to characterize the regulation in an "all-or-none" manner: the role a gene plays in certain cell line or biological processes is characterized as either "independent" or "dependent"; the relation between genes is characterized as either "correlated" or "uncorrelated". However, such an approach may not be adequate for the comprehension of intricate regulatory networks, especially when a much more subtle quantitative interrelationship among multiple genes and their corresponding products are examined and interrogated through technologies, such as quantitative real time PCR and microarray.

In this study, we examined the complex molecular mechanism of GSNO-induced mouse thymocyte apoptosis at molecular level via *S*-nitrosylation proteins and gene expression of the related genes, p53, ING1 and bax. The GSNO-induced mouse thymocyte apoptosis was detected and confirmed with MoFlo cell sorter, and the expressions of p53, ING1 and bax were measured with real time PCR (the raw data of which were processed by our new normalization method). Meanwhile, protein *S*-nitrosylation was detected by biotin-switch and western blotting.

To quantify the characteristics of the cooperative effect of p53 and ING1 in regulating the expression of bax, we employed a statistical non-parametric conditional expectation model via neural networks to identify the nonlinear response surface that accurately predicted the expression level of bax based on the corresponding p53 and ING1 levels. This regulatory network model was trained based on our real time PCR data and showed consistent cross-platform and cross-experimental performance gain in predicting the expression of bax. When the model was applied to four publicly available mouse microarray data sets and pre-processed similarly using our normalization procedure, similar consistency and performance were displayed.

Such consistency suggests that, provided that proper statistical models and methods are used for data preprocessing and analysis, it is plausible to precisely quantify the behavior of an individual gene product based on the gene-expression patterns in a network of interacting gene products. Here, we demonstrated that selecting the proper features (explanatory and response variables) and normalization methods for real time PCR and microarray data is crucial in data analysis and interpretation. In this study, we used relative gene expression measurements, which show better consistency in cross-platform data [14, 15], and the baselines for cross-platform measurements were also carefully adjusted.

Materials and methods

Experiments

Details on treatments and experiments are given in supplementary Material (http://ctb.pku.edu.cn/~wanlin/result/apoptosis/experiment_note.pdf).

Data

Raw data of our real time PCR are available from http://ctb.pku.edu.cn/~wanlin/result/apoptosis/rawdata.xls.

All 4 mouse microarray data sets were obtained from Gene Expression Omnibus (GEO) repository at the National Center for Biotechnology Information (NCBI) with the Accession No. GDS658 (thymocyte selection by agonist [16], GSE2128 (thymocyte negative selection [17], GSE3039 (innate vs. adaptive lymphocyte gene expression [18]) and GDS882 (neuromedin U effect on type-2 Th cells [19]). All these 4 studies were conducted on Affymetrix platform GPL81. The expressions of ING1, p53 and bax were obtained from probe-sets "94396-at" (ING1), "104154-at" (TRP53), and "93536-at" (bax), while the expressions of beta-actin were obtained by taking the averages of probe-sets "AFFX-b-ActinMur/M12481-3-at", "AFFX-b-ActinMur/M12481-5-at" and "AFFX-b-ActinMur/M12481-M-at" (ACTB) of each sample.

Pre-processing and normalization of real time PCR data

A novel normalization method for absolute quantification data of real time PCR is described as follows with a step by step procedure.

1. We perform a logarithm transformation to remove nonlinear effect in the raw data [20]. The log-transformed data shown in Fig. 1b are aligned more closely to a regression line in each plot while the raw data (Fig. 1a) are not.
2. For the log-transformed data, we conduct a novel normalization algorithm based on a statistical linear model with additive measurement errors [21]:

$$X_{ki}^{g} = \alpha_k S_k^g + \beta_k + \varepsilon_g + \varepsilon_g^i \tag{1}$$

where genes are indexed by g, treatments k, replicated/repeated measures i; S_k^g and X_{ki}^g denote the underlying and the i-th observed expression level of gene g under condition k, respectively; the systematic parameters α_k and β_k are gene-independent and account for the various differences in intensity measures between experiments. For example, α_k is mainly for modeling amplification efficiency in real time PCR experiments and β_k for assay-to-assay baseline drift. The ε_g accounts for variability between genes and ε_g^i is measurement error between the individual experiments within each gene g.

The underlying gene expression of *beta*-actin $S_k^{g^b}$, commonly used as the reference gene for the entire experiment, is considered biologically unchanged across assays and is thus denoted as S^{g^b}. For gene g and beta-actin g^b, we have

$$\begin{cases} X_{kj}^{g} = \alpha_k S_k^g + \beta_k + \varepsilon_g + \varepsilon_g^i & \text{(2a)} \\ X_{kj}^{g^b} = \alpha_k S_k^{g^b} + \beta_k + \varepsilon_{g^b} + \varepsilon_{g^b}^i & \text{(2b)} \\ S_k^{g^b} = S^{g^b} \end{cases} \tag{2}$$

