

# A Mixture Model with Random-Effects Components for Clustering Correlated Gene-Expression Profiles:

## Supplementary text

S.K. Ng<sup>a</sup>, G. J. McLachlan<sup>a,b,c,\*</sup>, K. Wang<sup>c</sup>, L. Ben-Tovim Jones<sup>b</sup>, S-W. Ng<sup>d</sup>

<sup>a</sup>Department of Mathematics, University of Queensland, Brisbane, QLD 4072, Australia, <sup>b</sup>Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD 4072, Australia, <sup>c</sup>ARC Centre for Complex Systems, University of Queensland, Brisbane, QLD 4072, Australia, <sup>d</sup>Laboratory of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Boston, MA 02115, USA.

### 1 THE E-STEP

In the EM framework for the mixture of linear mixed-effects models (LMMs) described in Section 2 of the main text, the log likelihood on the basis of the complete data  $(\mathbf{y}^T, \mathbf{z}^T, \mathbf{b}^T, \mathbf{c}^T)^T$  is given by  $\log f(\mathbf{z}) + \log f(\mathbf{c} | \mathbf{z}) + \log f(\mathbf{y}, \mathbf{b} | \mathbf{z}, \mathbf{c})$ , where the symbol  $f(\cdot)$  is being used generically to denote a density. With the last term in the complete-data log likelihood above, it is assumed that the joint vectors  $\mathbf{y}$  and  $\mathbf{b}$  are conditionally independent given  $\mathbf{z}$  and  $\mathbf{c}$ . Thus, we have

$$\log f(\mathbf{y}, \mathbf{b} | \mathbf{z}, \mathbf{c}) = \sum_{j=1}^n \sum_{h=1}^g z_{hj} \log f(\mathbf{y}_j, \mathbf{b}_{hj} | \mathbf{c}_h), \quad (1)$$

where  $f(\mathbf{y}_j, \mathbf{b}_{hj} | \mathbf{c}_h)$  is the multivariate normal density of dimension  $(m + q_b)$ , corresponding to the distribution,

$$N \left( \begin{bmatrix} \mathbf{X}\beta_h + \mathbf{V}\mathbf{c}_h \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{A}_h & \theta_{ch}\mathbf{V} \\ \theta_{ch}\mathbf{V}^T & \theta_{bh}\mathbf{I}_{q_b} \end{bmatrix} \right),$$

where  $\mathbf{A}_h = \text{diag}(\mathbf{W}\phi_h)$  is a diagonal matrix constructed from the vector  $(\mathbf{W}\phi_h)$  with  $\phi_h = (\sigma_{h1}^2, \dots, \sigma_{hq_e}^2)^T$  and  $\mathbf{W}$  a known  $m \times q_e$  zero-one design matrix. The complete-data log likelihood is therefore given by

$$\begin{aligned} \log L_c(\Psi) &= \sum_{h=1}^g \left[ \sum_{j=1}^n z_{hj} \log \pi_h - \frac{1}{2} \left\{ \sum_{j=1}^n z_{hj} q_b \log \theta_{bh} + \right. \right. \\ &\quad \left. \left. q_c \log \theta_{ch} + \sum_{j=1}^n z_{hj} \log |\mathbf{A}_h| + \right. \right. \\ &\quad \left. \left. \frac{\mathbf{b}_h^T \mathbf{b}_h}{\theta_{bh}} + \frac{\mathbf{c}_h^T \mathbf{c}_h}{\theta_{ch}} + \boldsymbol{\epsilon}_h^T \boldsymbol{\Omega}_h \boldsymbol{\epsilon}_h \right\} \right], \quad (2) \end{aligned}$$

apart from an additive constant, as described in Section 3 of the main text. In (2),  $\boldsymbol{\Omega}_h = \mathbf{I}_n \otimes \mathbf{A}_h^{-1}$  for  $h = 1, \dots, g$ . Since  $\mathbf{b}_h^T \mathbf{b}_h$ ,  $\mathbf{c}_h^T \mathbf{c}_h$ ,  $\boldsymbol{\epsilon}_h^T \boldsymbol{\epsilon}_h$ , and  $(\mathbf{y}_j - \mathbf{U}\mathbf{b}_{hj} - \mathbf{V}\mathbf{c}_h)$  are sufficient statistics (McCulloch and Searle, 2001; Searle *et al.*, 1992) for the complete model (2), the conditional expectation of the complete-data log likelihood is effected simply by replacing these sufficient statistics in (2) by their

