Supporting Online Material

Materials and Methods

Construction of library strains

The library strains are created via λ-RED recombination using a universal primer (Fig. S10) as follows. Strains from an existing chromosomal fusion library (*S1*) were grown up to $O.D_{600}$ ~0.6 in LB from an overnight culture. The cultures were transferred to Mini-Tubes (Bel-Art F37857-0000) for a 15-minute, 42°C heatshock, and then chilled on ice. The cells from each strain were centrifuged at 3,800 rpm × 10 minutes at 4°C (Sorvall Super T21, ST-H750 rotor), to remove the supernatant and then washed twice with cold water (Millipore) before they were concentrated and transferred to 1 mm gap electroporation cuvettes (VWR International). The venus-chloramphenicol (CAM) resistence cassette was prepared with Platinum PCR Supermix (Invitrogen) using primers VenF-SpaF (aactactgctagcgagaatttgtattttcagggtgagctcagcaagggcgaggagctgttcac) and CamR-KanR (ggcgtcgcttggtcggtcatttcgaaccccagagtcccgcctgccactcatcgcagtactgttgt) (Integrated DNA Technologies) and the PCR product was cleaned up using DNA Clean & Concentrator-5 columns (ZymoResearch). Electroporation was performed at 1600V (BTX ECM399). Cells were allowed to recover for at least 3 hours before plating with chloramphenicol for selection.

Colonies were screened for insertion by PCR. The PCR product from positive colonies was sequenced to confirm correct insertion using the Biopolymer Facility at Harvard Medical School.

Microfabrication

Photolithography and softlithography techniques were used to produce the microfluidic platform (*S2*). Poly-dimethylsiloxane (PDMS), a low-cost, optically-transparent silicon elastomer, was selected as the matrix of microfluidic chip. The chips were made by molding PDMS on a microfabricated silicon wafer. To prepare the mold, we designed a microfluidic pattern on AutoCAD 2004 (Autodesk Inc.) and output it into a photomask film using a commercial photoploting service (CAD/Art Services, Inc.) with a resolution of 20,000 dpi. We spin-coated an UV-curable epoxy (SU8-2025, Micro-Chem) with a 25 μm thick on a test-grade silicon wafer (University wafer). The designed microfluidic pattern was developed by exposing UV-light to the wafer through the photomask and immersing it in a developer solution. The PDMS was molded on the fabricated wafer by curing at 60°C in 45-60 minutes. ϕ0.75 mm holes were punched through the replicated PDMS sheet at the inlet/outlet positions. A coverslip (0.17 mm thick, 48 × 60 mm, Brain research laboratories) and the PDMS sheet were treated by an oxygen plasma cleaner and were bonded each other.

Microscopy

Single molecule fluorescent experiments were done on an inverted microscope (IX71, Olympus Americas, Inc.). Epi-illumination was provided by an Ar laser at 514 nm (Innova 300, Coherent) for Venus excitation and a fiber laser at 580 nm (VFL-P-Series, MPB Communications Inc.) for mCherry and Atto 594 excitation. Phase contrast illumination by a halogen lamp was also provided to identify the cell position and shape. Images were taken on an EM-CCD camera (Cascade 512B, Photometrics) with a 100 msec exposure through a 100× phase-contrast objective lens (NA = 1.35, Olympus). Samples were placed on a motorized 3D translational

stage (MS2000, Applied Scientific Instrumentation). For real-time experiments, a temperature controller was attached to the sample (FCS2, Bioptechs). The light source was switched by mechanical shutters (VMM-D3, Uniblitz) and a dichroic mirror wheel. Automatic measurements were done by Metamorph software (Molecular Devices), which synchronizes the stage scanning, shutter control and camera acquisition.

Cell preparation

Cells were grown in LB media with 20 μg/ml Chloramphenicol and subsequently were inoculated into M9 media supplemented with 0.4 % glucose, amino acids and vitamin with 1:400 dilution. The cells were incubated at 30°C for 11-12 hours and were grown to OD_{600} = 0.1-0.5. To check that 4-5 cell divisions in M9 is sufficient to generate cells in a steady state, we grew 20 randomly selected strains for >24 hours in M9 to confirm that the measured abundances were the same. Deep 2 ml 96-well plates (VWR) were used to culture many samples at once. During culturing, the plates were tightly capped and were placed on the side in a shaker to provide sufficient aeration. The doubling time was 150 minutes. Before imaging the cells were spun down in a tabletop centrifuge (Sorvall super T21, Kendro Laboratory Products) for 10 minutes at 3,800 rpm and washed once with 0.85% NaCl solution.

