

Fig. 7. SVD of the normalized and sorted elutriation data. Raster display of  $\hat{e}_N = \hat{u}_N \hat{\varepsilon}_N \hat{v}_N^T$  with overexpression (red), no change in expression (black), and underexpression (green) around the steady state of expression, showing linear transformation of the data from the 5,981-genes  $\times$  14-arrays space to the reduced diagonalized 14-eigenarrays  $\times$  14-eigengenes space using the 5,981-genes  $\times$  14-eigenarrays and 14-eigengenes  $\times$  14-arrays basis sets.

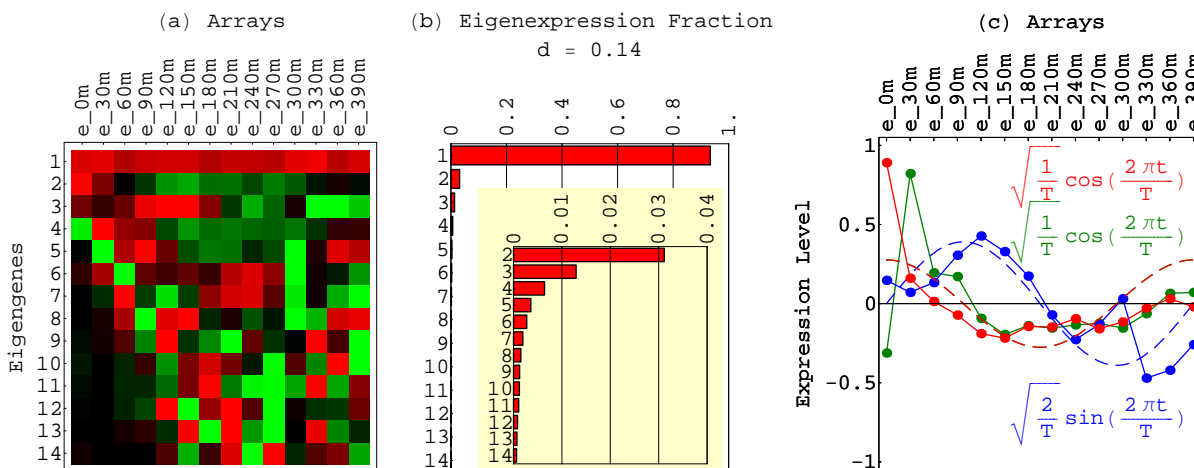


Fig. 8. Elutriation eigengenes. (a) Raster display of  $\hat{v}^T$ , the expression of 14 eigengenes in 14 arrays, with overexpression (red), no change in expression (black), and underexpression (green) around the steady state, which can be associated with the first eigengene,  $|\gamma_1\rangle$ . (b) Bar chart of the fraction of eigenexpression  $p_l$  of each eigengene  $|\gamma_l\rangle$ , showing more than 90% of the overall relative expression in  $|\gamma_1\rangle$ , about 3%, 1.5%, and 0.5% in  $|\gamma_2\rangle$ ,  $|\gamma_3\rangle$ , and  $|\gamma_4\rangle$ , respectively, and a low entropy  $d = 0.14 \ll 1$ . (c) Line-joined graphs of the expression levels of  $|\gamma_2\rangle$  (red),  $|\gamma_3\rangle$  (blue), and  $|\gamma_4\rangle$  (green) in the 14 arrays, and dashed graphs of normalized cosine (blue) and sine (red) of period  $T$ .

## APPENDIX

### BIOLOGICAL DATA ANALYSIS: $\alpha$ FACTOR-SYNCHRONIZED CELL CYCLE AND *CLB2* AND *CLN3* OVERACTIVATIONS

**Pattern Inference and Data Normalization.** Consider the 22 eigengenes of the  $\alpha$  factor, *CLB2*, and *CLN3* dataset (Fig. 11a in supplemental material at [www.pnas.org](http://www.pnas.org) and at <http://genome-www.stanford.edu/SVD/>). The entropy of this dataset  $d = 0.25$  (Fig. 11b), is higher than that of the elutriation dataset, because the  $\alpha$  factor, *CLB2*, and *CLN3* dataset combines three experiments, and therefore is less redundant. The first eigengene  $|\gamma_1\rangle$ , which captures about 80% of the overall relative expres-

sion, describes time (and experiment) invariant relative expression (Fig. 11c). We infer that  $|\gamma_1\rangle$  represents the steady-state expression, and filter it out, approximately centering the dataset. The eigengenes  $|\gamma_2\rangle$  and  $|\gamma_3\rangle$  describe initial transient increase and decrease, respectively, in relative expression superimposed over time-invariant relative expression during the cell cycle. We infer that  $|\gamma_2\rangle$  and  $|\gamma_3\rangle$  represent the responses to synchronization by  $\alpha$  factor, and filter them out. The eigengenes  $|\gamma_{10}\rangle$  and  $|\gamma_l\rangle$  for  $12 \leq l \leq L = 22$  all show rapidly varying expression during the cell cycle and steady-state expression in the *CLB2*- and *CLN3*-overactive arrays (Fig. 11a), and we filter them out as well, such that