\*to whom correspondence should be addressed

conditional expectations, where the expectation is with respect to the joint distribution of the missing data given the observed data and the current estimate of  $\Psi$ . On the  $(k + 1)$ th iteration of the EM algorithm, the conditional expectation,  $E_{\Psi^{(k)}}(z_{hj} | \mathbf{y})$  is given by

$$E_{\Psi^{(k)}}(z_{hj} | \mathbf{y}) = \int \tau_h(\mathbf{y}_j, \mathbf{c}; \Psi^{(k)}) f(\mathbf{c} | \mathbf{y}) d\mathbf{c},$$

where

$$\begin{aligned} \tau_h(\mathbf{y}_j, \mathbf{c}; \Psi^{(k)}) &= \text{pr}_{\Psi^{(k)}} \{Z_{hj} = 1 | \mathbf{y}_j, \mathbf{c}\} \\ &= \frac{\pi_h^{(k)} f(\mathbf{y}_j | z_{hj} = 1, \mathbf{c}_h; \boldsymbol{\psi}_h^{(k)})}{\sum_{i=1}^g \pi_i^{(k)} f(\mathbf{y}_j | z_{ij} = 1, \mathbf{c}_i; \boldsymbol{\psi}_i^{(k)})}, \quad (3) \end{aligned}$$

is the posterior probability that the  $j$ th gene belongs to the  $h$ th component of the mixture given  $\mathbf{y}_j$  and  $\mathbf{c}$ , and

$$\begin{aligned} \log f(\mathbf{y}_j | z_{hj} = 1, \mathbf{c}_h; \boldsymbol{\psi}_h^{(k)}) &= -\frac{1}{2} \left\{ \log |\mathbf{B}_h^{(k)}| + \right. \\ &\quad \left. (\mathbf{y}_j - \mathbf{X}\beta_h^{(k)} - \mathbf{V}\mathbf{c}_h)^T \mathbf{B}_h^{(k)-1} (\mathbf{y}_j - \mathbf{X}\beta_h^{(k)} - \mathbf{V}\mathbf{c}_h) \right\} \quad (4) \end{aligned}$$

is the log  $h$ th component density of  $\mathbf{y}_j$  conditional on  $\mathbf{c}_h$ , apart from an additive constant. Here  $\mathbf{B}_h^{(k)} = \mathbf{A}_h^{(k)} + \theta_{bh}^{(k)} \mathbf{U}\mathbf{U}^T$ . Based on standard normal theory (Searle *et al.*, 1992), after integrating over the conditional distribution of the random effects  $\mathbf{c}$  given  $\mathbf{y}$ , we have

$$E_{\Psi^{(k)}}(z_{hj} | \mathbf{y}) \propto \pi_h^{(k)} f(\mathbf{y}_j | z_{hj} = 1; \boldsymbol{\psi}_h^{(k)}), \quad (5)$$

where

$$\begin{aligned} \log f(\mathbf{y}_j | z_{hj} = 1; \boldsymbol{\psi}_h^{(k)}) &= -\frac{1}{2} \left\{ \log |\mathbf{B}_h^{(k)} + \mathbf{D}_h^{(k)}| + \right. \\ &\quad \left. (\mathbf{y}_j - \mathbf{X}\beta_h^{(k)})^T [\mathbf{B}_h^{(k)} + \mathbf{D}_h^{(k)}]^{-1} (\mathbf{y}_j - \mathbf{X}\beta_h^{(k)}) \right\}, \end{aligned}$$

apart from an additive constant, is the marginal log density of  $\mathbf{y}_j$  given that it belongs to the  $h$ th component ( $h = 1, \dots, g$ ), and where  $\mathbf{D}_h^{(k)} = \theta_{ch}^{(k)} \mathbf{V}\mathbf{V}^T$ . From (5), it follows that the conditional

expectation  $E_{\Psi^{(k)}}(z_{hj} | \mathbf{y})$  is given by

$$\begin{aligned} E_{\Psi^{(k)}}(z_{hj} | \mathbf{y}) &= \tau_{hj}^{(k)} \\ &= \frac{\pi_h^{(k)} f(\mathbf{y}_j | z_{hj} = 1; \boldsymbol{\psi}_h^{(k)})}{\sum_{i=1}^g \pi_i^{(k)} f(\mathbf{y}_j | z_{ij} = 1; \boldsymbol{\psi}_i^{(k)})}. \end{aligned} \quad (6)$$

