Table iB. FuzzyK parameter optimization

Trial	Clustering	k per	Total	Total	Hierarc	Hierarchical clusters ^f	
Name ^a	cycles ^b	cycle ^c	K_{exp}^{d}	K_{obs}^{e}	0.7	0.9	
A	1	100	100	37	34	15	
В	2	50	100	45	39	18	
C	3	34	102	43	42	23	
D	1	102	102	45	36	15	
E	3	34	102	78	43	22	
E2	3	26	78	61	44	22	
F	3	25	75	61	44	22	
G	3	34	102	66	41	25	
Н	3	20	60	45	42	22	
I*	3	40	120	91	46	21	
J	3	60	180	109	46	22	
K	3	100	300	122	47	22	
L	3	34	102	91	51	23	

This table summarizes the number of unique centroids identified by each fuzzy k-means clustering trial and the number of hierarchical cluster means that were correlated greater than the indicated cutoff to the centroids identified by fuzzy k-means clustering.

- A. One round of fuzzy clustering initialized with k PCA eigen vectors
- B. Same as A. but perform 2 rounds of clustering on data subset, no weight recalculation
- C. Same as A. but perform 3 rounds of clustering on data subsets, no weight recalculation
- D. Same as A. but seed with the eigen vectors used in C.
- E. Same as C. but recalculate gene and array weights after first round of clustering
- E2. Same as E but k = 26
- F. Same as E2. but do not discard any eigen vectors after seeding
- G. Same as E but use comlete dataset (Dataset A) during entire clustering procedure
- H. K. Same as E. but increase k in each trial
- I. These parameters were used to analyze the yeast gene expression data, as described in the text
- L. Standard k-means ("hard") clustering (see Materials and Methods for details)

^b Number of fuzzy clustering cycles in the trial

^c Maximum number of k centroids per cycle of fuzzy clustering

^d Total number of K centroids expected in the clustering trial

^e Number of unique centroids identified in the clustering trial

f Number of hierarchical cluster means that were correlated >0.7 and >0.9 to the centroids identified in the trial

^a Trial name and description (see also Table i):