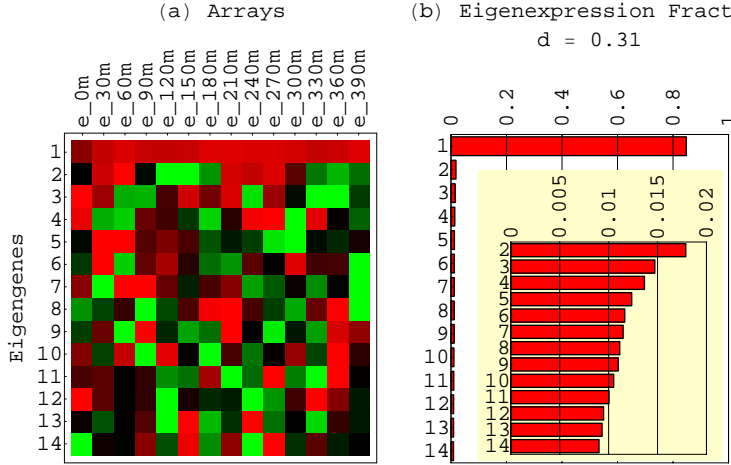


Fig. 9. Eigengenes of the natural logarithm of the variances in elutriation expression  $\hat{e}_{LV}$ . (a) Raster display of  $\hat{v}_{LV}^T$ ;  $|\gamma_1\rangle_{LV}$  is inferred to represent the steady scale of expression variance. (b)  $|\gamma_1\rangle_{LV}$  captures more than 80% of the overall information in this dataset.

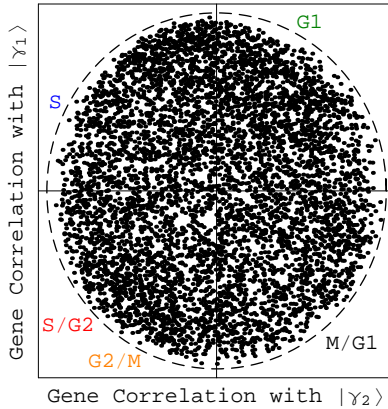


Fig. 10. Normalized elutriation expression in the subspace associated with the cell cycle. Correlation of each gene with  $|\gamma_1\rangle_N$  vs. that with  $|\gamma_2\rangle_N$  for all 5,981 genes, where classification of these genes based on their angular distance from the  $x$ -axis gives 1,582 in M/G<sub>1</sub>, 1,524 in G<sub>1</sub>, 1,344 in S, 688 in S/G<sub>2</sub>, and 843 in G<sub>2</sub>/M.

$$\hat{e} \rightarrow \hat{e}_C = \hat{e} - \sum_{l=1}^3 \varepsilon_l |\alpha_l\rangle \langle \gamma_l| - \varepsilon_{10} |\alpha_{10}\rangle \langle \gamma_{10}| - \sum_{l=12}^{22} \varepsilon_l |\alpha_l\rangle \langle \gamma_l| = \sum_{l=4}^9 \varepsilon_l |\alpha_l\rangle \langle \gamma_l| + \varepsilon_{11} |\alpha_{11}\rangle \langle \gamma_{11}|.$$

The remaining eigengenes  $|\gamma_l\rangle$  for  $4 \leq l \leq 9$  and  $|\gamma_{11}\rangle$  show expression oscillations of two periods during the  $\alpha$  factor-synchronized cell cycle, from  $t = 7$  to 119 min for a duration of  $T = 112$  min (Fig. 11 *d-f*), suggesting a delay of 7 min from the start of the experiment to the start of the progress of a cell cycle of period  $T/2 = 66$  min in the initially arrested culture. We infer that these

eigengenes represent cell cycle expression oscillations and assume that the corresponding eigenarrays represent the corresponding cell cycle cellular states.

Let  $\hat{e}_{LV}$  tabulate the natural logarithm of the variances in expression in the  $\alpha$  factor, *CLB2*, and *CLN3* experiments, such that each element of  $\hat{e}_{LV}$  satisfies  $\langle n | \hat{e}_{LV} | m \rangle \equiv \log(e_{C,nm}^2)$ , and consider the eigengenes of  $\hat{e}_{LV}$  (Fig. 12*a*). The first eigengene  $|\gamma_1\rangle_{LV}$ , which captures about 90% of the overall information in this dataset (Fig. 12*b*), describes a time (and experiment) invariant scale of expression variance, and we filter it out,  $\hat{e}_{LV} \rightarrow \hat{e}_{CLV} = \hat{e}_{LV} - \varepsilon_{1,LV} |\alpha_1\rangle_{LV} \langle \gamma_1|$ .