Chip preparation

The microfluidic chip we designed integrates 96 independent pathways in parallel and can hold 96 kinds of cell samples in parallel channels on a single coverslip (Fig. S11). The measurements were automatically performed by scanning the microfluidic chip under a microscope capable of single molecule detection with a PC-controlled 3D translational stage. The size of the pathway is 150 μm (width) × 10 mm (length) × 25 μm (height). Samples with fewer than 500 cells, or that include many long unhealthy cells or lysed cells, were discarded and re-measured later. All samples were measured at least twice on two different days. The microfluidic chips used only once for the experiment.

A multi-channel pipette (12 channels, Rainin) was used to inject solution into the channels. For this purpose, the spacing between every 4 channel inlets was designed to match the spacing between pipette tips. The elasticity of PDMS works sufficiently to seal between the ϕ0.75 mm inlet and disposable plastic tips to inject the solutions. To immobilize bacterial cells on the microchannels, the channels were pre-coated with 0.1 % poly-L-lysine. After pre-coating, cells were injected into the channels and were incubated for more than 45 minutes for stable binding to the channel surface. Floating cells in the channels were washed out by injecting 0.85 % NaCl solution, resulting in a single cell layer on the coverslip surface.

Scanning measurement

An automated scanning measurement was performed as shown in Fig. S12. A combination of a phase contrast image and fluorescent image was taken at different positions along the channel profile. Typically, 10 sets of image were acquired per channel, resulting in 500-20,000 cells observation. To prevent over-saturated images, the camera gain is decreased for subsequent images if a very bright pixel was observed in the first image. The relationship between camera gain and obtained pixel value has been calibrated in advance. The elapsed time was 25 seconds per one channel. The recorded images were saved in 16-bit TIFF format.

As controls for microscopy measurement, fluorescein solution and plain NaCl solution in separate channels were imaged respectively. The former was used to compensate for the heterogeneity of laser illumination in the image field. The latter was used to subtract the dark noise count due to the EMCCD property and the autofluorescence background from the microchannel and immersion oil.

Image processing

Automated image analysis was done by LabVIEW software (National instruments). Obtained fluorescence images were subtracted with the background image and were flattened using the fluorescein solution image (Fig. S13). Phase contrast images were processed through a closed filter and a sharpen convolution filter and were thresholded to create binary images. The binary images were segregated into particles, which were filtered by an area, minor and major caliper lengths (Fig. S14A), in order to exclude overlapped cells, long unhealthy cells, cell debris and the other unexpected objects from the analysis. The integrated fluorescence intensity within the entire cell area was obtained for each cell and was normalized by cell volume: (2/3)*(cell area)*(cell minor caliper length).

Determination of protein copy number distribution

The auto-fluorescence distribution, obtained from cells lacking fluorophore, was deconvoluted from the resulting fluorescence distribution to provide the true protein fluorescence distribution (Fig. S14C). The bin number of the histogram was set as the maximum bin number where all values of bins in the deconvoluted histogram were positive. Because the standard deviation in single molecule fluorescence intensity is small (65%) (Fig. S1A), the YFP intensity distribution was not deconvolved from the measured distribution. We expect that the contribution of the YFP fluorescence distribution to be negligible to the final result. We determined mean (μ), standard deviation (σ), skewness, kurtosis, an inverse of noise ($\mu^2/\sigma^2 \equiv \alpha$) and Fano factor ($\sigma^2/\mu \equiv \beta$) from the deconvoluted histogram. The errors of these parameters were provided by a bootstrap resampling method with 1,000 samples. All the distributions were fitted with both a gamma distribution and a sum of two gamma distributions. A least chi-square fitting was done based on error values for individual bins of the histogram estimated by a bootstrap. To check for bimodality, a likelihood-ratio test was performed using the difference in the chi-square residual values between one gamma and two gamma distributions. Data of the strains that have bimodal distributions were re-measured, and were discarded if they were not reproducible. Bimodal distributions are mainly observed in samples that have a mixture of healthy and unhealthy cells.

