Single-cell systems biology by super-resolution imaging and combinatorial labeling

Eric Lubeck1 & Long Cai1,2

Fluorescence microscopy is a powerful quantitative tool for exploring regulatory networks in single cells. However, the number of molecular species that can be measured simultaneously is limited by the spectral overlap between fluorophores. Here we demonstrate a simple but general strategy to drastically increase the capacity for multiplex detection of molecules in single cells by using optical super-resolution microscopy (SRM) and combinatorial labeling. As a proof of principle, we labeled mRNAs with unique combinations of fluorophores using fluorescence in situ hybridization (FISH), and resolved the sequences and combinations of fluorophores with SRM. We measured mRNA levels of 32 genes simultaneously in single Saccharomyces cerevisiae cells. These experiments demonstrate that combinatorial labeling and super-resolution imaging of single cells is a natural approach to bring systems biology into single cells.

The aim of systems biology researchers is to quantitatively understand the interactions between many biological components. To do so, researchers use two distinct technical approaches: genomics1–4, to analyze all genes and proteins in a system simultaneously, and single-cell biology5–10, to follow gene expression in individual cells in their native spatial context. However, these approaches have complementary limitations: genomics approaches report an average over many cells, which masks heterogeneity and spatial complexity in populations, whereas single-cell methods have been limited to the study of a few genes at a time.

To unify the two approaches, we propose to use super-resolution microscopy to bring genomics into single cells. SRM11–14 has been a powerful tool for cell biology and can resolve subcellular structures down to 10–20 nm. This extraordinary resolution can be harnessed for systems biology: for the typical yeast cell of 100 μm3, the 10–20 nm resolution of SRM translates into ~108 independent volume elements (voxels) per cell. These voxels provide enough space to detect large numbers of fluorophore-based barcodes attached to molecular species. Abundances of each molecule can be quantified by counting the number of times the corresponding barcode is observed in the super-resolution image of the cell, in the native cellular and intercellular contexts.

To test the feasibility of this approach, we performed proof-of-principle experiments to detect multiple mRNA species in single S. cerevisiae cells. We based our approach on the single-molecule FISH (smFISH) technique15,16. We used smFISH to combinatorially barcode transcripts, taking advantage of the high labeling specificity of oligonucleotide probes. We used two different barcoding strategies: spatial and spectral. The first relied on the spatial ordering of fluorophores from many small oligonucleotide probes bound to transcripts. The second relied only on the combination of colors used in the probes. Although the first strategy allows higher multiplexing, the second strategy is less technically demanding and more robust. We therefore used spectral barcoding to profile transcripts from 32 stress-responsive genes in single S. cerevisiae cells to test the method. In comparison to previous multiplex FISH approaches that labeled chromosomal loci and transcriptional active sites17,18, our approach directly barcoded single mRNAs.

RESULTS

Spatial coding of single mRNAs

The first strategy we explored for combinatorial labeling directly resolved the spatial ordering of different fluorescent oligonucleotide probes on individual mRNAs. As 20-mer probes are ~7 nm long, multiple regions of the mRNA to which 4–5 probes are hybridized and are separated by 100 nucleotides can in principle be resolved by SRM with its 10–20 nm resolution. By hybridizing probes labeled with different fluorophores to mRNA in a specific pattern, a nanoscopic barcode can be imparted on each transcript and resolved by SRM (Fig. 1).

We targeted the PUN1 mRNA in single S. cerevisiae cells with three sets of oligo probes labeled with different fluorophores. We tiled these probes along the mRNA in a 5′ to 3′ spatially ordered fashion. Hybridized mRNAs appeared as co-localized and diffraction-limited spots (Fig. 1a). We observed that 96 ± 2% (n = 29 molecules, ±s.e.m.) of spots that appeared in Alexa Fluor 594, Cy5 and Cy3 detection channels also co-localized among all three channels, indicating efficient hybridization of the PUN1 probes. We found that different probe sets hybridized with similar efficiency. We quantified the hybridization efficiency of individual FISH probes of the CMK2 probe set to be Hindividual = 67.5 ± 9.1% by photobleaching (Supplementary Fig. 1, n = 173, ±s.e.m.).

1Program in Biochemistry and Molecular Biophysics, California Institute of Technology, Pasadena, California, USA. 2Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California, USA. Correspondence should be addressed to L.C. (lcai@caltech.edu).