The normalized  $\alpha$  factor, *CLB2*, and *CLN3* dataset  $\hat{e}_N$ , where each of its elements satisfies  $\langle n | \hat{e}_N | m \rangle \equiv \text{sign}(e_{CS,nm}) \sqrt{\exp(e_{CLV,nm})}$ , tabulates for each gene and array expression patterns that are approximately centered at the steady-state expression level (i.e., of approximately zero arithmetic means), with variances which are approximately normalized by the steady scale of expression variance (i.e., of approximately unit geometric means).

Note that the entropy of  $\hat{e}_N$  is  $d = 0.62$ , suggesting that  $\hat{e}_N$  is less redundant (and maybe more informative) than  $\hat{e}$ . The eigengenes  $|\gamma_1\rangle_N$ ,  $|\gamma_2\rangle_N$ , and  $|\gamma_3\rangle_N$  of  $\hat{e}_N$  (Fig. 13*a*), which are of similar significance, capture together about 60% of the overall normalized expression (Fig. 13*b*). Their time variations fit normalized sine and cosine functions of two  $T/2 = 66$ -min periods superimposed on normalized sine function of one  $T = 112$ -min period from  $t = 7$  to 119 min during the cell cycle (Fig. 13*c*). While  $|\gamma_1\rangle_N$  and  $|\gamma_3\rangle_N$  describe underexpression in both *CLB2*-overactive arrays  $|a_{19}\rangle$  and  $|a_{20}\rangle$ , and overexpression in both *CLN3*-overactive arrays  $|a_{21}\rangle$  and  $|a_{22}\rangle$ ,  $|\gamma_2\rangle$  describes the antiparallel expression pattern of overexpression in  $|a_{19}\rangle$  and  $|a_{20}\rangle$  and underexpression in  $|a_{21}\rangle$  and  $|a_{22}\rangle$ .