We divided the obtained fluorescence count by the single molecule fluorescence, measured as described in Supplementary data section 1, to obtain copy numbers.

Characterization of protein localization

The characterized proteins were divided into (i) membrane/cytoplasm localization and (ii) punctate localization. (i) Membrane/cytoplasm localization was characterized by comparing the average of fluorescence values at the contour of cell and at the inside of cell. We calculated the ratio of fluorescence detected on the edge compared to inside of the cell (E/I) and obtained their mean and standard deviation. (ii) Punctate localization was characterized by counting spots in cells. We applied an open filter to the fluorescence image and calculate the difference between images before and after filtering (Fig. S15), which highlights only punctate localization. The

differential image was transformed to a binary image and filtered, and the number of particles within the cell area was counted for each cell.

Parameters obtained from the automated image analysis

From the image analysis, we determined the following parameters.

(1) Number of analyzed cells
(2) Mean, SD, skewness and kurtosis of single-cell fluorescence distribution
(3) $\alpha = \text{Mean}^2/\text{SD}^2$ and $\beta = \text{SD}^2/\text{Mean}$
(4) The error value of parameters in (2) and (3) obtained by a bootstrap
(5) A and B value and their errors in fitting a i) unimodal or ii) bimodal gamma distribution to the histogram
    i) discrete unimodal gamma distribution:

$$f_1(x_i, A, B) = \frac{x_i^{A-1} \exp\left(-\frac{x_i}{B}\right)}{\sum_j x_j^{A-1} \exp\left(-\frac{x_j}{B}\right)}$$

    ii) discrete bimodal gamma distribution:

$$f_2(x_i, A_1, B_1, A_2, B_2, p) = p \times f_1(x_i, A_1, B_1) + (1-p) \times f_1(x_i, A_2, B_2)$$

(6) Chi square values and *p*-values for the fit of a unimodal and bimodal distribution to the histogram
(7) The 2× logarithm of likelihood ratio of bimodal to unimodal distribution (= difference in chi square values) and *p*-value for unimodality
(8) Result of bimodality test with a type I error rate of 0.1% ($T \geq 12.8$)
(9) Maximum bin number where all values of bins in the deconvoluted histogram are positive (maximum: 45)
(10) Ratio of fluorescence from the edge to inside of the cell.
(11) The number of identified fluorescent spots per cell
(12) Cell area, width, length, and volume (= 2/3 × area × width)

Z-score calculation

To characterize the bias of parameters for each functional category, we determined the *Z*-score to describe the statistical probability of selecting the datasets using the category as a criteria. The *Z*-score is defined as:

$$Z = \frac{X - \mu}{\sigma},$$

where X is the average rank of a parameter for a data subset corresponding to genes belonging in a certain functional category. μ and σ are the mean and standard deviation of the average ranks simulated by a 1000 times bootstrap resampling where an number of random subsets of genes equal to the size of the entire population were selected from the population. Z scores of more than 3 (indicated by red) represent a significantly larger than average quantity compared with the whole genome distribution with >99.9% confidence, and Z scores of less than -3 (indicated by blue) represent a significantly smaller than average quantity.

We also used the *Z*-score to characterize the deviation of the rank of parameters within a subset of genes to examine similarity of parameters in a category. Here μ and σ were the mean and standard deviation of the rank deviation obtained by a bootstrap.

In addition, the correlation between parameters was characterized by the *Z*-score, where *X* is the average of a correlation coefficient for a data pair subset in a category. μ and σ are simulated by a 1000 times bootstrap where random subsets of data pairs were selected.