RECEIVED 2 MARCH; ACCEPTED 7 MAY; PUBLISHED ONLINE 3 JUNE 2012; DOI:10.1038/NMETH.2069
Barcoding with conventional fluorescence microscopy is only useful in cases where transcript levels are low. When the density of transcripts is not high enough to achieve a sufficient number of co-localized spots, super-resolution imaging of mRNA barcodes can be utilized to resolve the spatial order of fluorophores on mRNAs with higher resolution.

**Figure 1** Spatial ordering of fluorophores on mRNAs can be resolved by Gaussian centroid localization. (a) Fluorescence images of PUN1 probes hybridized in a single budding yeast cell, shown for each channel. (b) Schematic of labeled 25-mer oligonucleotides hybridized to PUN1 mRNA. (c) Reconstructions of centroids in boxed regions 1 and 2 in a after localization by Gaussian fitting and image alignment. (d) Percentage of co-localized PUN1 probes labeled with all three fluorophores that can be reconstructed in a with the intended fluorophore order: blue, magnet, green (n = 28, correct order of fluorophores = 74 ± 8%, s.e.m.). (e) Schematic of the probe set hybridized to GFP mRNA with different order and distances between probe positions compared to fluorophore positions shown in b. (f) Representative Gaussian fitting reconstruction of probe set in e. (g) Distance between the fluorophore positions (d₁ = 27.93 ± 14 nm, d₂ = 56 ± 33 nm; ±s.d.) as determined by Gaussian fitting is proportional to the intramolecular distance between barcode positions (190 bp and 350 bp). (h) Frequency of barcode identification for this probe set (n = 327, correct order = 76 ± 2%, s.e.m.).

The position of labeled probes on mRNA can be determined to a much higher resolution by Gaussian fitting²⁹ of the co-localized spots. We observed the correct spatial order in 74 ± 8% (n = 28, ±s.e.m.) of the mRNA that were hybridized by probes labeled with all three fluorophores after alignment of the imaging channels. Error in detecting correct fluorophore order in the barcode (26%) may result from a combination of factors, including localization error, lack of z-dimension resolution and mRNA secondary structure (Supplementary Note). Barcode read-out fidelity was robust to switching the barcode order (Fig. 1e, h), independent of mRNA identity (Supplementary Figs. 2–4).

**Figure 2** Super-resolution imaging enables combinatorial labeling of individual transcripts. (a) Each probe pair consisted of an activator-labeled and an emitter-labeled oligonucleotide probe hybridized adjacent to each other. (b) Schematic of colors used for barcoding, each of which results from an oligo labeled with an activator (Alexa Fluor 405 (A405), Alexa Fluor 488 (A488) and Cy3) binding adjacent to an oligo labeled with an emitter (Cy5, Alexa Fluor 680 (A680) and Alexa Fluor 750 (A750)). (c–d) PUN1 mRNA (c) and YPS1 mRNA (d) three-position spatial barcode, with the schematic (top) illustrating the order of colors in the barcode, each of which results from four sets of pairs of probes labeled with emitter-activator pairs as indicated in b. In the localization scatterplot (bottom) each dot represents an activation of a fluorophore pair. (e) RCN2 mRNA spectral three-color barcode; probes hybridize throughout the mRNA (top). Localization scatterplots are as in c. Number of integrated peak pixel counts throughout the STORM movie detected for each fluorophore pair is plotted in histograms. Cy5-A405, Cy5-A488, Cy5-Cy3 and A750-Cy3 were detected with 6195, 471, 6881 and 235 integrated counts, respectively. Cy5-A488 (asterisk) was present due to cross-talk from Cy5-Cy3 and was rejected based on the threshold measurements in **Supplementary Figure 10**. Note that the A750-based dye pairs give fewer photons than Cy5 dye pairs but were readily detected with less cross-talk. (f) YLR194C mRNA spectral three-color barcode. Cy5-A488, Cy5-Cy3, A750-Cy3 and A680-Cy3 were detected with 773, 999, 130 and 92 integrated peak counts, respectively. A680-Cy3 (asterisk) was due to cross-talk from Cy5-Cy3 and was rejected.
transcripts is high, diffraction-limited fluorescence spots of the same color will overlap and make barcode readout impossible. Super-resolution imaging using sparse activation of subsets of overlapping fluorophores is essential for high-density multiplex barcoding when signals from multiple transcripts overlap.

