Electronic Supplementary Material (ESI) for Molecular BioSystems.

Bayesian Supplementary Material

Naruemon Pratanwanich and Pietro Lio'

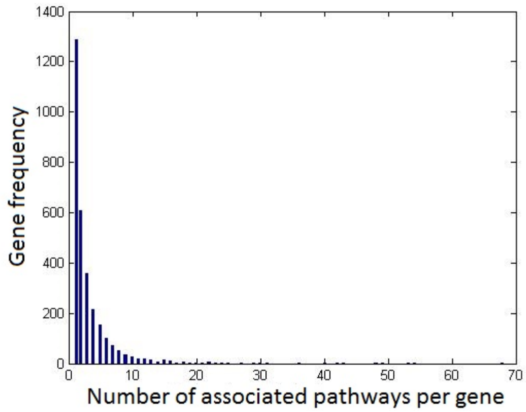
1 Data

Pre-processing for CMap data

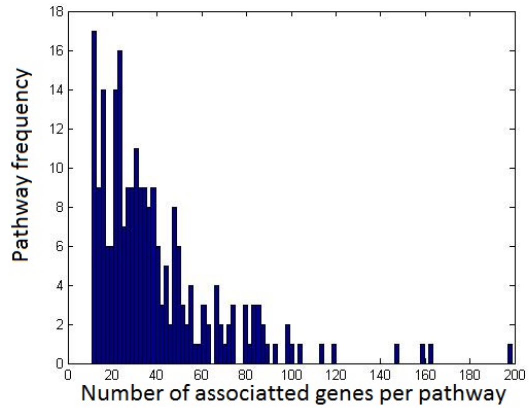
We followed the standard microarray data analysis as follows. Firstly, we performed quality assessment of each individual array downloaded from CMap database in the format of raw .CELL files of the human breast epithelial adenocarcinoma cell line derived from pleural effusion (MCF7) and the epithelial cell line established from human prostate adenocarcinoma (PC3) from the microarray gene chip HT_HG_U133A. The arrays with bad quality according to the NUSE (normalised Unscaled Standard Error) and RLE (Relative Log Expression) were filtered out. Next, we normalised across the remaining microarrays and then calculated the differential gene expression in terms of log fold change (the logarithm of the ratio between treatment and control). Finally, only probes mapped to genes with one-to-one relationship were retained for further analysis. To summarise, we used the following R packages in the preprocessing steps: *affyPLM* package for quality control, *gcrma* function from *affy* package for multiple arrays normalization, and *limma* package for the calculation of differential gene expression. As a result, the matrix of differential gene expression was created, where rows and columns correspond to gene identifiers and conditions (drugs), respectively.

Data from KEGG database

Since the REST-style version of KEGG API allows us to collect the information and execute some database operations via the internet browser, we downloaded the gene-pathway association with the operation called "link" provided in the API.



(a) Histogram of gene



(b) Histogram of pathway

Figure 1. The distribution of genes over the number of associated pathways (Figure 1(a)) and the distribution of pathways over the associated genes (Figure 1(b)) in matrix $\mathbf{K}$.

2 Model distributions

$$
\begin{aligned}
P(\mathbf{X}|\mathbf{B},\mathbf{S},\tau_\epsilon) &= \prod_{g=1}^{G}\prod_{d=1}^{D}\mathcal{N}(x_{gd}|\mathbf{b}_g\mathbf{s}_d,\tau_\epsilon^{-1}) \\
P(\tau_\epsilon|\alpha_\epsilon,\beta_\epsilon) &= \Gamma(\tau_\epsilon|\alpha_\epsilon,\beta_\epsilon) \\
P(\mathbf{B}|\mathbf{K}) &= \prod_{(g,p)\in\mathcal{Z}}\mathcal{N}(b_{gp}|0,\tau_B^{-1}) \\
P(\tau_B|\alpha_B,\beta_B) &= \Gamma(\tau_B|\alpha_B,\beta_B) \\
P(\mathbf{S}|\mathbf{\Phi}) &= \prod_{d=1}^{D}\mathcal{N}(\mathbf{s}_d|\mathbf{0},\mathbf{\Phi}^{-1}) \\
P(\mathbf{\Phi}|\nu,\mathbf{\Psi}) &= \mathcal{W}(\mathbf{\Phi}|\nu,\mathbf{\Psi})
\end{aligned}
$$

where $\mathcal{Z} = \{(g,p) \in G \times P | k_{gp} = 1)\}$,
$\mathcal{N}(\cdot), \Gamma(\cdot)$, and $\mathcal{W}(\cdot)$ denotes a Gaussian, Gamma, and Wishart distribution respectively.