References for data used for Z-score calculations

Biological Property	**Reference**
Gene annotation and gene coding information	ECOCYC (http://ecocyc.org/links.html)
Mass spectroscopy and 2D gel data	Lu, P. et al., Nature Biotechnology, 2007 (*S4*)
Operon	Regulon DB (http://regulondb.ccg.unam.mx/html/Data_Sets.jsp)
CAI index	Sharp, P. M., Nucleic Acids Res., 1987 (*S5*)
Gene essentiality	Profiling of *E. coli* chromosome (http://www.shigen.nig.ac.jp/ecoli/pec/index.jsp)
TF gene interactions	Regulon DB (http://regulondb.ccg.unam.mx/html/Data_Sets.jsp)
PPI information	Butland, G. et al., Nature, 2005 (*S1*)
DNA and RNA folding energy	UNAFold (http://dinamelt.bioinfo.rpi.edu/unafold/)
Conserved protein number	Comprehensive microbial resource (http://cmr.jcvi.org/)

***p*-value of essential genes copy number**

We confirmed that, given the distribution of a mean abundance for all genes in Figure 2A, the observation of 108 out of 121 essential genes having a mean abundance greater than 10 molecules per cells is statistically significant compared to what is expected for 121 random genes. We estimated the *p*-value to be $< 10^{-11}$ from the chi-square test ($\chi^2(1) = 50$).

Real-time experiment of protein expression

Real-time observation of library strains were done as described previously (*S6*). Cells were centrifuged and placed between a 3% agarose gel pad and a glass coverslip. For low copy strains, the gel pad was made with M9 media supplemented with 0.4% glucose, 0.05 % casamino acid, 0.15 μg/ml biotin and 1.5 μM thiamine. For high copy strains, the gel pad was made with M9 media supplemented with 0.4% glucose, amino acids, vitamin and 0.2% casamino acid. The gel was set in an imaging chamber (FCS2, Bioptechs) and was kept at 30°C during observation. Image acquisition was done every 5 minutes for 5-9 hours. For low copy strains, a higher laser power (~600 W/cm^2) was used to image single molecules where 50 ms exposure followed immediately by 2 additional images to completely photobleach existing fluorophores. In contrast, for high copy strains, a lower laser power (~1.3 W/cm^2) was used for the fluorescence excitation to prevent photobleaching. To exclude the frame-to-frame variation of fluorescence intensity due

to the inconstancy of auto-focusing, shutter timing and a stage drift, we normalized the fluorescence values of cells by their average for each frame.

Probe design for fluorescence in situ hybridization (FISH)

The FISH probe is comprised of a 20-mer oligodeoxynucleotide (Venus495r, 5'-TCCTCGATGTTGTGGCGGAT -3') with a covalently linked a dye molecule (Atto 594) on the 5' end. The oligonucleotide sequence was chosen such that it is the reverse complement of a region on the *yfp* mRNA that has the least frequency of secondary structures. Atto 594 (Atto-tec GmbH) was chosen for its brightness, photostability, and the reduced nonspecific binding during in situ hybridization (data not shown). The dye is linked to the oligonucleotide via NHS ester reaction, followed by RNase-free HPLC purification (custom made by Sigma-Aldrich).

Cell preparation for FISH

Library strains were inoculated and grown under the same condition stated above. At ~ 0.3 OD, 950 μl of each cultured strain in the deep 2 ml 96-well plate was rapidly mixed with 950 μl of pre-chilled 2X fixation solution (7.4% formaldehyde and 2X RNase-free PBS in DEPC-treated water (Ambion)). The mixture was shaked vigorously briefly and incubated on ice for 15 min. The cells were then pelleted with a tabletop centrifuge (Sorvall Super T21) for 10 min at 3,800 rpm, and washed twice with ice-cold RNase-free PBS solution (Ambion). After the wash, the cells were resuspended in 70% ethanol and incubated at RT for 1 hour. Finally, the cells were spun down and washed with the Wash Buffer (25% formamide (Ambion) and 2X SSC (Ambion) in RNase-free water (Ambion)). The cells were resuspended in ~20 μl of the Wash Buffer.

FISH probe hybridization

The hybridization protocol was originally adapted from Femino et al (*S7*), Raj et al (*S8*), and Zong et al (submitted). The condition was further optimized for the YFP probe (Venus495r) in *E. coli*. We used only one single oligonucleotide probe with only one dye molecule. This strategy offers an advantage of counting overlapping spots using either intensity or photobleaching steps. This is important for us when measuring the mRNA-protein correlation, because it requires the knowledge of the absolute copy number in each cell. It is still advantageous to use multiple probes in most cases, especially in bigger cells where single fluorophore cannot be easily detected.