To perform super-resolution barcoding, we turned to stochastic optical reconstruction microscopy (STORM) imaging with cyanine dye–based photoswitchable dye pairs. We used singly labeled probes that only generated photoswitchable pairs when an emitter-labeled and an activator-labeled oligonucleotide probe hybridized next to one another on the mRNA (Fig. 2a,b). As both probes are required for the fluorophore to be reactivated, background is dramatically reduced because nonspecifically bound emitter probes cannot reactivate. We used the activator-emitter dye pairs to construct our barcodes, with Cy5, Alexa Fluor 680 and Alexa Fluor 750 as emitters (Supplementary Fig. S5) and Alexa Fluor 405, FITC and Cy3 as activators.

We labeled the PUN1 mRNA (Fig. 2c) using Cy5-based probe pairs. Each probe pair was repeated four times per code position, ensuring a high frequency of hybridized probe-pair formation. With four probe pairs per barcode position, we observed that all three positions of the barcode were present at a rate of 61 ± 8% (n = 85, s.e.m.), consistent with the theoretical hybridization efficiency of Hsuperbar = 75% (Supplementary Note and Fig. 1). Hybridization efficiency can be improved by using a higher redundancy in coding: for example, with six probe pairs per barcode position, the hybridization rate can be increased to 92% for a three-position barcode. Of the barcodes that completely hybridized, we identified the correct barcode order by STORM imaging with 72 ± 10% (n = 50, s.e.m.) of the time. This was consistent with results obtained using non-photoswitchable fluorophores and independent of mRNA species (Fig. 2c,d and Supplementary Figs. 6–8).

For two-dimensional (2D) imaging the spatial barcoding scheme necessitates the linearization of mRNA. To do so, we compressed cells between coverglass slips, extending the mRNA using the lateral force of compression. Such treatment can be readily applied to single-cell organisms and embryos but may destroy spatially complex samples (Supplementary Note).

### Table 1: Comparative advantages of SRM barcoding techniques

| Barcode type | Hybridization pattern | Spatial resolution fidelity | Resolution requirement | Minimum required fluorophore emission | Linearization required | Multiplex scaling |
|--------------|-----------------------|---------------------------|-----------------------|--------------------------------------|------------------------|------------------|
| Spectral     | Distributed           | 100%                      | 100 nm                | ~400 photons                         | No                     | p!(p−n)/n!       |
| Spatial      | Localized             | 74%                       | 20 nm                 | ~3,000 photons                       | Yes                    | p!(p−n)/2        |

*aCalculations use p as the number of fluorophores and n as the number of positions.*

The principal drawback of spatial barcoding is its high-resolution requirement. It is difficult to use dyes that emit fewer photons, such as Alexa Fluors 680 and 750, when high (~20 nm) intramolecular resolution is necessary. We anticipate that the continued development of improved fluorophores, 3D microscopy and nucleic acid self-assembly will enable spatial coding to reach its full potential. In the meantime an alternative strategy, which we call spectral barcoding, is more appropriate for most applications.

### Spectral coding

In spectral barcoding (Fig. 2e,f) the mRNA is identified by the combination of hybridized fluorophore-labeled probes, ignoring spatial order. As long as all the fluorophore pairs are present (as determined by the integrated photon counts over the STORM imaging cycle) and can be identified above the cross-talk tolerances, the barcode assignment can be made confidently even if the total amount of collected photons is low. This comparatively low resolution requirement enables the use of dim fluorophores, as probes need only to be localized to within the ~100 nm area of a single mRNA (Table 1). Probes labeled with the same fluorophore can be distributed throughout an mRNA, making spectral coding more robust to heterogeneities in hybridization and partial degradation of mRNA. In contrast to spatial coding, molecules do not need to be linearized to be faithfully identified, potentially permitting other molecules, such as proteins, to be multiplexed under super-resolution imaging.

### Profiling stress response genes in single yeast cells

As a proof of principle of our technique, we profiled transcripts from 32 known stress-responsive genes (Supplementary Table 1) in single S. cerevisiae cells in response to extracellular calcium stress. We have previously shown that the master transcription factor Crz1 translocates in and out of the nucleus in short (2–3 min) well-defined pulses in the presence of extracellular calcium. These pulses occur stochastically in time and activate target gene expression in a probabilistic fashion. Calcium stress also triggers the Msn2 pathway, a general stress-response regulator that also translocates between the cytoplasm and nucleus in pulses.