Taking the ratio $\frac{(2a)-(2b)}{(2b)-(\beta_k+\varepsilon_{g^b})}$, we can have

$$\begin{aligned} \frac{X_{kj}^{g} - X_{kj}^{g^b}}{X_{kj}^{g^b} - (\beta_k + \varepsilon_{g^b})} &= \frac{\alpha_k\left(S_k^g - S^{g^b}\right) + (\varepsilon_g - \varepsilon_{g^b}) + \left(\varepsilon_g^i - \varepsilon_{g^b}^i\right)}{\alpha_k S^{g^b} + \varepsilon_{g^b}^i} \\ &= \left(\frac{S_k^g}{S^{g^b}} - 1\right) + \frac{(\varepsilon_g - \varepsilon_{g^b}) + \left(\varepsilon_g^i - \varepsilon_{g^b}^i\right)}{\alpha_k S^{g^b}} + o(\varepsilon) \end{aligned} \tag{3}$$

where $o(\varepsilon)$ represents a term at the order of $\varepsilon_{g^b}^i$, a small quantity. Multiplying (3) by $\left[X_{ki}^{g^b} - (\beta_k + \varepsilon_{g^b})\right]$ and denote $\frac{S_k^g}{S^{g^b}}$ by R_k^g, we can have

$$\begin{aligned} X_{kj}^{g} &= \frac{S_k^g}{S^{g^b}}\left[X_{kj}^{g^b} - (\beta_k + \varepsilon_{g^b})\right] + \left(\beta_k + \varepsilon^{g^b}\right) \\ &\quad + \frac{(\varepsilon_g - \varepsilon_{g^b}) + \left(\varepsilon_g^i - \varepsilon_{g^b}^i\right)}{\alpha_k S^{g^b}}\left[X_{kj}^{g^b} - (\beta_k + \varepsilon_{g^b})\right] \\ &= R_k^g X_{ki}^{g^b} + (1 - R_k^g)(\beta_k + \varepsilon_{g^b}) \\ &\quad + \frac{(\varepsilon_g - \varepsilon_{g^b}) + \left(\varepsilon_g^i - \varepsilon_{g^b}^i\right)}{\alpha_k S^{g^b}}\left(\alpha_k S^{g^b} + \varepsilon_{g^b}^i\right) \\ &= R_k^g X_{ki}^{g^b} + (1 - R_k^g)(\beta_k + \varepsilon_{g^b}) \\ &\quad + \left[(\varepsilon_g - \varepsilon_{g^b}) + \left(\varepsilon_g^i - \varepsilon_{g^b}^i\right)\right]\left(1 + \frac{\varepsilon_{g^b}^i}{\alpha_k S^{g^b}}\right) \\ &= R_k^g X_{ki}^{g^b} + (1 - R_k^g)(\beta_k + \varepsilon_{g^b}) + o(\varepsilon). \end{aligned} \tag{4}$$

Thus for replicate measures, the observed expression of gene g is a linear function of that of the reference gene g^b with a constant slope R_k^g for intersection $(1 - R_k^g)\beta_k$, plus a residue term $(1 - R_k^g)\varepsilon_{g^b} + o(\varepsilon)$.

Normalize the real time PCR data for each sample based on Eq. 4. We define the normalized relative expression level as

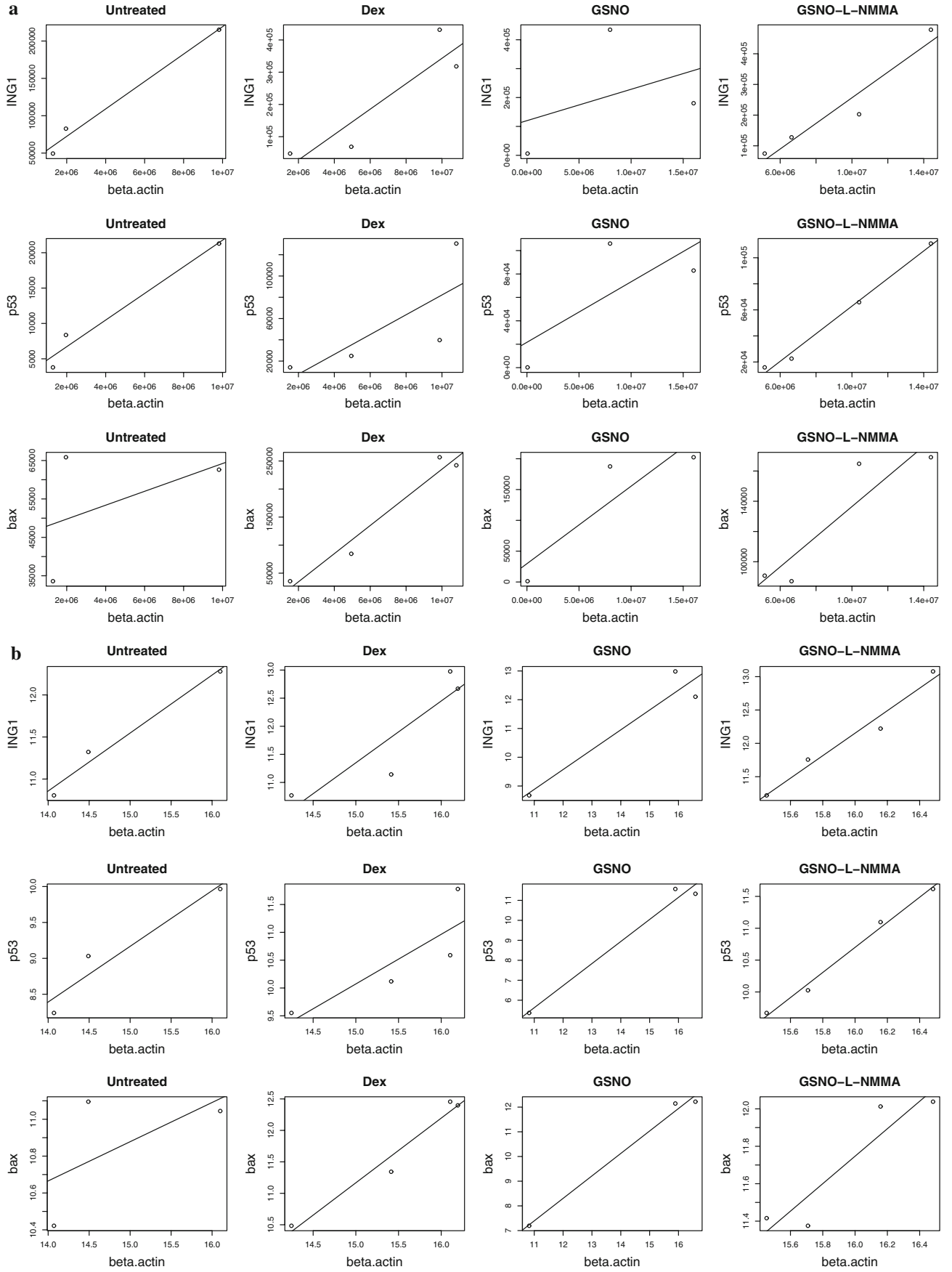
$$(R^g)_{ki} \equiv \frac{X_{ki}^{g} - (1 - R_k^g)\beta_k}{X_{ki}^{g^b}} \tag{5}$$