The Hybridization Buffer consists of 25% formamide, 2X SSC, 10% dextran sulfate (Sigma Aldrich), 0.2 mg/ml BSA (New England Biolabs), 2 mM ribonucleoside vanadyl complex (Sigma Aldrich), and 0.1% *E. coli* tRNA (Sigma Aldrich). 10 μl of the cells prepared in the previous step was mixed with 50 μl Hybridization Buffer and 2.5 μl of 30 nM the FISH probes (Atto594-Venus495r) dissolved in RNase-free water and 0.2 mg/ml BSA. The hybridization mixture was incubated in a 96-well plate at 30°C for 9 hours.

Following the 30°C incubation, the cells were washed with the Wash Buffer twice, incubated in Wash Buffer at 30°C for 1 hour, and then washed once again with the Wash Buffer and with PBS once. The cells were resuspended in 10 μl PBS.

mRNA imaging sample preparation

The hybridized and washed cells were immediately applied to poly-D-lysine (Sigma Aldrich) coated glass coverslips. The coverslips (25 mm × 75 mm, Belco) were cleaned and prepared as follows: 30 min sonication in 1 M potassium hydroxide, 30 min sonication in purified water (Millipore), 1 min blow dry by Nitrogen gas, 10 min in plasma sterilizer, 40 min in 0.03 % poly-D-lysine (Sigma Aldrich), 1 min rinse in deionized water, and 1 min blow dry by Nitrogen gas. Each coverslip is then adhered to a 16-well silicone gasket (FlexWells, Grace Bio-labs).
The cells were allowed to adhere to poly-D-lysine coated coverslip for 30 min while covered to prevent evaporation. Each well is then washed extensively with RNase-free PBS. After the final wash, each well is filled with 140 mM 2-mercaptoethanol (Sigma Aldrich) in PBS, and covered with a cleaned glass slide.

Microscopy for FISH experiment
Single molecule imaging was performed as described previously (*S6, S9, S10*). An Olympus IX71 microscope with a 100X NA 1.35 phase-contrast objective and an EMCCD (Cascade 512B, Photometrics) is used in this study. For Atto 594 imaging, a 580 nm fiber laser (MPB Communications Inc.) was used. An achromatic quarter waveplate (Thorlabs) was used to create near-circular polarization at the objective imaging plane. The fluorescence filter set includes an excitation filter (HQ575/50X, Chroma), a dichroic mirror (z594rdc, Chroma), and an emission filter (D635/55M, Chroma). The laser intensity at the image plane is ~100 W/cm^2. Each image was recorded in 1 s. For YFP imaging, a 514 nm laser (Innova 300, Coherent) was used. The filter set includes an excitation filter (D510/20X, Chroma), a 525 nm longpass dichroic mirror, and an emission filter (HQ545/30M, Chroma). The laser intensity at the image plane is ~100 W/cm^2. Each image was recorded in 100 ms. No statistically significant crosstalk was observed between the YFP channel and the Atto 594 channel.

Automated image acquisition using Metamorph (Molecular Devices) allows sequential imaging of each 16-well coverslip, as described earlier. YFP and FISH images were recorded for 20-30 field-of-views for each strain, with an average of ~1000 cells total. Images without laser excitation were recorded to serve as offsets of the actual fluorescence images. Images of dilute dye solutions were recorded to correct for the slight inhomogeneity of the field-of-view.

Image analysis for mRNA counting
The phase-contrast images and the YFP images were analyzed as described in the earlier section. The FISH images were subtracted with camera offset, and the field-of-view was flattened using the image of dilute dye solution described in the previous section. Fluorescence spots corresponding to localized mRNA were identified with a peak-searching algorithm written in Matlab (The MathWorks). The algorithm searches for pixels that have both (i) pixel intensity above a pre-defined threshold and (ii) image curvature above a pre-defined threshold. The thresholds are adjusted so that all fluorescent spots are identified via visual inspection, and that all identified peaks correspond to actual spots. For each fluorescent spot, the fluorescence intensity above background was calculated in the 5-by-5 pixels (corresponding to 800 × 800 nm) surrounding the peak. If more than one peak are identified within the 5-by-5 region in a same cell, the masks are merged so that each pixel is counted only once. For each cell, the following information were recorded: the total FISH signal, the total YFP signal, the size of the cell, and the lengths of the major and minor axes of the cell. The accuracy of the analysis method is

illustrated in the following sections, where we discuss the false-negative and false-positive rates of the assay.