For mRNA profiling we selected 14 genes that are regulated by Crz1, 5 general stress-response genes as well as 13 other aging and stress markers. To ensure that we observed the products resulting from individual transcription factor translocation between the
Figure 4 | Single-cell expression profiles of 32 mRNAs. (a,b) Each column corresponds to the expression profile of a single yeast cell. Cells and genes in a were clustered using agglomerative hierarchical clustering on the correlation between species using Ward’s criterion. P values for secondary clusters we calculated by bootstrap resampling and are indicated adjacent to these clusters. Genes can be broadly clustered into two classes, one largely containing genes regulated by both Crz1 and Msn2 (P = 0.09, upper cluster) and one largely containing genes regulated by Crz1 (P = 0.08, lower cluster) (a). Cells are grouped in two distinct clusters, one showing correlations amongst the expression of all genes regulated by Crz1 (P = 0.2, left cluster), the other with large expression correlations among combinatorial genes (P = 0.16, right clusters). Data for additional measured genes are shown in b.

Cells and genes in a were clustered using agglomerative hierarchical clustering on the correlation between species using Ward’s criterion. P values for secondary clusters were calculated by bootstrap resampling and are indicated adjacent to these clusters. Genes can be broadly clustered into two classes, one largely containing genes regulated by both Crz1 and Msn2 (P = 0.09, upper cluster) and one largely containing genes regulated by Crz1 (P = 0.08, lower cluster) (a). Cells are grouped in two distinct clusters, one showing correlations amongst the expression of all genes regulated by Crz1 (P = 0.2, left cluster), the other with large expression correlations among combinatorial genes (P = 0.16, right clusters). Data for additional measured genes are shown in b.

Combinatorial regulation in the calcium stress network

Cells often need to turn on the expression of many genes in response to external stress. We analyzed the expression of several genes in the calcium stress regulon that respond to nuclear localization pulses of the Crz1 and Msn2 transcription factors. In Figure 4, each column represents a single yeast cell grown in medium supplemented with 50 mM CaCl2 and each row represents the mRNA abundances measured for each gene. From these data it is clear that the transcriptional response to calcium varied widely among individual cells.

To characterize the heterogeneity in regulon expression in single yeast cells, we first examined the variability in the abundance of each mRNA species. The distribution of expression levels for most genes included a low basal state and a long-tailed high-expression mode corresponding to bursts of transcription (Supplementary Fig. 13), similar to those observed in previous smFISH experiments in yeast25. As these bursts of transcription are triggered by a transcription-factor translocation pulse in our experiment, we asked whether for our target genes the bursts occurred synchronously. Previous two-color smFISH experiments26 in yeast found that targets of a constitutively active transcription factor in the galactose-response network as well as metabolic genes measured by two different scrambled barcode schemes for 20 genes (Supplementary Table 2). We measured mRNAs with a range of average copy numbers from 1 to 10 copies per cell. Some mRNAs (CMK2 and NPT1) were expressed at levels comparable to or higher than the abundant tubulin subunit genes. We did not measure the 12 remaining genes with low copy numbers (~1 per cell) because they were within the false detection threshold of our measurements. Results from the two spectral barcoding schemes agreed with an R2 = 0.88 (n = 1,871 for the standard code, n = 1,523 for the switched code; Fig. 3c, Supplementary Table 2 and Supplementary Fig. 12) for 19 of the 20 compared genes. As switching scrambled the barcode assignments for genes with different expression levels, the consistent results in this experiment showed that barcodes of high-copy-number genes did not bias the quantification of barcodes of low-copy-number genes through misidentification.

Combinatorial regulation in the calcium stress network

Cells often need to turn on the expression of many genes in response to external stress. We analyzed the expression of several genes in the calcium stress regulon that respond to nuclear localization pulses of the Crz1 and Msn2 transcription factors. In Figure 4, each column represents a single yeast cell grown in medium supplemented with 50 mM CaCl2 and each row represents the mRNA abundances measured for each gene. From these data it is clear that the transcriptional response to calcium varied widely among individual cells.

To characterize the heterogeneity in regulon expression in single yeast cells, we first examined the variability in the abundance of each mRNA species. The distribution of expression levels for most genes included a low basal state and a long-tailed high-expression mode corresponding to bursts of transcription (Supplementary Fig. 13), similar to those observed in previous smFISH experiments in yeast25. As these bursts of transcription are triggered by a transcription-factor translocation pulse in our experiment, we asked whether for our target genes the bursts occurred synchronously. Previous two-color smFISH experiments26 in yeast found that targets of a constitutively active transcription factor in the galactose-response network as well as metabolic genes
were well correlated ($R^2 = 0.5\text{–}0.7$). We found a range of correlation coefficients from $-0.11$ to $0.78$ between pairs of calcium-responsive genes (Supplementary Figs. 14 and 15), indicative of heterogeneous coordination throughout the regulon.