where $(1 - R_k^g)\beta_k$ can be easily estimated as the intercept of a linear regression model based on replicates $\left(X_{ki}^{g^b}, X_{ki}^{g}\right)$ under the same treatment k.

Our normalization algorithm potentially highlights the following advantages: (1) the statistical linear model in our normalization algorithm was based on incorporating a baseline drift term $\beta_k + \varepsilon_g$, which is not negligible in real world data as shown in Fig. 2; (2) the normalization algorithm only requires a small number of replicate/repeated measures (usually 3–4 are sufficient); (3) the normalized relative gene expression $(R^g)_{ki}$ does not depend on the term α_k, hence the effects of both amplification efficiency and baseline drift are eliminated through our normalization algorithm.

Pre-processing and normalization of microarray data

To perform cross-experiment and cross-laboratory data fusion and meta-analysis, we took the following pre-processing and normalization procedures that are crucial to ensure the data from various sources are transformed so that all the data sets share a common scale and baseline. A normalization procedure similar to the above-mentioned pre-processing step for real time PCR data was taken with the following slight modification: Instead of using repeated experimental data to estimate $(1 - R_k^g)\beta_k$ in formula (5),


a
Untreated
Dex
GSNO
GSNO–L–NMMA
ING1
p53
bax
beta.actin
b
Untreated
Dex
GSNO
GSNO–L–NMMA
ING1
p53
bax
beta.actin

◀ **Fig. 1** Plots of replicated copy numbers of genes of interest versus those of gene beta-actin under different treatments. **a** Raw data from real time PCR; **b** Logarithm transformed data. The line in each plot is the linear regression line of the data in the plot

quantile normalization method [22] ("normalize.quantile" function of "affy" package in Bioconductor (version 2.1)) was applied to the logarithm transformed microarray data to remove the baseline bias term. Thus we can obtain the normalized relative expression level simply as

$$(R^g)_k \equiv \frac{X_k^g}{X_k^{g^b}} \tag{6}$$

where X_k^g and $X_k^{g^b}$ are logarithm transformed and quantile normalized values of gene g and beta-action in sample k.

To further align the microarray data to real time PCR data so that the microarray data have the same gene-specific baseline as the real time PCR data, we performed a cross-platform normalization by adjusting all four microarray data sets by the mean difference between the untreated thymocyte samples in 4 microarray data and the PCR data.

Statistical analysis

The neural network model was fitted with the function "nnet" in "MASS" package in R (version 2.6.1) with a constraint of positive output. To avoid overfitting due to the limited samples, we used a randomization technique in the learning step to obtain 50 networks with random initial values. The reported final neural network was obtained by taking the mean of these 50 networks to achieve a reliable and robust model. The two linear models were implemented by the "lm" function in R. Source code of our procedure is available upon request.

Results

Inducing thymocyte apoptosis by GSNO and Dexamethasone (Dex)

We conducted various (Dex-treated, GSNO-treated, and GSNO-L-NMMA-co-treated) treatments on mouse