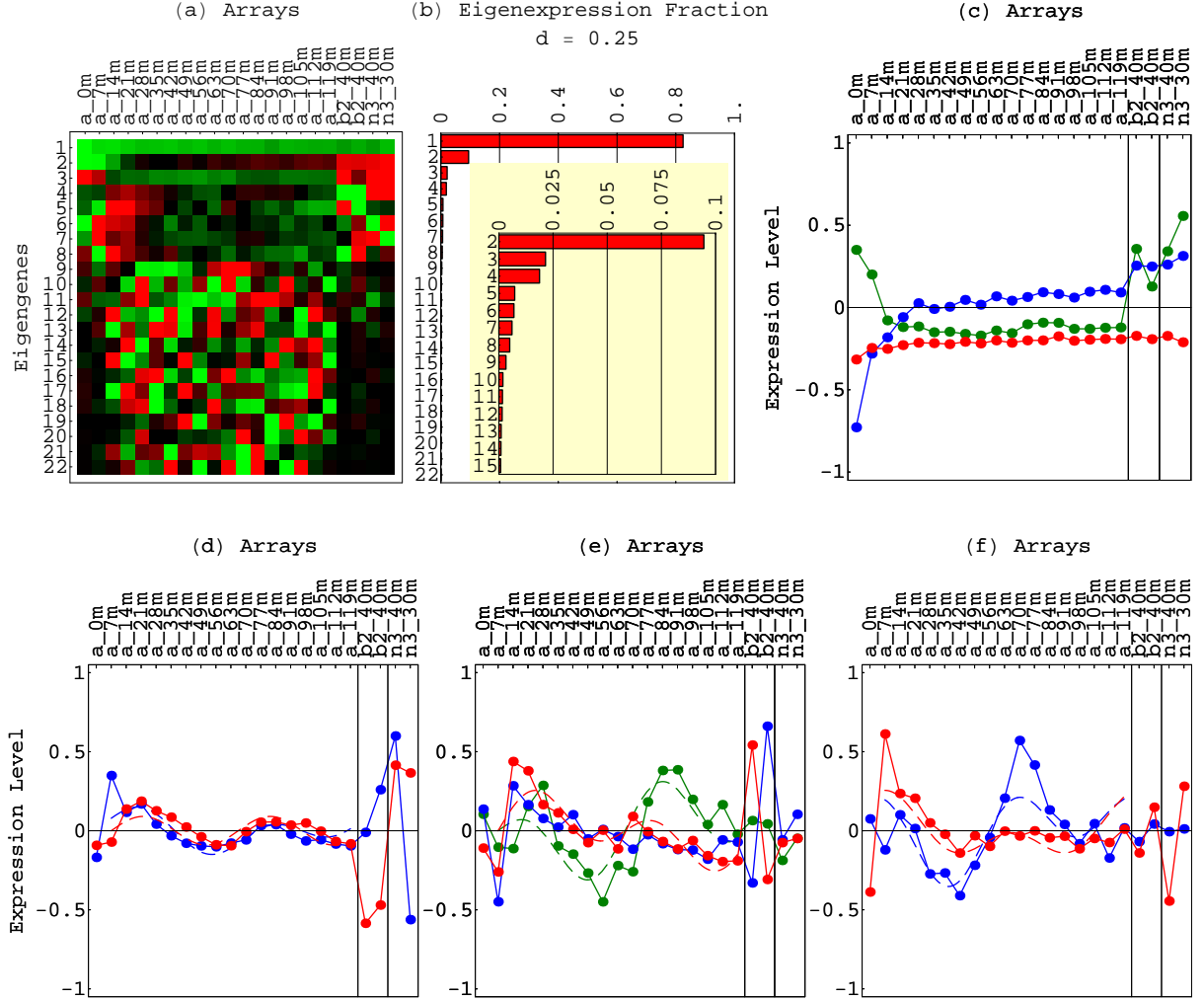


Fig. 11.  $\alpha$  factor, *CLB2*, and *CLN3* eigengenes. (a) Raster display of  $\hat{v}^T$ , the expression of 22 eigengenes in 22 arrays. (b) Bar chart of the fractions of eigenexpression, showing that  $|\gamma_1\rangle$  captures about 80% of the overall relative expression. (c) Line-joined graphs of the expression levels of  $|\gamma_1\rangle$  (red), which is inferred to represent the steady expression state, and  $|\gamma_2\rangle$  (blue) and  $|\gamma_3\rangle$  (green), which are inferred to represent responses to synchronization in the cell cycle. (d) Expression levels of  $|\gamma_4\rangle$  (red) and  $|\gamma_7\rangle$  (blue). (e) Expression levels of  $|\gamma_5\rangle$  (red),  $|\gamma_8\rangle$  (blue), and  $|\gamma_{11}\rangle$  (green). (f) Expression levels of  $|\gamma_6\rangle$  (red) and  $|\gamma_9\rangle$  (blue). All fit dashed graphs of periodic functions with period  $T/2 = 66$  min superimposed on periodic functions with period  $T = 112$  min from  $t = 7$  to  $t = 119$  min during the cell cycle.

**Degenerate Subspace Rotation.** We approximate the eigenexpression levels  $\varepsilon_{1,N} \approx \varepsilon_{2,N} \approx \varepsilon_{3,N}$  with  $\varepsilon_{1,RN} = \varepsilon_{2,RN} = \varepsilon_{3,RN} = \sqrt{(\varepsilon_{1,N}^2 + \varepsilon_{2,N}^2 + \varepsilon_{3,N}^2)/3}$ . First, we rotate the eigengenes and eigenarrays

$$\begin{aligned} |\gamma_1\rangle_N &\rightarrow \hat{R}_1|\gamma_1\rangle_N = \cos \rho_1|\gamma_1\rangle_N + \sin \rho_1|\gamma_2\rangle_N, \\ |\gamma_2\rangle_N &\rightarrow \hat{R}_1|\gamma_2\rangle_N = -\sin \rho_1|\gamma_1\rangle_N + \cos \rho_1|\gamma_2\rangle_N, \\ |\alpha_1\rangle_N &\rightarrow \hat{R}_1|\alpha_1\rangle_N = \cos \rho_1|\alpha_1\rangle_N + \sin \rho_1|\alpha_2\rangle_N, \\ |\alpha_2\rangle_N &\rightarrow \hat{R}_1|\alpha_2\rangle_N = -\sin \rho_1|\alpha_1\rangle_N + \cos \rho_1|\alpha_2\rangle_N. \end{aligned}$$

Requiring the rotated second eigengene  $\hat{R}_1|\gamma_2\rangle_N$  to describe equal expression in the *CLB2*-overactive array  $|a_{20}\rangle$  and the *CLN3*-overactive array  $|a_{21}\rangle$ , both mea-

sured at  $t = 40$  min after the start of overactivation, such that  ${}_N\langle a_{20}|\hat{R}_1|\gamma_2\rangle_N = {}_N\langle a_{21}|\hat{R}_1|\gamma_2\rangle_N$ , gives the unique rotation angle  $\rho_1 \approx -\pi/6$ . Second, we rotate the eigengenes and eigenarrays

$$\begin{aligned} \hat{R}_1|\gamma_1\rangle_N &\rightarrow \hat{R}_2\hat{R}_1|\gamma_1\rangle_N = \cos \rho_2\hat{R}_1|\gamma_1\rangle_N + \sin \rho_2|\gamma_3\rangle_N, \\ |\gamma_3\rangle_N &\rightarrow \hat{R}_2|\gamma_3\rangle_N = \sin \rho_2\hat{R}_1|\gamma_1\rangle_N - \cos \rho_2|\gamma_3\rangle_N, \\ \hat{R}_1|\alpha_1\rangle_N &\rightarrow \hat{R}_2\hat{R}_1|\alpha_1\rangle_N = \cos \rho_2\hat{R}_1|\alpha_1\rangle_N + \sin \rho_2|\alpha_3\rangle_N, \\ |\alpha_3\rangle_N &\rightarrow \hat{R}_2|\alpha_3\rangle_N = \sin \rho_2\hat{R}_1|\alpha_1\rangle_N - \cos \rho_2|\alpha_3\rangle_N. \end{aligned}$$