Given this heterogeneous regulon expression profile, we next asked whether there were subgroups of genes that had simultaneous bursts of transcription and whether these correlated groups of genes correspond to different regulatory architectures. We resolved two distinct clusters for the Crz1-responsive genes ($P = 0.09, 0.08$ respectively; Fig. 4 and Supplementary Fig. 16). In the first cluster, Crz1-responsive genes were preferentially clustered based on their coexpression with Msn2-responsive genes. The promoter sequences of many of these genes, such as PUN1, YLR194C, RCN2 and NPT1, contained Msn2-binding sites along with Crz1-binding sites. Genes in the other cluster predominantly contained only Crz1-binding sites, with the exception of YPS1 and PMC1, which also contained Msn2-binding sites. The lower average correlations between Crz1-responsive genes and genes combinatorially regulated by Crz1 and Msn2 (here referred to as ‘combinatorial’ targets) ($R^2 = 0.36, \text{s.d.} = 0.16, n = 62$ cells) compared to those among pure Crz1-responsive targets ($R^2 = 0.55, \text{s.d.} = 0.17, n = 62$) or among the combinatorial targets ($R^2 = 0.52, \text{s.d.} = 0.11, n = 62$ cells) indicated different sources of inputs on the target genes.

As Crz1 and Msn2 nuclear localization pulses are not synchronized in individual cells$^{23}$ at the low calcium (50 mM) levels used in this experiment, we hypothesized that the two cell clusters in the data correspond to cells responding to either a Crz1 pulse or a Msn2 pulse (Fig. 4 and Supplementary Note). We tested this by knocking out the Crz1 and Msn2 pathway separately and visualizing the loss of the associated expression clusters (Supplementary Figs. 17–20). In cells treated with FK506, an inhibitor of the Crz1 phosphotase Calcineurin, the overall mRNA copy number was lower (Fig. 5a). Conversely, in the Msn2− and MSN4− cells, all targets genes should be expressed in response only to Crz1 translocation pulses (Supplementary Note). This is consistent with the observed coordination between pure Crz1 and combinatorial targets (Fig. 5b,c) as well as the disappearance of the cell cluster with only combinatorial targets expressed (Supplementary Fig. 18). This experiment indicates a modulatory effect for Msn2 on Crz1 target genes.

**DISCUSSION**

Super-resolution barcoding can be dramatically scaled up with the technique we presented here. We found that the typical transcript size was about 100 nm and that transcripts were uniformly random distributed in yeast cells (Supplementary Fig. 21) at a mean density of $1.9\pm1.5$ barcodes $\mu$m$^2$ ($n = 2463$, s.d.). There were local regions that contained multiple mRNA barcodes in a diffraction-limited spot (Supplementary Fig. 9 and Supplementary Note) necessitating the use of super-resolution imaging. However, globally, super-resolution barcodes of the 32 profiled genes occupy less than 2% of the 2D super-resolution space available in a yeast cell, leaving a large amount of space open for additional multiplexing. With one of the additional available emitters$^{27,12}$, $C_2 = 792$ genes can be spectrally coded.

Spectral barcoding has a reconstruction fidelity dependent only on the binding of fluorophore-labeled probes, does not require linearization of the coded molecule, has a low photon requirement and allows the molecule of interest to be labeled in a distributed pattern to increase robustness. These advantages make it an excellent technique for barcoding mRNA and other molecules with unknown structures with the existing SRM dye palette. In comparison, spatial barcoding has a localized hybridization pattern, a lower reconstruction fidelity owing to the much higher photon requirements and the tertiary structure of the barcoded molecule. It has the advantage of scaling and may be more applicable to chromosome and splice isoform barcoding.

Scaling up beyond 1,000 genes will require the implementation of 3D SRM$^{28}$ to improve the axial resolution by a factor of 20 and the synthesis of an additional far-infrared fluorophore as an emitter (to spectrally code $C_{18,532} = 18,532$ genes). At this higher barcode density, more sophisticated computational algorithms will also be necessary to identify barcodes in an intelligent and automatic fashion (Supplementary Note).