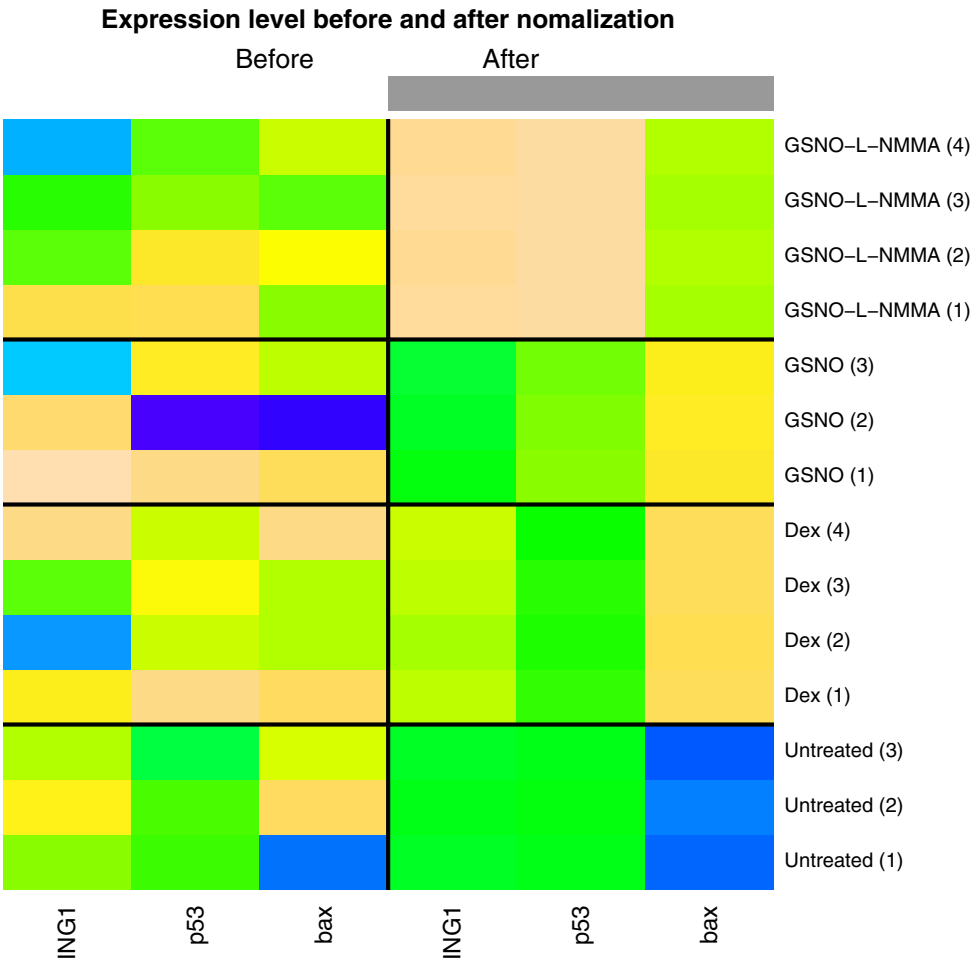


Fig. 2 Gene expression levels before and after normalization. Before: gene expression = log (copy number of interested gene)/log (copy number of beta-actin). After: normalized gene expressions by our algorithm show consistent levels between replicates of each gene within treatment

thymocyte to induce apoptosis [23]. Levels of thymocyte apoptosis (percentage) detected by a MoFlo cell sorter were untreated (2.6%), Dex-treated (16.1%), GSNO-treated (8.6%) and GSNO-L-NMMA-co-treated (7.8%). Treatment with GSNO or GSNO-L-NMMA markedly increased the apoptosis level with the latter being slightly inhibited by L-NMMA. Meanwhile, Dex, as a positive control [24], showed the strongest activity in inducing apoptosis (Fig. 3a).

The expression of p53, ING1 and bax were measured by real time PCR to the same thymocyte samples in parallel (Fig. 3b). The expression levels of bax were in accordance with the percentages of apoptosis: the higher the expression levels of bax, the more progressive the apoptosis (Fig. 3). Hence, the expression level of bax can be considered as an indicator of the level of both Dex-induced and GSNO-induced mouse thymocyte apoptosis.

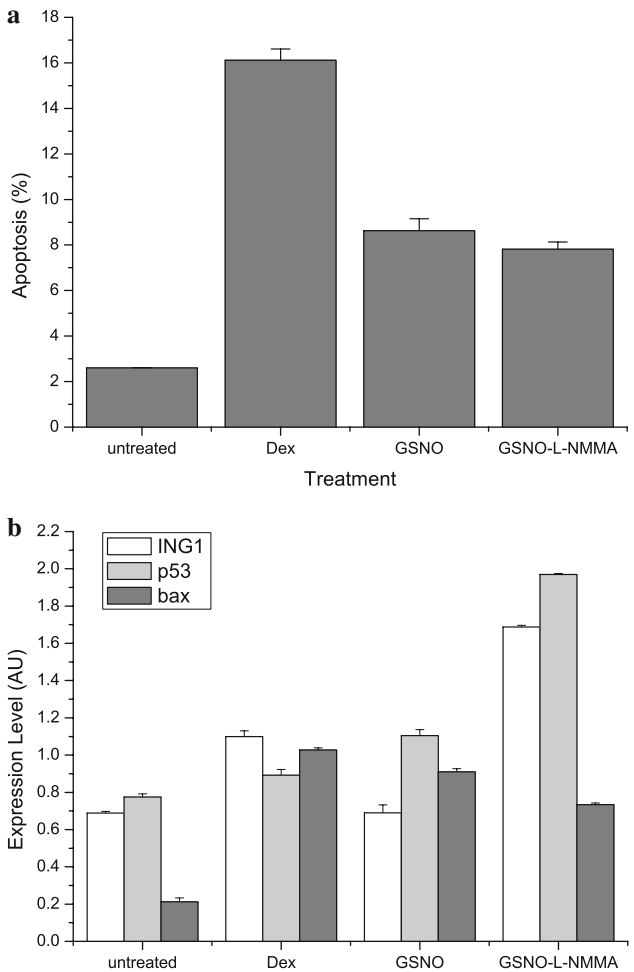


Fig. 3 Apoptosis and gene expression. **a** Apoptosis percentage (mean ± SD) in different treatments measured by MoFlo cell Sorter. Results are representative of three independent experiments performed in triplicate. Error bars represent one SD; **b** Expression levels of p53, ING1 and bax under untreatment, Dex-treatment, GSNO-treatment and GSNO-L-NMMA-treatment

Production of protein *S*-nirosylation in thymocyte

S-Nitrosylated proteins in the GSNO-treated thymocytes in absent/present L-NMMA treated were detected by biotin-switch and western blotting (Fig. 4). At least five *S*-nitrosylated protein bands are apparently seen in the apoptosis thymocytes induced by GSNO. There is significant difference of protein *S*-nitrosylation between the untreated and the GSNO-treated thymocytes. The protein *S*-nitrosylation was significantly inhibited by the NOS inhibitor L-NMMA, which is consistent with the protection of L-NMMA from apoptosis. We also found that proteins of moderate molecular weight (20–60 kDa) are more sensitive to GSNO and are strongly *S*-nitrosylated. NO has long been known to mediate cell death, and *S*-nitrosylation conveys a large part of the influence of NO on cellular signal transduction by affecting the activity, the conformation and the interaction of proteins.

