

Supporting Information

Peng et al. 10.1073/pnas.0912090107

SI Results

Dependence of Fluorescence in Each Object Type on Probe Concentration.

We assume that the total fluorescence in an image of a fundamental pattern is directly proportional to the probe concentration (over an appropriate range of subsaturating concentrations), and that objects extracted from a cell exposed to a higher concentration of probe will exhibit a higher intensity. That this assumption holds is illustrated in Fig. S2, which shows the approximately linear relationship between fluorescence within each object type and the concentration of probe added. The average correlation coefficient for the relationship between fluorescence in each type and probe concentration was 0.900 for Mitotracker and 0.745 for Lysotracker across all object types.

Concentration Estimation. In many cases, estimating the fraction of probe present in each pattern type is the desired goal. For our test dataset, this fraction is not directly known but must be inferred from the concentrations of probes added (see for example, Fig. 3). However, we can use the fluorescence fraction approach to directly estimate the concentration of probe that was added. For mixture-pattern images, all fundamental patterns can contribute to the total fluorescence F with different fractions in different wells (according to the concentration of the probes). Hence, the total fluorescence can be expressed as the weighted sum of the fluorescence of the fundamental patterns. This can be inverted so that the probe concentrations can be estimated given the

mixture fractions and the total fluorescence (*Materials and Methods*). Fig. S3 shows estimated concentrations of each probe for each well in the test dataset as a function of the actual concentration. The average correlation coefficient for the two patterns combined is 0.93 (0.99 and 0.60 for the separate Mitotracker and Lysotracker patterns, respectively).

Compute Time. The average CPU time per training image to train a mixture model, including calculating object features and model learning, is 84 s; average CPU time to unmix a testing image is 76 s. Each image contains 10–30 cells.

Removal of Nondiscriminative Object Types. Some of the clustered object types have very similar object frequencies or fluorescence fractions across different fundamental patterns. Exclusion of these nondiscriminative types of objects might be expected to boost the unmixing accuracy. We tested this hypothesis by removing object types in the order of their ability to distinguish the two fundamental patterns, training a model using just the retained object types (recording the accuracy on the training data), and calculating correlation coefficients between expected fractions and unmixed fractions. As shown in Fig. S4, the accuracy on the training data and the unmixing accuracy on testing data both remained high even when up to eight object types were removed.

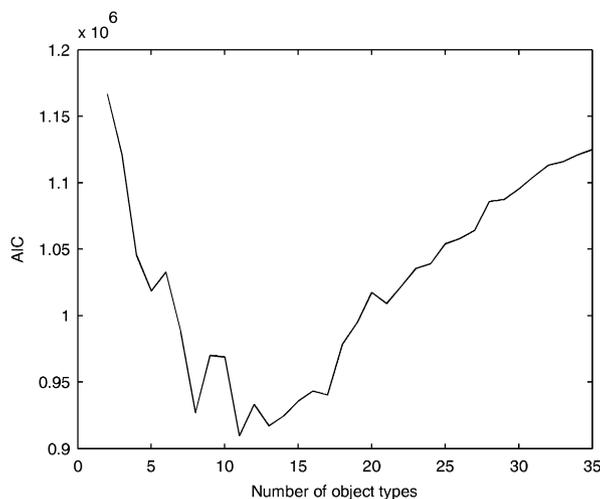


Fig. S1. Learning the number of object types. k -means clustering was carried out for varying numbers of clusters (k) for all objects in the combination of images for U2OS cells receiving only Mitotracker or Lysotracker. The Akaike Information Criterion (AIC) value (which balances the tightness of the clustering against the number of clusters required to achieve it) was calculated for each clustering. The optimal number of clusters (minimum AIC value) is found to be 11.