Concerning the conditional expectation,  $E_{\Psi^{(k)}}(\mathbf{b}_h^T \mathbf{b}_h | \mathbf{y})$ , it is given by

$$E_{\Psi^{(k)}}(\mathbf{b}_h^T \mathbf{b}_h | \mathbf{y}) = \sum_{j=1}^n \tau_{hj}^{(k)} \mathbf{b}_{hj}^{(k)T} \mathbf{b}_{hj}^{(k)} + \zeta_{bh}^{(k)}, \quad (7)$$

where

$$\begin{aligned} \mathbf{b}_{hj}^{(k)} &= E_{\Psi^{(k)}}(\mathbf{b}_{hj} | \mathbf{y}) \\ &= \theta_{bh}^{(k)} \mathbf{U}^T \mathbf{B}_h^{(k)-1} \left[ (\mathbf{y}_j - \mathbf{X} \boldsymbol{\beta}_h^{(k)}) - \mathbf{D}_h^{(k)} \mathbf{M}_h^{(k)} \sum_{l=1}^n \tau_{hl}^{(k)} (\mathbf{y}_l - \mathbf{X} \boldsymbol{\beta}_h^{(k)}) \right], \end{aligned}$$

and where

$$\mathbf{M}_h^{(k)} = \left[ \mathbf{B}_h^{(k)} + \sum_{j=1}^n \tau_{hj}^{(k)} \mathbf{D}_h^{(k)} \right]^{-1}.$$

In (7),  $\zeta_{bh}^{(k)}$  is the trace of the current conditional covariance matrix  $\text{cov}(\mathbf{b}_h | \mathbf{y})$  and is given by

$$\begin{aligned} \zeta_{bh}^{(k)} &= \sum_{j=1}^n \tau_{hj}^{(k)} q_b \theta_{bh}^{(k)} - \left( \sum_{j=1}^n \tau_{hj}^{(k)} - 1 \right) \theta_{bh}^{(k)2} \text{trace}(\mathbf{U}^T \mathbf{B}_h^{(k)-1} \mathbf{U}) \\ &\quad - \theta_{bh}^{(k)2} \text{trace}(\mathbf{U}^T \mathbf{M}_h^{(k)} \mathbf{U}). \end{aligned}$$

Concerning the conditional expectation  $E_{\Psi^{(k)}}(\mathbf{c}_h^T \mathbf{c}_h | \mathbf{y})$ , we have

$$E_{\Psi^{(k)}}(\mathbf{c}_h^T \mathbf{c}_h | \mathbf{y}) = \mathbf{c}_h^{(k)T} \mathbf{c}_h^{(k)} + \zeta_{ch}^{(k)}, \quad (8)$$

where

$$\begin{aligned} \mathbf{c}_h^{(k)} &= E_{\Psi^{(k)}}(\mathbf{c}_h | \mathbf{y}) \\ &= \theta_{ch}^{(k)} \mathbf{V}^T \mathbf{M}_h^{(k)} \sum_{l=1}^n \tau_{hl}^{(k)} (\mathbf{y}_l - \mathbf{X} \boldsymbol{\beta}_h^{(k)}) \end{aligned}$$

and

$$\zeta_{ch}^{(k)} = q_c \theta_{ch}^{(k)} - \sum_{j=1}^n \tau_{hj}^{(k)} \theta_{ch}^{(k)2} \text{trace}(\mathbf{V}^T \mathbf{M}_h^{(k)} \mathbf{V})$$

is the trace of the conditional covariance matrix  $\text{cov}(\mathbf{c}_h | \mathbf{y})$ .

The last conditional expectation  $E_{\Psi^{(k)}}(\boldsymbol{\epsilon}_h^T \boldsymbol{\Omega}_h \boldsymbol{\epsilon}_h | \mathbf{y})$  is given by

$$E_{\Psi^{(k)}}(\boldsymbol{\epsilon}_h^T \boldsymbol{\Omega}_h \boldsymbol{\epsilon}_h | \mathbf{y}) = \sum_{j=1}^n \tau_{hj}^{(k)} \boldsymbol{\epsilon}_{hj}^{(k)T} \mathbf{A}_h^{(k)-1} \boldsymbol{\epsilon}_{hj}^{(k)} + \zeta_h^{(k)}, \quad (9)$$