A novel algorithm for normalizing the raw real time PCR data

Few methods were developed to normalize absolute quantification data of real time PCR. We propose a novel method here to normalize the raw data from real time PCR (absolute quantification data). It can efficiently remove the systematic variances and the baseline drafts by real time PCR, and yields accurate relative expression levels with small variance among replicated measurements.

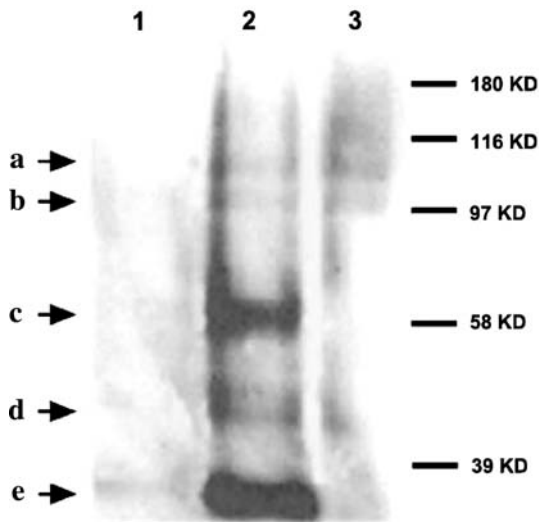


Fig. 4 Analysis of *S*-nitrosylated proteins in thymocytes by biotin-switch and western blotting method. *S*-Nitrosylated proteins are labeled with arrows on the left and the molecular markers are shown on the right. *Lane 1*, *S*-nitrosylated proteins in thymocytes without GSNO treatment; *Lane 2*, *S*-nitrosylated proteins in thymocytes treated with GSNO; *Lane 3*, *S*-nitrosylated proteins in thymocytes treated with L-NMMA and GSNO. The results are representative of three individual experiments

We obtained the copy numbers of genes p53, ING1 and bax, from real time PCR assays with 3–4 separated experiments. Figure 1a shows the copy numbers of genes of interest under the various treatments versus the copy numbers of the housekeeping gene beta-actin. Large variability between replicates was shown in most plots, which is due to the following effects: (1) nonlinear effects of the amplification process and fluorescence detection; and (2) varying experimental conditions, such as fluorescence detection and baseline drift in replicated measures. While the former can be taken care of with improved modeling, the latter cannot be controlled even with the use of endogenous genes, and thus needs to be adjusted computationally. We thus propose a novel normalization method which takes into account the two effects and can also be applied to microarray data (see "Materials and methods" for details).

Gene expressions between separated experiments obtained by our normalization method have less variance comparing those by the conventional method (Fig. 2). This demonstrates that a suitable normalization method for real time PCR data is crucial in effective data analysis and interpretation, and expression levels of p53, ING1, bax obtained by our algorithm are more reliable for further analysis of the intricate gene interactions.

p53 and ING1 may cooperate to regulate bax expression

Many studies have suggested that bax is a p53 primary-response gene, involved in a p53-up-regulated pathway for the induction of apoptosis [4, 5]. The expression of p53 and bax increased the following all three apoptosis-inducing treatments (Fig. 3b). However, p53 and ING1 increased the most following the treatment with GSNO-L-NMMA, while bax increased the least. Furthermore, p53 and bax show few correlations in data set from our real time PCR (Table 1). These findings seem to contradict the biological results. The problem lies in that a nonlinear relationship was characterized in a linear form (correlation coefficient here), which leads to misleading results. All the above experimental results suggest that bax is not regulated linearly by p53 or by ING1 alone, but rather through cooperative interactions among the gene products in a nonlinear network.

Table 1 Pairwise Pearson correlation coefficients of gene pairs

Gene pairs	Pearson correlation coefficient	*P* value
p53 and ING1	0.871	<0.0001
Gene expression	0.123	0.6748
ING1 and bax	0.248	0.3927

Nonlinear network model for p53 and ING1 regulating bax

Recent studies employing neural network-like supervised learning methods have successfully predicted the clinical outcome of breast cancer [25] and multiple time-dependent apoptotic responses [26]. These models are capable of modeling several nonlinear relationships between the response variable and the independent variables simultaneously through a set of flexible non-parametric statistical models. These models suggest promising approaches to explore molecular mechanism on thymocyte apoptosis which is obviously nonlinear.

We modeled the regulatory network of these three genes with a feed-forward neural network model: the expression levels of the genes p53 and ING1 (independent variables) as the inputs and the expression levels of bax as the output (response variable). The feed-forward neural network has a hidden layer of two units and was trained by the normalized relative measurements from real time PCR data. The response surface for bax was obtained with our neural network model and demonstrates a nonlinear relationship between p53 and ING1 in regulating bax cooperatively (Fig. 5).