Requiring the rotated third eigengene  $\hat{R}_2|\gamma_3\rangle_N$  to describe equal expression in the arrays  $|a_{20}\rangle$  and  $|a_{21}\rangle$ , such that  ${}_N\langle a_{20}|\hat{R}_2|\gamma_3\rangle_N = {}_N\langle a_{21}|\hat{R}_2|\gamma_3\rangle_N$ , gives the unique rotation angle  $\rho_2 \approx \pi/4$ .

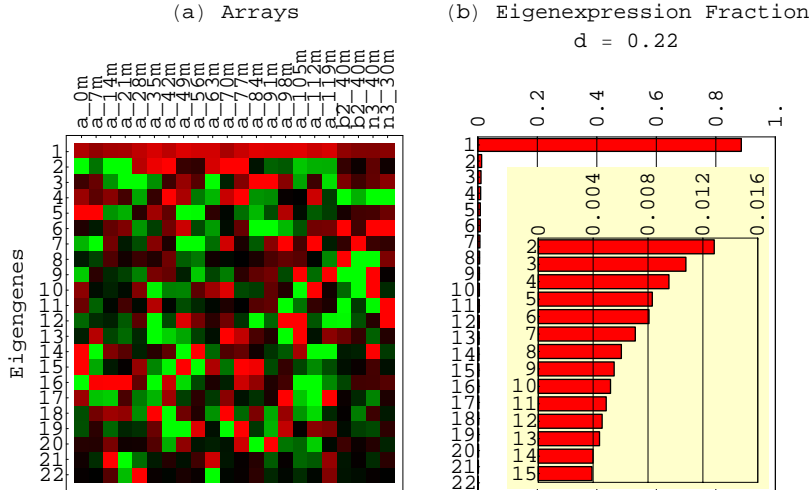


Fig. 12. Eigengenes of the natural logarithm of the variances in expression in the  $\alpha$  factor, *CLB2*, and *CLN3* experiments  $\hat{e}_{LV}$ . (a) Raster display of  $\hat{v}_{LV}^T$ ;  $|\gamma_1\rangle_{LV}$  is inferred to represent the steady scale of expression variance. (b)  $|\gamma_1\rangle_{LV}$  captures about 90% of the overall information in this dataset.

After these rotations (Fig. 4), the time variations of  $|\gamma_1\rangle_{RN} = \hat{R}_2 \hat{R}_1 |\gamma_1\rangle_N$  and  $|\gamma_2\rangle_{RN} = \hat{R}_1 |\gamma_2\rangle_N$  fit normalized sine and cosine functions of two  $T/2 = 66$ -min periods during the cell cycle, from  $t = 7$  to 119 min, and initial phase  $\theta \approx \pi/4$ ,  $\sqrt{2/T} \sin(4\pi t/T - \theta)$ , and  $\sqrt{2/T} \cos(4\pi t/T - \theta)$ , respectively. Upon sorting of the genes (and arrays) according to  $|\gamma_1\rangle_{RN}$  and  $|\gamma_2\rangle_{RN}$  (and  $|\alpha_1\rangle_{RN} = \hat{R}_2 \hat{R}_1 |\alpha_1\rangle_N$  and  $|\alpha_2\rangle_{RN} = \hat{R}_1 |\alpha_2\rangle_N$ ), the initial phase  $\theta \approx \pi/4$  can be interpreted as a delay of 7 min between the start of the cell cycle progress at  $t = 7$  min and that of the cell cycle stage  $G_1$  at  $t = 14$  min. The time variation of  $|\gamma_3\rangle_{RN} = \hat{R}_2 |\gamma_3\rangle_N$  fits normalized sine function of one  $T = 112$ -min period from  $t = 7$  to 119 min and initial phase  $-\theta/2 \approx -\pi/8$ ,  $\sqrt{2/T} \sin(2\pi t/T + \theta/2)$ , suggesting differences in expression during the first cell cycle 66-min period and the successive 66-min period, which may be due to dephasing of the initially synchronized yeast culture.