where

$$\boldsymbol{\epsilon}_{hj}^{(k)} = \mathbf{y}_j - \mathbf{X} \boldsymbol{\beta}_h^{(k)} - \mathbf{U} \mathbf{b}_{hj}^{(k)} - \mathbf{V} \mathbf{c}_h^{(k)}$$

and

$$\begin{aligned} \zeta_h^{(k)} &= \sum_{j=1}^n \tau_{hj}^{(k)} \text{trace}(\mathbf{A}_h^{(k)}) - \text{trace}(\mathbf{A}_h^{(k)T} \mathbf{M}_h^{(k)} \mathbf{A}_h^{(k)}) - \\ &\quad \left( \sum_{j=1}^n \tau_{hj}^{(k)} - 1 \right) \text{trace}(\mathbf{A}_h^{(k)T} \mathbf{B}_h^{(k)-1} \mathbf{A}_h^{(k)}). \end{aligned}$$

## 2 THE M-STEP

As described in Section 3 of the main text, the M-step updates the estimates that maximize the  $Q$ -function with respect to  $\Psi$ . With reference to (2) and the conditional expectations given by (6) to (9), the updating formulae for  $\Psi^{(k+1)}$  are given by

$$\pi_h^{(k+1)} = \frac{\sum_{j=1}^n \tau_{hj}^{(k)}}{n}, \quad (10)$$

$$\boldsymbol{\beta}_h^{(k+1)} = \boldsymbol{\beta}_h^{(k)} + \frac{\mathbf{G}_h^{(k)} \sum_{j=1}^n \tau_{hj}^{(k)} (\mathbf{y}_j - \mathbf{X} \boldsymbol{\beta}_h^{(k)})}{\sum_{j=1}^n \tau_{hj}^{(k)}}, \quad (11)$$

$$\theta_{bh}^{(k+1)} = \frac{\sum_{j=1}^n \tau_{hj}^{(k)} \mathbf{b}_{hj}^{(k)T} \mathbf{b}_{hj}^{(k)} + \zeta_{bh}^{(k)}}{(q_b \sum_{j=1}^n \tau_{hj}^{(k)})}, \quad (12)$$

$$\theta_{ch}^{(k+1)} = \frac{\mathbf{c}_h^{(k)T} \mathbf{c}_h^{(k)} + \zeta_{ch}^{(k)}}{q_c}, \quad (13)$$

and

$$\sigma_{hl}^{(k+1)2} = \frac{\sum_{j=1}^n \tau_{hj}^{(k)} \boldsymbol{\epsilon}_{hj}^{(k)T} \mathbf{A}_h^{(k)-1} \boldsymbol{\epsilon}_{hj}^{(k)} + \zeta_h^{(k)}}{(\mathbf{W}_l^T \mathbf{W}_l \sum_{j=1}^n \tau_{hj}^{(k)})} \quad (14)$$

for  $l = 1, \dots, q_e$ , and where

$$\mathbf{G}_h^{(k)} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{M}_h^{(k)} \mathbf{A}_h^{(k)}$$

and  $\mathbf{W}_l$  is the  $l$ th column vector of  $\mathbf{W}$  ( $l = 1, \dots, q_e$ ).

The E- and M-steps are alternated repeatedly until convergence of the EM sequence of iterates (Ng *et al.*, 2004), as described in Section 3 of the main text. It can be seen that an initial parameter value of the vector of unknown parameters  $\Psi$ ,  $\Psi^{(0)}$ , has to be specified. The monograph of ? provides an in-depth account of the choice of initial values and the effects of different starting strategies on parameter estimates. Briefly, with mixture models the likelihood typically will have multiple maxima. Hence in practice the EM algorithm needs to be started from a variety of initial values for the parameter vector  $\Psi$ . With our method, we adopt the automatic approach used in the EMMIX program of ? to obtain initial values  $\Psi^{(0)}$  for  $\Psi$  (EMMIX program is available online at the website <http://www.maths.uq.edu.au/~gjm/emmix/emmix.html>). With this automatic approach, the EMMIX algorithm is run from ten starts corresponding to clusterings of the data

by  $k$ -means ( $k$ -means clustering-based starts). The parameter estimates, which correspond to the clustering that produces the highest log likelihood, are adopted as the initial estimates  $\Psi^{(0)}$  for the purposes of starting the EM algorithm.