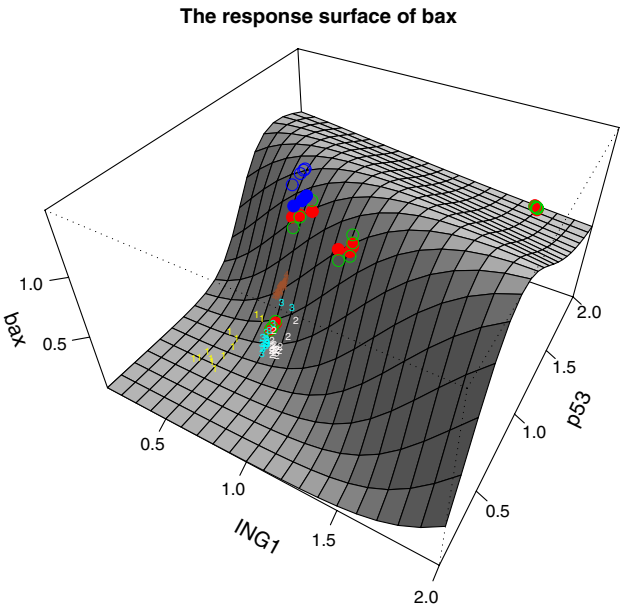


Fig. 5 Response surface for Bax. Response surface for bax (*z*-axis) by the inputs of p53 and ING1: *Red dots*, training data from real time PCR assays; *Green circles*, the predicted bax for real time PCR assays by the neural network model; *Blue dots*, testing data from real time PCR assays (GSNO-SRS treatment); *Blue circles*, the predicted bax for real time PCR assays of GSNO-SRS treatment by the neural network model. Colored symbols present predicted levels for 4 microarray data by neural network model. Yellow *1* data set *TS* White *2* data set *TN* Cyan *3* data set *IA* Brown *4* data set *NU*

Validations by cross-platform and cross-experimental data

To avoid overfitting and to validate the neural network response surface model constructed above, an independent data from real time PCR and four completely independent publicly available microarray data sets were used for testing.

Sinapine was discovered for many years [27] and was intensively studied in plants [28]. Sinapic acid shows anxiolytic-like effects in mice [29]. To explore the role of Sinapine in apoptosis, we measured the expression levels of the 3 genes (p53, ING1 and bax) in mouse thymocytes which were co-treated by GSNO and Sinapine of Raphani sativus semen (SRS) with 4 separated experiments and processed by the same procedures as our proposed method. This data were used as our testing data.

Meanwhile, four publicly available microarray data sets obtained from thymocyte and/or T cell studies were used also. The four data sets were: thymocyte selection by agonist (TS) [16], thymocyte negative selection (TN) [17], innate vs. adaptive lymphocyte gene expression (IA) [18], and neuromedin U effect on type-2 Th cells (NU) [19]. They were preprocessed as described in "Materials and methods". It is shown that the expression level of bax predicted by the response surface of neural network model approximated very well the experimental data (Figs. 5, 6).

Furthermore, as comparisons, a linear model with p53 as input and a linear model with p53 and ING1 together as inputs were also constructed to predict the expression levels of bax. These two linear models were also trained with our real time PCR data. It is shown that the prediction by the response surface of neural network model is much better (Fig. 6).

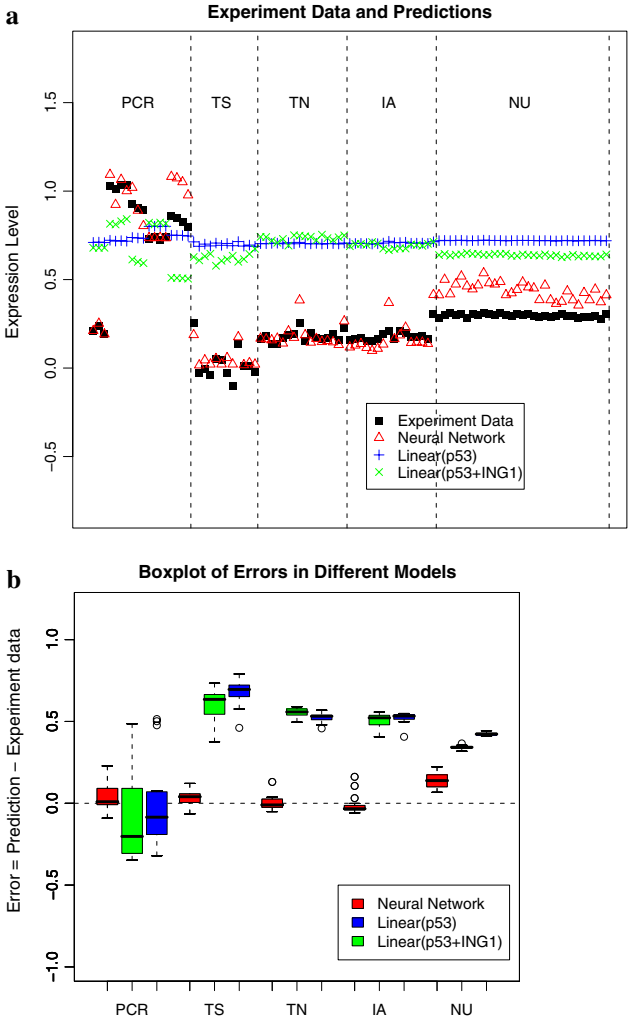


Fig. 6 Validation of models. **a** The scatter plot of experiment measured values of bax for different data sources and their corresponding predictions by neural network model and the two linear models. The right most 4 samples in the PCR data region were sample of GSNO–SRS treated mouse thymocytes; **b** Boxplot of prediction errors (= predicted value-experiment measured value) of three models

Discussion

In this study, we explored the quantitative relationship among the expression of three vital genes during mouse thymocyte apoptosis. Both GSNO-induced and Dex-induced apoptosis were investigated. The expression level of bax was proportional to the rate of apoptosis under various GSNO and Dex treatments. In examining the relation between bax and its upstream genes p53 and ING1, we found that bax is not regulated linearly on a one-on-one basis, but rather by a gene network through nonlinear cooperative interactions.