While  $|\gamma_2\rangle_{RN}$  and  $|\gamma_3\rangle_{RN}$  describe steady-state expression in the *CLB2*- and *CLN3*-overactive arrays,  $|\gamma_1\rangle_{RN}$  describes underexpression in the *CLB2*-overactive arrays and overexpression in the *CLN3*-overactive arrays. We, therefore, infer that  $|\gamma_1\rangle_{RN}$  represents cell cycle expression oscillations that are *CLB2*- and *CLN3*-dependent, whereas  $|\gamma_2\rangle_{RN}$  represents cell cycle expression oscillations that are *CLB2*- and *CLN3*-independent, and  $|\gamma_3\rangle_{RN}$  represents variations in the cell cycle expression from the first period to the second, which also appear to be *CLB2*- and *CLN3*-independent. We assume that  $|\alpha_1\rangle_{RN}$ ,  $|\alpha_2\rangle_{RN}$ , and  $|\alpha_3\rangle_{RN}$  represent the corresponding cellular states. With this we also infer that the subspace spanned by  $|\gamma_1\rangle_{RN}$  and  $|\gamma_2\rangle_{RN}$  approximately represents all cell cycle expression oscillations, and we assume that the subspace spanned by  $|\alpha_1\rangle_{RN}$  and  $|\alpha_2\rangle_{RN}$  approximately represents all cell cycle cellular states.

**Data Sorting.** Consider the normalized expression of the 22  $\alpha$  factor, *CLB2*, and *CLN3* arrays  $\{|a_m\rangle\}$  in the subspace spanned by  $|\alpha_1\rangle_{RN}$  and  $|\alpha_2\rangle_{RN}$ , which is assumed to represent approximately all cell cycle cellular states (Fig. 5a). All arrays have at least 25% of their normalized expression in this subspace, except for  $|a_{16}\rangle$ , which measures  $t = 105$  min in the cell cycle, and  $|a_{19}\rangle$ ,  $|a_{20}\rangle$ ,  $|a_{21}\rangle$ , and  $|a_{22}\rangle$ , the *CLB2*- and *CLN3*-overactive arrays, respectively. The sorting of the arrays according to their phases  $\{\phi_m\}$ , which describes the transition from the expression pattern  $|\alpha_2\rangle_{RN}$  to  $|\alpha_1\rangle_{RN}$  and back to  $|\alpha_2\rangle_{RN}$ , gives an array order that is similar to that of the cell cycle time points measured by the arrays, an order that describes the progress of the cell cycle expression from the M/ $G_1$  stage through  $G_1$ , S, S/ $G_2$ , and  $G_2$ /M and back to M/ $G_1$  twice. Even though the arrays measure equally spaced cell cycle time points, the arrays are not equally spaced, suggesting that the initially synchronized culture dephases in time, or that the cell cycle may not progress linearly in time.

Since  $|\alpha_1\rangle_{RN}$  is correlated with  $|a_{13}\rangle$  and is  $|a_{14}\rangle$ , and anticorrelated with  $|a_9\rangle$  and  $|a_{10}\rangle$ , we associate  $|\alpha_1\rangle_{RN}$  with the cellular state of the transition from the cell cycle stage  $G_1$  to the stage S, and  $-|\alpha_1\rangle_N$  with the transition from  $G_2$ /M to M/ $G_1$ . Similarly,  $|\alpha_2\rangle_{RN}$  is correlated with the arrays  $|a_2\rangle$ ,  $|a_3\rangle$ ,  $|a_{11}\rangle$ , and  $|a_{12}\rangle$ , and we associate  $|\alpha_1\rangle_{RN}$  with the cellular state of the transition from M/ $G_1$  to  $G_1$ . Also,  $|\alpha_2\rangle_{RN}$  is anticorrelated with  $|a_6\rangle$ ,  $|a_7\rangle$ ,  $|a_{16}\rangle$ , and  $|a_{17}\rangle$ , and therefore we associate  $-|\alpha_1\rangle_{RN}$  with the cellular state of the transition from S to S/ $G_2$ . We also associate the subspace spanned by  $|\alpha_1\rangle_{RN}$  and  $|\alpha_2\rangle_{RN}$  with all cell cycle cellular states. With these associations the phase of  $|a_1\rangle$ ,  $\phi_1 \approx -\pi/4$ , corresponds to the 7-min delay between the start of the progress of the cell cycle at  $t = 7$  min and the start of the cell cycle stage