Super-resolution barcoding provides several advantages over existing transcriptional profiling techniques. First, direct imaging of the sample preserves the spatial information both within cells and among cells. With the application of light-sheet microscopy$^{29}$, this technique can be extended to analyze optically thick samples without photobleaching associated with z-dimension sectioning in epi-fluorescence microscopy. This advantage makes it a powerful tool in studying signaling in heterogeneous systems such as microbial ecosystems, tissue and embryos, where interactions among different cellular populations have an essential role in cellular decisions. Second, because of the single-molecule and *in situ* nature of the technique, the method is quantitative and avoids intrinsic bias in RNA extraction and conversion to cDNA. Third, many cells can be imaged simultaneously under...
a microscope quickly, and throughput can be scaled up without considerable costs, compared to the high cost and long waiting time for sequencing nucleic acid from single cells. After the initial cost of the probe-set synthesis, the probe set can be hybridized many thousands of times in wild-type and mutant organisms. Super-resolution barcoding even at its current throughput can be a useful follow-up to existing high-throughput transcriptomics techniques by allowing genes of interest to be monitored with single-cell resolution in spatially complex samples.

This labeling scheme can be applied to many types of molecules in situ. It is a short leap to consider combinatorial labeling of chromosomes and proteins for single-cell proteomics and chromatin immunoprecipitation experiments. We hypothesize that for many types of biochemical techniques, such as microarrays, there is an equivalent in situ single-cell experiment possible through super-resolution barcoding, removing the need for spatial separation traditionally performed using gels or dilution on a chip.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**ACKNOWLEDGMENTS**

We credit B. Wold with discussions that led to this work. We thank M. Elowitz for lending space and equipment in his laboratory, T. Zhilyentayev, H.-O. Li and X. Wang for assistance with experiments, A. Raj for technical assistance with FISH, X.W. Zhuang and her group for STORM, and A. Eldar, S. Fraser, G.W. Li, J. Levine and J. Locke for discussion and reading of the manuscript. This work was supported by a Beckman Institute seed grant and a US National Institutes of Health New Innovator Award 1DP2OD008530.

**AUTHOR CONTRIBUTIONS**

E.L. and L.C. performed the experiments, carried out the analysis and wrote the manuscript. L.C. conceived the idea and designed the experiments.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

Published online at [http://www.nature.com/doifinder/10.1038/nmeth.2069](http://www.nature.com/doifinder/10.1038/nmeth.2069). Reprints and permissions information is available online at [http://www.nature.com/reprints/index.html](http://www.nature.com/reprints/index.html).

1. Johnson, D.S., Mortazavi, A., Myers, R.M. & Wold, B. Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**, 1497–1502 (2007).
2. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods* **5**, 621–628 (2008).
3. Nagalakshmi, U. et al. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344–1349 (2008).
4. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470 (1995).
5. Elowitz, M.B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
6. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. & Prasher, D. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805 (1994).
7. Golding, I., Paulsson, J., Zawilski, S.M. & Cox, E.C. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025–1036 (2005).
8. Cai, L., Friedman, N. & Xie, X.S. Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358–362 (2006).
9. Yu, J., Xiao, J., Ren, X., Lao, K. & Xie, X.S. Probing gene expression in live cells, one protein molecule at a time. *Science* **311**, 1600–1603 (2006).
10. Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309 (2006).
11. Hell, S.W. Far-field optical nanoscopy. *Science* **316**, 1153–1158 (2007).
12. Betzig, E. et al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).
13. Rust, M.J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–795 (2006).
14. Hess, S.T., Girirajan, T.P.K. & Mason, M.D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.* **91**, 4258–4272 (2006).
15. Femino, A.M. Visualization of single RNA transcripts in situ. *Science* **280**, 585–590 (1998).
16. Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879 (2008).
17. Lowenstein, M.G. Long-range interphase chromosome organization in *Drosophila*: a study using color barcoded fluorescence in situ hybridization and structural clustering analysis. *Mol. Biol. Cell* **15**, 5678–5692 (2004).
18. Levsky, J.M., Shenoy, S.M., Pezo, R.C. & Singer, R.H. Single-cell gene expression profiling. *Science* **297**, 836–840 (2002).
19. Thompson, R.E., Larson, D.R. & Webb, W.W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82**, 2775–2783 (2002).
20. Yildiz, A. et al. Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* **300**, 2061–2065 (2003).
21. Bates, M., Dempsey, G.T., Chen, K.H. & Zhuang, X. Multicolor super-resolution fluorescence imaging via multi parameter detection. *ChemPhysChem* **13**, 99–107 (2012).
22. Barish, R.D., Schulman, R., Rothemund, P.W.K. & Winfree, E. An information-bearing seed for nucleating algorithmic self-assembly. *Proc. Natl. Acad. Sci. USA* **106**, 6054–6059 (2009).
23. Cai, L., Dalal, C.K. & Elowitz, M.B. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature* **455**, 485–490 (2008).
24. Gasch, A.P. et al. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11**, 4241–4257 (2000).
25. Zenklusen, D., Larson, D.R. & Singer, R.H. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* **15**, 1263–1271 (2008).
26. Gandhi, S.J., Zenklusen, D., Lionnet, T. & Singer, R.H. Transcription of functionally related constitutive genes is not coordinated. *Nat. Struct. Mol. Biol.* **18**, 27–31 (2011).
27. Dempsey, G.T., Vaughan, J.C., Chen, K.H., Bates, M. & Zhuang, X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat. Methods* **8**, 1027–1036 (2011).
28. Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **319**, 810–813 (2008).
29. Cella Zanacchi, F. et al. Live-cell 3D super-resolution imaging in thick biological samples. *Nat. Methods* **8**, 1047–1049 (2011).
30. Agnew, H.D. et al. Iterative in situ click chemistry creates antibody-like protein-capture agents. *Angew. Chem. Int. Ed. Engl.* **48**, 4944–4948 (2009).
ONLINE METHODS