We employed a neural network response surface model to characterize the nonlinearity present in the cooperative interactions among the genes p53, ING1 and bax. This kind of nonlinearity, however, has not been revealed so far in most apoptosis studies utilizing microarray and real time PCR techniques. One might doubt the nonlinearity is only from overfitting in learning of neural networks. But the results on independent validation data showed that this is not the case. Typically, individual studies have only a limited range of gene expression under strictly controlled biological conditions, and their data are restricted in small ranges. Therefore linear approximation would be good enough, and thus pair-wise linear relationships within various pathways are often examined. When we integrate data from various cases and laboratories, the nonlinearity would appear. In addition, lack of suitable data preprocessing methods and large variability in cross-experimental, cross-laboratory, and cross-platform studies of gene expression data, often led to incomplete and even

contradictory interpretations of the gene expression data. Nonlinearities may only be detected through a large domain of gene expressions that cover different studies. It thus can be concluded that pair-wise gene interaction studies through correlation coefficients are only adequate locally in a small range of gene expressions. A rather sophisticated network model is needed to study the intricate and delicate relationships among genes for a thorough understanding of the p53 pathway and apoptosis. Moreover, the agreement of our predicted gene expression level with the microarray experimental data also indicates that employment of nonlinear models seems to be not just plausible but necessary.

It is worthwhile to note that the reliability of microarray technology is still in debate [30]. Major concerns were raised with inconsistent conclusions drawn from cross-platform experiments [15]. In fact, such inconsistency may not be solely due to the technology itself but is in part due to the normalization methods and analysis algorithms employed. Large variability may still exist and may lead to reliability concerns after the standard normalization methods. For example, the Pearson correlation coefficients between bax and p53 vary largely across these 4 data sets (0.698 in TS, 0.349 in TN, 0.741 in IA, and 0.471 in NU). However, after applying our normalization method of taking the ratios of each gene to beta-actin, the Pearson correlation coefficients lie at a comparable level: 0.775, 0.509, 0.747, and 0.503, respectively.

Furthermore, it is noted that different probes of the same target sequences (genes) with the same copy numbers may yield different intensities, because the binding affinities of probes in microarray are non-negligibly different and depend on probe sequences [31]. Although the relative expression levels have been used for genes (p53, ING1 and bax) in microarray data, they still differ from our normalized real time PCR data by a gene-specific baseline. Thus, a cross-platform normalization procedure is required to achieve the consistency between microarray gene expression data and the real time PCR gene expression data.

Although correlation has been used to integrate cross-platform microarray data [32], it was not suitable to our analysis here for quantitative study. We thus propose a method to eliminate cross-platform differences, a cross-platform normalization procedure, see "Materials and methods". This cross-platform normalization method yielded remarkable consistency of the predicted bax level by the neural network response surface model using both real time PCR and microarray experimental data. This indicates that gene expression levels from different platforms can be consistent if they are properly preprocessed with reference genes. Our results exemplify the feasibility of integrating cross-laboratory, cross-experimental and cross-platform gene expression data.

In summary, we exemplified an integrated approach to study nonlinear relationships of 3 genes in p53-dependent networks. We indicate that for complex interdependent regulatory networks, pair-wise gene interaction studies through correlation coefficients are often only adequate locally in a small range of gene expressions. A rather sophisticated network model is needed to study the intricate and delicate relationship among genes for a thorough understanding of the gene regulatory networks. Our results also demonstrate that gene expression levels from different platforms, such as real time PCR and microarrays, can be consistent if they are properly preprocessed with appropriate reference genes.

Acknowledgments This work is supported by the National Natural Science Foundation of China (No. 30570425, No. 10721403, No. 39770202), the National High Technology Research and Development of China (No. 2006AA02Z331, No. 2008AA02Z306), the National Key Basic Research Project of China (No. 2003CB715903, No. 2006CB911001), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry. This work was completed during the visiting of authors (LW and MQ) to Michigan State University. We also acknowledge Lizhen Gu for providing Sinapine, and Xiaoyou Su, Ruihua Liu, Shuang Cao and Lijin Shi for technical assistance.