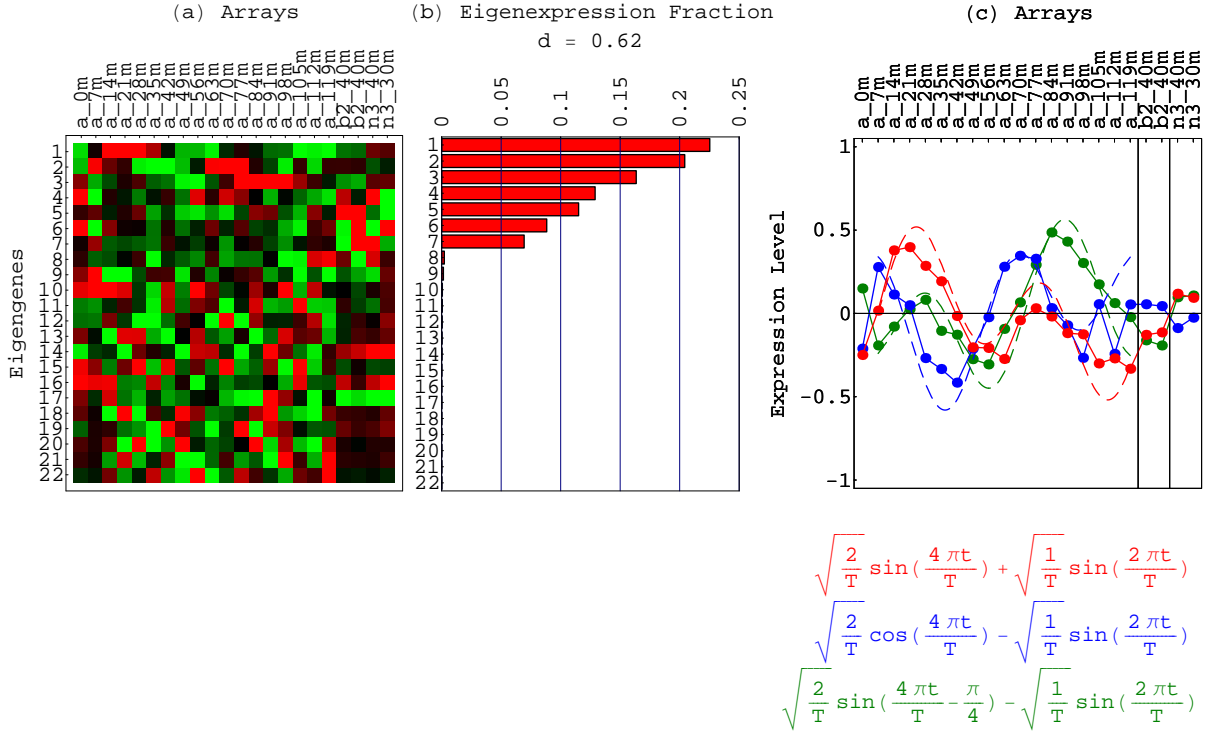


Fig. 13. Normalized  $\alpha$  factor, *CLB2*, and *CLN3* eigengenes. (a) Raster display of  $\hat{v}_N^T$ . (b) Bar chart of the fractions of eigenexpression, showing that  $|\gamma_1\rangle_N$ ,  $|\gamma_2\rangle_N$ , and  $|\gamma_3\rangle_N$  capture more than, about and less than 20% of the overall normalized expression, respectively, and span an approximately degenerate subspace. (c) Line-jointed graphs of the expression levels of  $|\gamma_1\rangle_N$  (red),  $|\gamma_2\rangle_N$  (blue), and  $|\gamma_3\rangle_N$  (green) fit dashed graphs of periodic functions with period  $T/2 = 66$  min superimposed on periodic functions with period  $T = 112$  min from  $t = 7$  to  $t = 119$  min during the cell cycle.

$G_1$  at  $t = 14$  min, which is also present in the inferred cell cycle expression oscillations  $|\gamma_1\rangle_{RN}$  and  $|\gamma_2\rangle_{RN}$ .

The arrays  $|a_{19}\rangle$  and  $|a_{20}\rangle$ , which measure the *CLB2*-overactive samples, and the arrays  $|a_{21}\rangle$  and  $|a_{22}\rangle$ , which measure the *CLN3*-overactive samples, are uncorrelated with  $|\alpha_2\rangle_{RN}$ , and therefore we associate  $|\alpha_2\rangle_{RN}$  with a *CLB2*- and *CLN3*-independent cell cycle cellular state. The arrays  $|a_{19}\rangle$  and  $|a_{20}\rangle$  are anticorrelated with  $|\alpha_1\rangle_{RN}$ , and we associate  $-|\alpha_1\rangle_{RN}$  with *CLB2*-overactive samples. Also, the arrays  $|a_{21}\rangle$  and  $|a_{22}\rangle$  are correlated with  $|\alpha_1\rangle_{RN}$ , and we associate  $|\alpha_1\rangle_{RN}$  with *CLN3*-overactive samples. We also associate the subspace spanned by  $|\alpha_1\rangle_{RN}$  and  $|\alpha_2\rangle_{RN}$  with all cell cycle cellular states, both *CLB2*- and *CLN3*-dependent and independent. The correlation of  $|a_{21}\rangle$  and  $|a_{22}\rangle$  with  $|\alpha_1\rangle_{RN}$  suggests, therefore, that overactivation of *CLN3*, which encodes a  $G_1/S$  cyclin, approximately simulates the transition from the cell cycle stage  $G_1$  to the stage S. The anticorrelation of  $|a_{19}\rangle$  and  $|a_{20}\rangle$  with  $|\alpha_1\rangle_{RN}$  suggests, therefore, that overactivation of *CLB2*, which encodes a  $G_2/M$  cyclin, approximately simulates the half-period advanced (or delayed) cell cycle stage that corresponds to the transition from the cell cycle stage  $G_2/M$  to the stage  $M/G_1$ .