Probe design, purification and hybridization. We designed 25-mer oligonucleotide probes (Biosearch) to have the same melting temperature whenever possible (Supplementary Table 3). FISH probes were designed with 2-base-pair spacing between probes to allow efficient reactivation of the pair dyes, often leading to varying melting temperature of probes. Alexa Fluor 405, Alexa Fluor 488 and Cy3 were used as the activators, and Cy5, Alexa Fluor 680 and Alexa Fluor 750 (Invitrogen) as the switchable dye. Probes were labeled and purified following the protocol in ref. 16. Yeast cells were grown in yeast extract peptone dextrose (YPD) medium with 50 mM CaCl₂ and fixed in log growth phase. Cells were also treated with 0.1% NaBH₄ for 30 min before the ethanol-permeabilization step. We found the NaBH₄ treatment decreased the autofluorescence background of fixed yeast cells. Cells were stored at −20 °C in Eppendorf tubes and aliquoted out for hybridization experiments. Cells were hybridized with the probes overnight at room temperature in 20% formaldehyde and 10% dextran sulfate. For all smFISH experiments, 12 probes (Supplementary Table 3) were hybridized to mRNA. For super-resolution barcoding, sets of 4–7 probe pairs were used for Cy5, A680 and A750, respectively. After hybridization, cells were washed in 10% formaldehyde and 0.2× SSC solution 3 times and imaged.

Imaging. For conventional epi-fluorescence microscopy, images were acquired on an Olympus IX81 with a 100× sapphire objective with laser illumination at 532 nm, 594 nm and 640 nm. Images were acquired with Micromanagement software and an Andor iKon-M DU934 BV charge-coupled device (CCD). Conventional fluorescence images were acquired in three different fluorescence channels (Semrock zero-line filters). The centroids of the diffraction-limited FISH spots were calculated in each channel, and the images were aligned by center of mass alignment of co-localized fluorescence spots between channels. This was sufficient for alignment without correcting for rotation and dilution.

Super-resolution imaging was performed on a Nikon TI-Eclipse microscope with PFS autofocus lock. Standard super-resolution imaging buffers with glucose oxidase and β-mercaptoethanol were used. The imaging lasers, a 640-nm laser along with a 30-nm 691-nm and a 30-nm 730-nm laser (Coherent Lasers) were brought to the sample through a Nikon Apo total internal reflection (TIR) 100× objective. The 405-nm, 473-nm and 556-nm lasers were used as activation lasers, and imaging automation was controlled by Micromanagement software. Two illumination pathways were used on the microscope: an imaging pathway containing the 640-nm, 691-nm and 730-nm lasers, and an activation pathway consisting of the remaining lasers. The activation pathway contained an automated filter wheel (Thorlabs) with neutral density (ND) filters 0.3, 0.6, 1 and 1.6 and an Uniblitz shutter, enabling automatic control of activation power by a custom written Micromanagement script. No excitation filters were used in the microscope, enabling both activation and imaging wavelengths to reach the sample. Exposure times for all STORM movies were 150 ms for Cy5, 400 ms for A680 and 300 ms for A750.

For all STORM experiments, the imaging routine was designed to minimize cross-talk and photobleaching. To reduce activator cross-talk, every imaging routine sequentially used activation lasers from 556 nm, to 473 nm, to 405 nm. As the absorption spectra of the activator dyes slightly overlap, this allowed us to minimize cross-activation by preferentially bleaching the overlapping dye before the routine reached the adjacent activation laser.