3 Joint posterior distribution

Let $\mathbf{\Omega} = \{\alpha_E, \beta_E, \alpha_B, \beta_B, \nu, \mathbf{\Psi}\}$ are predefined hyperparameters. The joint posterior distribution for the parameters $\{\tau_\epsilon, \tau_B, \mathbf{B}, \mathbf{S}, \mathbf{\Phi}\}$ is proportional to the product of the likelihood and the prior distributions as follows.

$$
\begin{aligned}
P(\tau_\epsilon,\tau_B,\mathbf{B},\mathbf{S},\mathbf{\Phi}|\mathbf{X},\mathbf{K},\mathbf{\Omega}) \propto\, & P(\mathbf{X}|\mathbf{B},\mathbf{S},\tau_\epsilon)P(\tau_\epsilon|\alpha_E,\beta_E) \\
& P(\mathbf{B}|\mathbf{K},\tau_B)P(\tau_B|\alpha_B,\beta_B) \\
& P(\mathbf{S}|\mathbf{\Phi})P(\mathbf{\Phi}|\nu,\mathbf{\Psi})
\end{aligned}
$$

4 Conditional posterior distributions for Gibbs sampling

1. $\tau_\epsilon \sim P(\tau_\epsilon|\mathbf{X},\mathbf{B},\mathbf{S},\alpha_\epsilon,\beta_\epsilon)$

$$\begin{aligned}
P(\tau_\epsilon|\mathbf{X},\mathbf{B},\mathbf{S},\alpha_\epsilon,\beta_\epsilon) &\propto \prod_{g=1}^{G}\prod_{d=1}^{D}\mathcal{N}(x_{gd}|\mathbf{b}_g\mathbf{s}_d,\tau_\epsilon^{-1})\Gamma(\tau_\epsilon|\alpha_\epsilon,\beta_\epsilon)\\
&\propto \prod_{g=1}^{G}\prod_{d=1}^{D}\tau_\epsilon^{\frac{1}{2}}\exp(-\frac{\tau_\epsilon}{2}(x_{gd}-\mathbf{b}_g\mathbf{s}_d)^2)\cdot\tau_\epsilon^{\alpha_\epsilon-1}\exp(-\beta_\epsilon\tau_\epsilon)\\
&\propto \tau_\epsilon^{\frac{GD}{2}}\exp(-\frac{\tau_\epsilon}{2}\sum_{g=1}^{G}\sum_{d=1}^{D}(x_{gd}-\mathbf{b}_g\mathbf{s}_d)^2)\cdot\tau_\epsilon^{\alpha_\epsilon-1}\exp(-\beta_\epsilon\tau_\epsilon)\\
&\propto \tau_\epsilon^{\alpha_\epsilon+\frac{GD}{2}-1}\exp(-\tau_\epsilon(\beta_\epsilon+\frac{1}{2}\sum_{g=1}^{G}\sum_{d=1}^{D}(x_{gd}-\mathbf{b}_g\mathbf{s}_d)^2))\\
&\propto \Gamma(\alpha_\epsilon^*,\beta_\epsilon^*)
\end{aligned}$$

;where $\alpha_\epsilon^* = \alpha_\epsilon + \frac{GD}{2}$ and $\beta_\epsilon^* = \beta_\epsilon + \frac{1}{2}\sum_{g=1}^{G}\sum_{d=1}^{D}(x_{gd}-\mathbf{b}_g\mathbf{s}_d)^2$.

2. $\tau_B \sim P(\tau_B|\mathbf{B},\alpha_B,\beta_B)$

$$\begin{aligned}
P(\tau_B|\mathbf{B},\alpha_B,\beta_B) &\propto \prod_{(g,p)\in\mathcal{Z}}\mathcal{N}(b_{gp}|0,\tau_B^{-1})\cdot\Gamma(\tau_B|\alpha_B,\beta_B)\\
&\propto \prod_{(g,p)\in\mathcal{Z}}\tau_B^{\frac{1}{2}}\exp(-\frac{\tau_B}{2}b_{gp}^2)\cdot\tau_B^{\alpha_B-1}\exp(-\beta_B\tau_B)\\
&\propto \tau_B^{\frac{|\mathcal{Z}|}{2}}\exp(-\frac{\tau_B}{2}\sum_{(g,p)\in\mathcal{Z}}b_{gp}^2)\cdot\tau_B^{\alpha_B-1}\exp(-\beta_B\tau_B)\\
&\propto \tau_B^{\alpha_B+\frac{|\mathcal{Z}|}{2}-1}\exp(-\tau_B(\beta_B+\frac{1}{2}\sum_{(g,p)\in\mathcal{Z}}b_{gp}^2))\\
&\propto \Gamma(\alpha_B^*,\beta_B^*)
\end{aligned}$$