Consider also the expression of the 4,579 genes in the

subspace spanned by  $|\gamma_1\rangle_{RN}$  and  $|\gamma_2\rangle_{RN}$  (Fig. 14). Of the 638 genes that were classified by Spellman *et al.* (3) as cell cycle regulated, 544 have more than 25% of their normalized expression in this subspace (Fig. 5c). We sort all 4,579 genes according to their phases  $\{\phi_n\}$ , to describe the transition from the expression pattern  $|\gamma_2\rangle_{RN}$  to that of  $|\gamma_1\rangle_{RN}$  and back to  $|\gamma_2\rangle_{RN}$ , starting at  $\phi_1 \approx -\pi/4$ , ordering the genes according to the stages in the cell cycle in which their patterns of expression peak, and describing the progress of the cell cycle along the genes. For the 638 cell cycle regulated genes this sorting gives a classification of the genes into the five cell cycle stages, which is in good agreement with the classification by Spellman *et al.*

Since  $|\gamma_1\rangle_{RN}$  is correlated with genes that peak late in the cell cycle stage  $G_1$  and early in the stage S, and anticorrelated with genes that peak late in  $G_2/M$  and early in  $M/G_1$ , we associate  $|\gamma_1\rangle_{RN}$  with the cell cycle expression oscillations that start at the transition from  $G_1$  to S, and  $-|\gamma_1\rangle_{RN}$  with the oscillations that start at the transition from  $G_2/M$  to  $M/G_1$ . Similarly,  $|\gamma_2\rangle_{RN}$  is correlated with genes that peak late in  $M/G_1$  and early in  $G_1$ , and is anticorrelated with genes that peak late in S and early in  $S/G_2$ , and therefore we associate  $|\gamma_2\rangle_{RN}$  with the cell cycle expression oscillations that start at

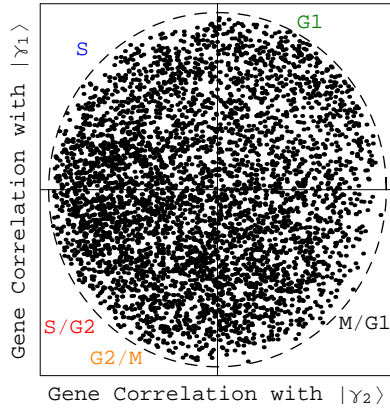


Fig. 14. Rotated normalized  $\alpha$  factor, *CLB2*, and *CLN3* expression in the subspace associated with the cell cycle. Correlation of each gene with  $|\gamma_1\rangle_{RN}$  vs. that with  $|\gamma_2\rangle_{RN}$  for all 4,579 genes, where classification of these genes based on their angular distance from the  $x$ -axis gives 1,020 in M/G<sub>1</sub>, 893 in G<sub>1</sub>, 1,196 in S, 807 in S/G<sub>2</sub>, and 663 in G<sub>2</sub>/M.

the transition from M/G<sub>1</sub> to G<sub>1</sub>, and  $-|\gamma_2\rangle_{RN}$  with the oscillations that start at the transition from S to S/G<sub>2</sub>. The expression patterns of *CLB2* and *CLN3* are uncorrelated with  $|\gamma_2\rangle_{RN}$ . The expression pattern of *CLB2* is anticorrelated with  $|\gamma_1\rangle_{RN}$ , while that of *CLN3* is correlated with  $|\gamma_1\rangle_{RN}$ . We, therefore, associate the eigengene  $|\gamma_1\rangle_{RN}$  with the observed genome-wide effect of *CLN3*, and  $-|\gamma_1\rangle_{RN}$  with the observed genome-wide effect of *CLB2*.

With all 4,579 genes sorted the gene variations of  $|\alpha_1\rangle_{RN}$  and  $|\alpha_2\rangle_{RN}$  fit normalized sine and cosine functions of period  $Z \equiv N - 1 = 4,578$  and initial phase  $\pi/8$ ,  $\sqrt{2/Z} \sin(2\pi z/Z - \pi/8)$ , and  $\sqrt{2/Z} \cos(2\pi z/Z - \pi/8)$ , respectively, where  $z \equiv n - 1$  (Fig. 6 *b* and *c*). The normalized and sorted cell cycle expression approximately fits a traveling wave, varying sinusoidally across both genes and arrays, such that the expression of the  $n$ th gene in the  $m$ th array satisfies  $\langle n|\hat{e}_N|m\rangle \propto 2 \cos(4\pi t/T - 2\pi z/Z - \pi/8)/\sqrt{ZT}$ . The normalized and sorted expression in the *CLB2*- and *CLN3*-overactive arrays approximately fits standing waves, constant across the arrays and varying sinusoidally across genes only, which appear similar to  $-|\alpha_1\rangle_{RN}$  and  $|\alpha_1\rangle_{RN}$ , respectively (Fig. 6*a*).