For the spatial coding images, samples were first imaged with only the 640-nm laser for 100 frames to switch off Cy5 and to determine the nonspecific blinking rate. Then, 100 frames were acquired in each activation channel by illumination with the activation and the imaging lasers. This reduced the cross-talk among the different activation channels.

For the spectral coding images, the samples were first bleached in all imaging channels for six frames. Samples were imaged in order of the activators, starting at 556 nm to 473 nm and then 405 nm. For the 556-nm laser activation, the microscope first imaged the 730-nm laser followed by either the 640-nm or 691-nm laser for 48 cycles. This pattern minimized the Cy5 and A680 signal loss due to bleaching of both dyes by the 640-nm and 691-nm lasers. For the remaining activation lasers, the microscope cyclically imaged the 730-nm laser followed by the 640-nm laser. Samples were only illuminated with activation light in the A750 channel. This imaging pattern minimized photobleaching of A750 while still allowing us to observe adequate activation of the A680 and Cy5 probe pairs. Throughout the imaging routine, for every four cycles of the imaging lasers, two cycles without the activator lasers were acquired. This enabled us to rule out much of the false positive and nonspecific blinking events in the images.

Activation powers were selected to maximize activation rate while avoiding cross-talk among the activators. After the imaging cycle with all emitters, the 640-nm laser was continuously used to image Cy5 for another 30 frames of specific activation with an ND filter added to the pathway to bring the activation laser power to 50% of the original activation. The acquisition sequence is illustrated in Supplementary Figure 11.

The activation lasers were controlled by an Arudino microcontroller board and a servo motor shutter. Fluorescent beads (Invitrogen F-8810) were used as fiducial markers to correct for stage drifts. The microscope stages (Prior and ASI) were automated and controlled by acquisition software to enable multiposition imaging.

Analysis. Images from diffraction-limited and super-resolution experiments were analyzed with a custom written Mathematica script, available upon request. In the analysis, the beads were first aligned to determine the stage drifts. Beads emit on the order of 50,000 photons per image, and could be localized to a few nanometers. Beads close to the cells were eliminated from analysis as the switching of fluorophores in cells can disrupt the bead alignment. Then, fluorophores were selected from each image by intensity thresholding (600 photons for Cy5, 400 photons for A680 and 250 photons for A750), and their centroid was calculated by Gaussian fitting. We did not reject activation events that involved multiple fluorophores in the analysis because they came predominantly from single clusters and allowed us to preserve more of the photons collected for the analysis. In principle, multiple activation frames can be compared to neighboring frames to extract localization information.

To determine which of the barcode colors are present instead of nonspecific activation, we compared the number of activation events occurring in the specific activation channel versus the activations observed in frames with no activation. If the frequency of
nonspecific activation events exceeded that of specific activation events, then that channel was rejected. As several of the activators and emitter channels can cross-talk into other channels, we quantified the cross-talk ratio and rejected activation events if they fell below our thresholds (Supplementary Note). Switching events that spatially clustered together were grouped to display the resolved barcode. For spectral barcodes, activations were clustered together on a 184-nm grid for simplicity because most RNAs were contained within a diameter of 100 nm. Activation events near grid vertices were assigned to a neighboring region containing localizations of the same fluorophore pair. For three-color spatial barcodes, the center color was determined by finding the position that was not one of the two localized colors separated by the longest distance. Cell positions were determined by manual segmentation. Barcodes were collected and tabulated for each single cell. Cross-correlation was calculated using the correlation function in Mathematica, and the standard errors were calculated from resampling the data 100 times.

Quantitative PCR measurements. For real-time PCR experiments, *S. cerevisiae* were grown in YPD medium and were induced with 50 mM CaCl$_2$ for 2 h. The cells were then spheroplasted with zymolyase, RNase inhibitor and β-mercaptoethanol. The RNA was isolated using the Trizol Plus RNA Purification System (Invitrogen). The extract was treated with recombinant DNase I (Roche) to digest genomic DNA. Reverse transcription of the RNA was performed using the iScript Reverse Transcription Supermix (Bio-Rad). Aliquots of the resulting cDNA were diluted 1:10 and 1:100 with water. QPCR was performed using Maxima Sybr Green and Flourescein qPCR Master Mix (Fermentas) on a Bio-Rad CFX machine. QPCR was performed on each gene with the undiluted cDNA, 1:10 dilution cDNA and 1:100 dilution cDNA, with each dilution repeated in triplicate (Fig. 3). The PCR efficiency for each gene was calculated from a linear fit of the three dilution points. RNA abundance was determined from the amplification efficiency. Primer sequences are listed in Supplementary Table 4.