To reduce the gene dosage effect on the mRNA copy number distribution, a small set of cells within a certain size range was used for further analyses. The area of the cells ranges from ~1.9 μm^2 to ~4 μm^2, depending on the stage of the cell cycle. We selected the cells whose sizes are between 1.92 μm^2 to 2.30 μm^2. The fluorescence signal histograms are computed for each strain, including a mock strain which contains no YFP gene. The resulting histograms are deconvolved from the histogram of the mock strain, which represents the nonspecific signal level. The fluorescence signal is then normalized to the signal from a single fluorophore to convert to the absolute number of probes. The 95% confidence level in determining the mean fluorescence level, the Fano factor of the distribution, and the mRNA-protein correlation were estimated by bootstrapping.

Ensemble mRNA copy number measurement

To independently confirm the fidelity of FISH measurement, we compared the average mRNA copy number per cell measured by FISH with the number measured by quantitative PCR in bulk. The comparison was done in the *E. coli* strain PC2a, in which the YFP expression is control under the *lac* promoter at the *lac* operon locus on chromosome. The bulk measurement is performed in courtesy of Professor Nam-Ki Lee, with the following procedures. The cells were grown overnight in M9 glycerol medium supplemented with amino acids and vitamins at 37C. On the next day the culture was diluted 200 times into fresh M9 glycerol medium supplemented with amino acids, vitamins, and 1 mM IPTG. When the O.D. of the culture reached 0.2-0.3, 300 µl of the culture was transferred into 600 µl bacterial RNA stabilizer solution (Invitrogen). Total RNA was extracted using the RNeasy kit (invitrogen) following the manufacturer's guide. The residual gDNAs was removed using the TurboDNA-free kit (Ambion). The cell density of the culture was measured using a cell counter (Hausser Scientific) (REF). Reverse transcription was performed using SuperScript III (Invitrogen) at 50°C for 1 hour with primer sequence 5'-CGTCGTCCTTGAAGAAGATGG. Quantitative PCR was performed using the 7500 Fast Real-Time PCR System (Applied Biosystems), with a Taqman probe (5'-(FAM)ATCGCCCTCGCCCTC(MGB)) and two primers (5'-CGTCGTCCTTGAAGAAGATGG and 5'–CCGACCACTACCAGCAGAACA). Calibration was done using mRNA generated by *in vitro* transcription. The *E. coli* strain BW25993 was harvested, and mixed with known amount of *venus* mRNA at the amount of 1.0×10^{10}, 1.0×10^{9}, 1.0×10^{8}, and 1.0×10^{7} molecules, respectively. We extracted mRNA three times independently and for each extracted mRNA six RT-PCR reactions were performed. The calibration was performed in parallel with each measurement.

RNA half-life measurement by RNA-seq

DY330 cells were grown in M9 media (0.4% glucose, vitamins, amino acids, thiamine, biotin) at 30°C with shaking until OD_{600} = 0.3. 8 ml of cells were removed to 900 µl of cold 90:10 EtOH:phenol. Rifampicin (Sigma Aldrich) was added to a final concentration of 500 µg/ml, and further 8 ml aliquots of cells were removed at 2, 4, 6 and 8 min post drug. The cells were then harvested by centrifugation and washed once with 0.85% NaCl solution before storing in -80°C.

Frozen pellets were resuspended in 1 mg/ml lysozyme TE buffer and lysed by an equal volume of Cell Lysis Buffer (Purgene). Acidic phenol/chloroform (OmniPur) was added, and the aqueous layer was collected. The aqueous layer was extracted once in chloroform, and RNA was collected in RNA Clean and Concentrator columns (Zymo Research) and eluted in water. Contaminating DNA was removed by DNase I (NEB) treatment for 30 min at 37°C, and the resulting RNA repurified.

Starting with 5 μg of RNA, rRNA was removed first using Ambion's MICROBExpress following manufacturer's protocol, except RNA was collected using Zymo's RNA columns. A second rRNA removal step was performed following the protocol described in Affymetrix Expression Handbook, substituting enzymes MMLV (Ambion), RNase H (NEB), and DnaseI (Amplification grade, Invitrogen). 150-300 ng of RNA remain.