;where
$\mathcal{Z} = \{(g,p)\in G\times P|k_{gp}=1)\}$
$\alpha_B^* = \alpha_B + \frac{|\mathcal{Z}|}{2}$ and $\beta_B^* = \beta_B + \frac{1}{2}\sum_{(g,p)\in\mathcal{Z}}b_{gp}^2$.

3. $\mathbf{b}_g \sim P(\mathbf{b}_g|\mathbf{X},\mathbf{S},\tau_\epsilon,\tau_B)$

$$P(\mathbf{b}_g|\mathbf{X},\mathbf{S},\tau_\epsilon,\tau_B) \propto \mathcal{N}(\mathbf{b}_g|\mu_B^*,(\mathbf{\Phi}_B^*)^{-1})\ \texttt{if}\ \ k_{gp}=1; p=1,2,3,...,P$$

where $\mu_B^* = (\mathbf{\Phi}_B^*)^{-1}(\tau_\epsilon\mathbf{S}^*\mathbf{x}_g^\top)$, $\mathbf{\Phi}_B^* = \tau_\epsilon\mathbf{S}^*(\mathbf{S}^*)^\top + \tau_B I_{|\mathbf{S}^*|}$ and S^* = the submatrix of matrix $\mathbf{S}$ with the row indices corresponding to the 1-entries of the vector $\mathbf{k}_g$.

4. $\mathbf{\Phi} \sim P(\mathbf{\Phi}|\mathbf{S}, \nu, \mathbf{\Psi})$

$$
\begin{aligned}
P(\mathbf{\Phi}|\mathbf{S}, \nu, \mathbf{\Psi}) &\propto \prod_{d=1}^{D} \mathcal{N}(\mathbf{s}_d|\mathbf{0}, \Phi^{-1}) \cdot \mathcal{W}(\mathbf{\Phi}|\nu, \mathbf{\Psi}) \\
&\propto |\mathbf{\Phi}|^{\frac{D}{2}} \exp(-\frac{1}{2}\sum_{d=1}^{D} \mathbf{s}_d^\top \mathbf{\Phi} \mathbf{s}_d) \cdot |\mathbf{\Phi}|^{\frac{\nu-P-1}{2}} \exp(-\frac{1}{2} trace(\mathbf{\Psi}^{-1}\mathbf{\Phi})) \\
&\propto |\mathbf{\Phi}|^{\frac{\nu+D-P-1}{2}} \exp(-\frac{1}{2}(trace(\sum_{d=1}^{D} \mathbf{s}_d^\top \mathbf{\Phi} \mathbf{s}_d + \mathbf{\Psi}^{-1}\mathbf{\Phi}))) \\
&\propto |\mathbf{\Phi}|^{\frac{\nu+D-P-1}{2}} \exp(-\frac{1}{2}(trace((\sum_{d=1}^{D} \mathbf{s}_d^\top \mathbf{s}_d + \mathbf{\Psi}^{-1})\mathbf{\Phi}))) \\
&\propto \mathcal{W}(\mathbf{\Phi}|\nu^*, \mathbf{\Psi}^*)
\end{aligned}
$$

;where $\nu^* = \nu + D$ and $\mathbf{\Psi}^* = (\mathbf{\Psi}^{-1} + \sum_{d=1}^{D}(\mathbf{s}_d\mathbf{s}_d^\top))^{-1}$