References

1. Bogdan C (2001) Nitric oxide and the immune response. Nat Immunol 2:907–916. doi:10.1038/ni1001-907
2. Vogelstein B, Lane D, Levine AJ (2000) Surfing the p53 network. Nature 408:307–310. doi:10.1038/35042675
3. Xu Y (2003) Regulation of p53 responses by post-translational modifications. Cell Death Differ 10:400–403. doi:10.1038/sj.cdd.4401182
4. Chipuk JE, Green DR (2004) Cytoplasmic p53: bax and forward. Cell Cycle 3:429–431
5. Miyashita T, Reed JC (1995) Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. Cell 80:293–299. doi:10.1016/0092-8674(95)90513-8
6. Yin C, Knudson CM, Korsmeyer SJ, Van Dyke T (1997) Bax suppresses tumorigenesis and stimulates apoptosis in vivo. Nature 385:637–640. doi:10.1038/385637a0
7. Garkavtsev I, Kazarov A, Gudkov A, Riabowol K (1996) Suppression of the novel growth inhibitor p33ING1 promotes neoplastic transformation. Nat Genet 14:415–420. doi:10.1038/ng1296-415
8. Helbing CC, Veillette C, Riabowol K, Johnston RN, Garkavtsev I (1997) A novel candidate tumor suppressor, ING1, is involved in the regulation of apoptosis. Cancer Res 57:1255–1258
9. Garkavtsev I, Grigorian IA, Ossovskaya VS, Chernov MV, Chumakov PM, Gudkov AV (1998) The candidate tumour suppressor p33ING1 cooperates with p53 in cell growth control. Nature 391:295–298. doi:10.1038/34675
10. Gonzalez L, Freije JM, Cal S, Lopez-Otin C, Serrano M, Palmero I (2006) A functional link between the tumour suppressors ARF and p33ING1. Oncogene 25:5173–5179
11. Cheung KJ Jr, Li G (2002) p33(ING1) enhances UVB-induced apoptosis in melanoma cells. Exp Cell Res 279:291–298. doi:10.1006/excr.2002.5610

12. Zhu JJ, Li FB, Zhou JM, Liu ZC, Zhu XF, Liao WM (2005) The tumor suppressor p33ING1b enhances taxol-induced apoptosis by p53-dependent pathway in human osteosarcoma U2OS cells. Cancer Biol Ther 4:39–47. doi:10.1158/1535-7163.MCT-04-0330
13. Levine AJ, Hu W, Feng Z (2006) The P53 pathway: what questions remain to be explored? Cell Death Differ 13:1027–1036. doi:10.1038/sj.cdd.4401910
14. Barczak A, Rodriguez MW, Hanspers K et al (2003) Spotted long oligonucleotide arrays for human gene expression analysis. Genome Res 13:1775–1785. doi:10.1101/gr.1048803
15. Tan PK, Downey TJ, Spitznagel EL Jr et al (2003) Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Res 31:5676–5684. doi:10.1093/nar/gkg763
16. Yamagata T, Mathis D, Benoist C (2004) Self-reactivity in thymic double-positive cells commits cells to a CD8 alpha alpha lineage with characteristics of innate immune cells. Nat Immunol 5:597–605. doi:10.1038/ni1070
17. Zucchelli S, Holler P, Yamagata T, Roy M, Benoist C, Mathis D (2005) Defective central tolerance induction in NOD mice: genomics and genetics. Immunity 22:385–396. doi:10.1016/j.immuni.2005.01.015
18. Yamagata T, Benoist C, Mathis D (2006) A shared gene-expression signature in innate-like lymphocytes. Immunol Rev 210:52–66. doi:10.1111/j.0105-2896.2006.00371.x
19. Johnson EN, Appelbaum ER, Carpenter DC et al (2004) Neuromedin U elicits cytokine release in murine Th2-type T cell clone D10.G4.1. J Immunol 173:7230–7238
20. Speed TP (2003) Statistical analysis of gene expression microarray data. Chapman & Hall/CRC, Boca Raton
21. Tsodikov A, Szabo A, Jones D (2002) Adjustments and measures of differential expression for microarray data. Bioinformatics 18:251–260. doi:10.1093/bioinformatics/18.2.251
22. Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19:185–193. doi:10.1093/bioinformatics/19.2.185
23. Lin DY, Ma WY, Duan SJ, Zhang Y, Du LY (2006) Real-time imaging of viable-apoptotic switch in GSNO-induced mouse thymocyte apoptosis. Apoptosis 11:1289–1298. doi:10.1007/s10495-006-7804-1
24. Nicoletti I, Migliorati G, Pagliacci MC, Grignani F, Riccardi C (1991) A rapid and simple method for measuring thymocyte apoptosis by propidium iodide staining and flow cytometry. J Immunol Methods 139:271–279. doi:10.1016/0022-1759(91)90198-O
25. van't Veer LJ, Dai H, van de Vijver MJ et al (2002) Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530–536. doi:10.1038/415530a
26. Janes KA, Albeck JG, Gaudet S, Sorger PK, Lauffenburger DA, Yaffe MB (2005) A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. Science 310:1646–1653. doi:10.1126/science.1116598
27. Tzagoloff A (1963) Metabolism of sinapine in mustard plants I. Degradation of sinapine into sinapic acid & choline. Plant Physiol 38:202–206
28. Weier D, Mittasch J, Strack D, Milkowski C (2008) The genes BnSCT1 and BnSCT2 from Brassica napus encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression. Planta 227:375–385. doi:10.1007/s00425-007-0624-x
29. Yoon BH, Jung JW, Lee JJ et al (2007) Anxiolytic-like effects of sinapic acid in mice. Life Sci 81:234–240. doi:10.1016/j.lfs.2007.05.007
30. Eisenstein M (2006) Microarrays: quality control. Nature 442:1067–1070. doi:10.1038/4421067a
31. Zhang L, Miles MF, Aldape KD (2003) A model of molecular interactions on short oligonucleotide microarrays. Nat Biotechnol 21:818–821. doi:10.1038/nbt836
32. Zhou XJ, Kao MC, Huang H et al (2005) Functional annotation and network reconstruction through cross-platform integration of microarray data. Nat Biotechnol 23:238–243. doi:10.1038/nbt1058