### 3 AN OUTRIGHT CLUSTERING

To effect a probabilistic or an outright clustering of the correlated genes into  $g$  components, we calculate the estimated posterior probabilities of component membership conditional on the cluster random-effects vector  $\mathbf{c}$  and the observed expression levels  $\mathbf{y}$  for the genes. The random effects  $\mathbf{c}_h$  ( $h = 1, \dots, g$ ) are, however, unobservable, so we use their estimated conditional expectations given the observed data,  $\hat{\mathbf{c}}_h = E_{\hat{\Psi}}(\mathbf{c}_h | \mathbf{y})$ , where  $\hat{\Psi}$  is the ML estimate of  $\Psi$ . Since the genes within a cluster are independently distributed given  $\mathbf{c}_h$ , it suffices to effect a clustering with each gene considered individually in terms of its (estimated) posterior probabilities of component membership given its profile vector  $\mathbf{y}_j$  and  $\mathbf{c}$ , which are given by  $\tau_h(\mathbf{y}_j, \hat{\mathbf{c}}; \hat{\Psi})$ , where

$$\tau_h(\mathbf{y}_j, \mathbf{c}; \Psi) = \text{pr}_{\Psi} \{z_{hj} = 1 | \mathbf{y}_j, \mathbf{c}\} \quad (15)$$

for  $h = 1, \dots, g; j = 1, \dots, n$ . Using Bayes' Theorem,  $\tau_h(\mathbf{y}_j, \mathbf{c}; \Psi)$  can be expressed as

$$\tau_h(\mathbf{y}_j, \mathbf{c}; \Psi) = \frac{\pi_h f(\mathbf{y}_j | z_{hj} = 1, \mathbf{c}_h; \psi_h)}{\sum_{i=1}^g \pi_i f(\mathbf{y}_j | z_{ij} = 1, \mathbf{c}_i; \psi_i)}, \quad (16)$$

where  $\log f(\mathbf{y}_j | z_{hj} = 1, \mathbf{c}_h; \psi_h)$  is equal to (4). An outright clustering of the correlated genes into  $g$  components is therefore achieved by assigning each gene to the component to which it has the highest estimated conditional posterior probability.

### 4 THE LOG LIKELIHOOD FUNCTION

The log likelihood for  $\Psi$  based on  $\mathbf{y}$  is given by

$$\log L(\Psi) = \log \left\{ \iint \prod_{j=1}^n f(\mathbf{y}_j | \mathbf{b}, \mathbf{c}; \Psi) f(\mathbf{b}, \mathbf{c}) d\mathbf{b} d\mathbf{c} \right\}, \quad (17)$$

where  $f(\mathbf{y}_j | \Psi; \mathbf{b}, \mathbf{c}) = \sum_{h=1}^g \pi_h f(\mathbf{y}_j | z_{hj} = 1, \mathbf{b}_h, \mathbf{c}_h; \psi_h)$  and  $f(\mathbf{b}, \mathbf{c})$  denote the  $(q_b + q_c)$ -dimensional normal density of the random effects  $\mathbf{b}$  and  $\mathbf{c}$ . The integration can be obtained in two stages where the inner expectation is with respect to  $\mathbf{b}$  only. This integration can be effected in closed form. However, the outer integration with respect to  $\mathbf{c}$  cannot be obtained in closed form. Here we proceed approximately by forming the likelihood as if the  $\mathbf{y}_j$  were all independently distributed. Using then standard normal theory (Searle *et al.*, 1992), it can be shown that

$$\begin{aligned} \log L(\Psi) &\approx \log \prod_{j=1}^n \sum_{h=1}^g \pi_h f(\mathbf{y}_j | z_{hj} = 1; \psi_h) \\ &= \sum_{j=1}^n \log \sum_{h=1}^g \pi_h f(\mathbf{y}_j | z_{hj} = 1; \psi_h), \end{aligned} \quad (18)$$

where

$$\log f(\mathbf{y}_j; \psi_h) = -\frac{1}{2} \left\{ m \log(2\pi) + \log |\mathbf{B}_h + \mathbf{D}_h| + (\mathbf{y}_j - \mathbf{X}\beta_h)^T [\mathbf{B}_h + \mathbf{D}_h]^{-1} (\mathbf{y}_j - \mathbf{X}\beta_h) \right\}.$$

## 5 RESULTS FOR SEARCH OF COMMON REGULATORY ELEMENTS

As described in Section 5.1 of the main text, we searched through the 700-bp upstream region of the start codon of each gene in Clusters 1, 3, 10, 11, and 12 for the presence of binding site sequences for any known yeast cell cycle transcription factors like MBF, SBF, Mcm1p-containing factors, and Swi5p factors. We found that the majority of the genes in these clusters share common promoter elements that are identifiable based on the published literature.