The RNA was fragmented using Ambion's Fragmentation Reagent at 70°C for 5 min, and collected by Zymo's RNA columns. RNA seq libraries were prepared according to Illumna's protocol, using NEB enzymes and barcoded adapters (Integrated DNA Technologies). The libraries were pooled and sequenced with an Illumina GA II machine (Center for Systems Biology, Harvard University). Sequences are available online (GEO accession number GSE21341).

After pooling barcodes and aligning the tags to the W3110 genome, the tag count for each gene was divided by the gene size to obtain the expression levels as the relative copy numbers of mRNA. Then, the sum of expression levels was normalized to give ~1350 molecules/cell (Bionumbers, R. Milo, et al, NAR, 2009).

Supplementary data

1. Calibration of single molecule fluorescence

The fluorescence count corresponding to single molecule fluorescence was calibrated in two ways (Fig. S1); one is a single molecule method (i) and the other is a bulk method (ii). We confirmed that the values from these two methods are consistent with each other.

(i) We measured SX4 strain expressing membrane-bound Tsr-Venus (*S9*). We obtained fluorescence counts from single Venus molecules by measuring the localized fluorescent spots. The obtained count was 161 counts/molecule/average cell volume. The cell volume is an average over the population.

(ii) We purified Venus protein and compared its fluorescence counts with that of a culture of a library cell strain (AcpP-Venus). The concentration of purified Venus ([*Venus*]) was measured by fluorometer (DU800, Beckman Coulter), and the density of cells (*C*) was obtained by a cell counter (Hasser Scientific Partnership). We injected cells, purified Venus, and 0.85% NaCl solution into different microfluidic channels pre-coated with BSA, and their fluorescence was observed with a 10× objective lens (Olympus). Comparison of fluorescence counts between those channels gives the protein number per cell for the cell sample, *n*, by the equation:

$$n = \frac{[Venus]\dfrac{F_c - F_w}{F_v - F_w} N_A}{C},$$

where N_A is the Avogadro's number, and F_c, F_v and F_w are the fluorescence counts from cells, Venus protein and NaCl solution, respectively. By comparing the value with the steady-state library data, we could obtain the calibration count consistent with the value determined in (i). The average fluorescence per cell for AcpP was 741,223, calculated as described in Methods; this gives 141 counts/molecule/average cell volume.

2. Consistency with single molecule localization method

Instead of counting protein molecules by detection by localization (*S9*), we used a deconvolution method to determine the protein count. The deconvolution method has several advantages: it can measure both low and high copy proteins, and is not affected by protein localization. To check if the deconvolution method is consistent with the results from localized single-molecule counting, we measured the abundance of localized Tsr-YFP molecules in strain SX701 (*S6*) as a function of an inducer, TMG. We found that the induction kinetics of tsr-venus is the same whether measured by counting localized molecules or by deconvolution (Fig. S2). This indicates that the deconvolution method has enough sensitivity and resolution to detect similar changes as the localization method.

3. Detection limit of the measurement system

We first confirmed the single molecule sensitivity of our microscope by observing a one-step photobleaching of low-copy, membrane-bound library strains (Fig. S14B). We checked the detection limit of our measurement system by deconvoluting the fluorescence histogram of wild type cells from the auto-fluorescence histogram of the control cells. The limit was determined to be 0.08 ± 0.08 /cell (mean ± SD, N = 15), which allows detection of 99.3% of the library data. We also verified the accuracy of our measurements by showing that measurements made on two separate occasions are largely reproducible (r = 0.92, Fig. S3A). We also determined the detection limit of protein noise to be ~0.01 by measuring control samples (Fig. S3B).