5. $\mathbf{s}_d \sim P(\mathbf{s}_d|\mathbf{X}, \mathbf{B}, \tau_\epsilon, \mathbf{\Phi})$

$$
\begin{aligned}
P(\mathbf{s}_d|\mathbf{X}, \mathbf{B}, \tau_\epsilon, \mathbf{\Phi}) &\propto \mathcal{N}(\mathbf{x}_d|\mathbf{B}\mathbf{s}_d, \tau_\epsilon^{-1}\mathbf{I}_G) \cdot \mathcal{N}(\mathbf{s}_d|\mathbf{0}, \mathbf{\Phi}^{-1}) \\
&\propto \exp(-\frac{1}{2}(\mathbf{x}_d - \mathbf{B}\mathbf{s}_d)^\top(\tau_\epsilon\mathbf{I}_G)(\mathbf{x}_d - \mathbf{B}\mathbf{s}_d)) \cdot \exp(-\frac{1}{2}\mathbf{s}_d^\top\mathbf{\Phi}\mathbf{s}_d) \\
&\propto \exp(-\frac{\tau_\epsilon}{2}(\mathbf{x}_d^\top - (\mathbf{B}\mathbf{s}_d)^\top)(\mathbf{x}_d - \mathbf{B}\mathbf{s}_d) - \frac{1}{2}\mathbf{s}_d^\top\mathbf{\Phi}\mathbf{s}_d) \\
&\propto \exp(-\frac{\tau_\epsilon}{2}(\mathbf{x}_d^\top - \mathbf{s}_d^\top\mathbf{B}^\top)(\mathbf{x}_d - \mathbf{B}\mathbf{s}_d) - \frac{1}{2}\mathbf{s}_d^\top\mathbf{\Phi}\mathbf{s}_d) \\
&\propto \exp(-\frac{\tau_\epsilon}{2}(\mathbf{x}_d^\top\mathbf{x}_d - \mathbf{x}_d^\top\mathbf{B}\mathbf{s}_d - \mathbf{s}_d^\top\mathbf{B}^\top\mathbf{x}_d + \mathbf{s}_d^\top\mathbf{B}^\top\mathbf{B}\mathbf{s}_d) - \frac{1}{2}\mathbf{s}_d^\top\mathbf{\Phi}\mathbf{s}_d) \\
&\propto \exp(\frac{\tau_\epsilon}{2}\mathbf{x}_d^\top\mathbf{B}\mathbf{s}_d + \frac{\tau_\epsilon}{2}\mathbf{s}_d^\top\mathbf{B}^\top\mathbf{x}_d - \frac{\tau_\epsilon}{2}\mathbf{s}_d^\top\mathbf{B}^\top\mathbf{B}\mathbf{s}_d - \frac{1}{2}\mathbf{s}_d^\top\mathbf{\Phi}\mathbf{s}_d) \\
&\propto \exp(\frac{\tau_\epsilon}{2}\mathbf{x}_d^\top\mathbf{B}\mathbf{s}_d + \frac{\tau_\epsilon}{2}\mathbf{s}_d^\top\mathbf{B}^\top\mathbf{x}_d - \frac{1}{2}\mathbf{s}_d^\top(\tau_\epsilon\mathbf{B}^\top\mathbf{B} + \mathbf{\Phi})\mathbf{s}_d) \\
&\propto \mathcal{N}(\mathbf{s}_d|\mu^*, (\mathbf{\Phi}^*)^{-1})
\end{aligned}
$$

;where $\mu^* = (\mathbf{\Phi}^*)^{-1}(\tau_\epsilon\mathbf{B}^\top\mathbf{x}_d)$ and $\mathbf{\Phi}^* = \tau_\epsilon\mathbf{B}^\top\mathbf{B} + \mathbf{\Phi}$

NOTE:

$$
\begin{aligned}
\mathcal{N}(\mathbf{x}|\mu, \Sigma^{-1}) &\propto \exp(-\frac{1}{2}(\mathbf{x} - \mu)^\top\Sigma(\mathbf{x} - \mu)) \\
&\propto \exp(-\frac{1}{2}(\mathbf{x}^\top - \mu^\top)\Sigma(\mathbf{x} - \mu)) \\
&\propto \exp(-\frac{1}{2}\mathbf{x}^\top\Sigma\mathbf{x} + \frac{1}{2}\mathbf{x}^\top\Sigma\mu + \frac{1}{2}\mu^\top\Sigma\mathbf{x} - \frac{1}{2}\mu^\top\Sigma\mu) \\
&\propto \exp(-\frac{1}{2}\mathbf{x}^\top\Sigma\mathbf{x} + \frac{1}{2}\mathbf{x}^\top\Sigma\mu + \frac{1}{2}\mu^\top\Sigma\mathbf{x})
\end{aligned}
$$