Genes in Clusters 1 and 10 show typical G1 peak expression in cell cycle and were the major members of the "CLN2" cluster described by Spellman *et al.* (1998). Cluster 1 contains mostly genes that are involved in DNA replication and repair. 91% of the genes contain either an SCB element or an MCB element in their promoter, commonly seen in genes expressed at the G1/S transition (Koch and Nasmyth, 1994). 80% of genes contain two or more of these elements, with 46% of the genes containing both kinds of elements in their promoters. These genes are usually strongly induced by CLN3, but are strongly repressed by CLB2, as the SBF and MBF transcription factors (which bind to the SCB and MCB elements) are posttranslationally activated by CLN3, and inactivated by CLB2 (Amon *et al.*, 1993). Cluster 10 (11 genes) contains 2 genes involved in DNA replication, with the remaining members encompassing other various functions. Members of this cluster also contain a high percentage of SCB and MCB elements (91% of the genes contain either one element, while 55% contain both). But there is a higher percentage (45%) of genes in Cluster 10 that also contain a Swi5p site, than in the 28% of genes in Cluster 1.

Genes in Cluster 3 contain genes previously clustered in the "CLB2" cluster of Spellman's work. These genes include CLB1, CDC5, and CDC20. They are genes involved in mitosis and their expressions have a peak in the M phase, and are regulated by the MCM1 and SFF transcription factors that induce genes during mitosis (Althoefer *et al.*, 1995; Sanders and Herskowitz, 1996).

Another two interesting clusters are Clusters 11 and 12. Cluster 11 contains members of the "Y" cluster described by Spellman *et al.* (1998), which are genes located in Y' elements located at chromosomal ends, and with very high sequence similarity (Spellman *et al.*, 1998). The expressions of these genes have a peak in the G1 phase, as do Clusters 1 and 10. Seventy-seven percent of these genes contain either a SCB or a MCB element. However, only 35% of genes have two or more elements in their promoters, and the function and regulation of these genes are still largely unknown to date. Cluster 12 is composed mostly of histone genes. Their expressions are tightly peaked in the S phase, and there are very high peak-to-trough ratios. Histones are known to be regulated by a negative element in the 3' region of the mRNAs that destabilizes the mRNAs except during S phase; and a repeated positive element ATGCGAAR, that activates transcription (Freeman *et al.*, 1992).

## REFERENCES

- Althoefer, H., Schleiffer, A., Wassmann, K., Nordheim, A., and Ammerer, G. (1995) Mcm1 is required to coordinate G2-specific transcription in *Saccharomyces cerevisiae*, *Mol. Cell. Biol.*, **15**, 5917-5928.
- Amon, A., Tyers, M., Fitcher, B., and Nasmyth, K. (1993) Mechanisms that help the yeast cell cycle clock tick: G2 cyclins transcriptionally activate G2 cyclins and repress G1 cyclins, *Cell*,

- 74, 993-1007.
- Freeman, K.B., Karns, L.R., Lutz, K.A., and Smith, M.M. (1992) Histone H3 transcription in *Saccharomyces cerevisiae* is controlled by multiple cell cycle activation sites and a constitutive negative regulatory element, *Mol. Cell. Biol.*, **12**, 5455-5463.
- Koch, C. and Nasmyth, K. (1994) Cell cycle regulated transcription in yeast, *Curr. Opin. Cell Biol.*, **6**, 451-459.
- McCulloch, C.E. and Searle, S.R. (2001) *Generalized, Linear, and Mixed Models*. Wiley, New York.
- Ng, S.K., Krishnan, T., and McLachlan, G.J. (2004) The EM algorithm. In Gentle, J., Hardle, W., and Mori, Y. (eds), *Handbook of Computational Statistics Vol. 1*. Springer-Verlag, New York, pp. 137-168.
- Sanders, S.L. and Herskowitz, I. (1996) The BUD4 protein of yeast, required for axial budding, is localized to the mother/BUD neck in a cell cycle-dependent manner, *J. Cell Biol.*, **134**, 413-427.
- Searle, S.R., Casella, G., and McCulloch, C.E. (1992) *Variance Components*. Wiley, New York.
- Spellman, P., Sherlock, G., Zhang, M.Q. *et al.* (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization, *Mol. Biol. Cell*, **9**, 3273-3297.