4. Miller assay

We use a reporter protein, LacZ, to check that fusing SPA-Venus (SPA is a scar sequence from the previous library) or Venus to proteins did not affect their expression levels. Assuming that the β-galactosidase (β-gal) activity of LacZ is unchanged by C-terminal protein fusions, we can use the Miller assay to report on LacZ protein abundance. We can vary inducer (IPTG or TMG) concentration to achieve different levels of LacZ expression, to mimic low or high copy proteins. The Miller assay was performed using the yeast β-galactosidase assay kit (Pierce). We measured three different constructs: 1) LacZ, 2) LacZ + Venus, 3) LacZ + SPA + Venus. The C-terminal fusions of Venus and SPA-Venus resulted in at most a 2-3 fold reduction at high expression levels of LacZ as measured by β-gal activity (Figure S4). There was less perturbation in protein expression when LacZ protein was at low concentrations. This suggests that the C-terminal SPA-Venus tag did not cause significant changes in protein expression levels. We also confirmed that the β-gal activity is proportional to the observed fluorescence value (Fig. S4C). This shows that there is no appreciable self-quenching of fluorescence at high expression levels. Furthermore, the direct comparison between the fluorescence and β-gal activity shows that the fluorescent reporter is linear for at least three orders of magnitude.

5. Global relationship between parameters.
Figure S5 shows global relationships between obtained parameters. The mean-skewness graph and the mean-kurtosis plots (Fig. S5) showed a monotonous decrease with a rise of mean, describing that the shapes of protein distributions are changed from a skewed distribution with a peak at zero to more symmetric distribution with a non-zero peak. The correlation coefficients and Z scores between these parameters are summarized in Table S2.

In addition, we characterized the preference of parameters for each functional category by *Z*-score. Some functional categories are strongly correlated with some parameters. For example, essential proteins have a strong correlation with high a (Z = 7.5) and high b (Z = 5.3). As expected, membrane proteins showed high edge/inside ratio (Z = 3.4-7.2), and transcriptional repressors indicated high punctuate localization (Z = 4.1). The mean expression levels are not strongly correlated with the chromosomal positions. These results describe that some protein expression characteristics are significantly correlated with functional properties. The results were summarized in Table S3

6. Comparison with yeast
Newman *et al.* has investigated for >2,500 high-abundance protein in yeast using a flow cytometry (*S11*). The work examined noise properties of a subset of cells in which cells were gated for size and granularity using forward/side scatter parameters. Figure S6A shows yeast's noise and Fano factor, which is roughly estimated from the yeast data, plotted over the *E. coli* data. The *E. coli* data at high abundance was similar to the ungated yeast data, but differed from the gated data.

We hypothesized that this difference is due to the fact that in *E. coil* the noise is not sensitive to cell size and granularity. Newman's data reported that cell size or cell growth state can affect noise properties by 200-300 % (*S11*). To check this in *E. coli*, we analyzed cell length dependence on the protein distributions. The data indicated that the noise property was almost independent of the cell length (Fig. S6B). This means that the heterogeneity of cell size and granularity is not a dominant factor in *E. coli* protein noise, unlike in yeast, and suggests that the maximum noise level without gating might be similar between *E. coli* and yeast.

7. The measurement limits of noise parameters
We confirmed that the detection limit of the microscope system is below the noise measured for all strains. For this purpose, a bright cell strain (open symbols in Fig. S3B) and a fluorescent dye solution (closed symbols) were measured with various laser intensities. This control measurement showed that the fluorescence intensity was linear with respect to the concentration, and the instrument noise was smaller than the biological noise for the entire range of data collection, indicating that the observed two-scaling is not caused by measurement limitations of the setup or shot noise from low signal intensities.

8. Noise propagation in high copy strains
By performing a real-time observation of protein levels for four randomly selected high copy library strains, we demonstrated that a slowly varying heterogeneity exists between cells in a population (Fig. 2C). To quantitatively understand the relationship between the variation within a

single cell over time and the variation among different cells at a particular moment, we analyzed the fluorescence time traces in two ways.

First, we used the F-test to show that the variation within a single cell cycle is much smaller than that of the population. Here F is defined as $F = \eta^2_{\text{cell-to-cell}}/\eta^2_{\text{single}}$, where η^2_{single} is the amplitude of protein level fluctuations during a cell cycle, $\eta^2_{\text{cell-to-cell}}$ is the cell-to-cell variation of protein levels. We confirmed a significant difference between dynamical noise of single cells and population noise ($F(2600, 2052) = 5.8\text{-}6.4$, $p < 0.0001$). This result justifies our assumption of static heterogeneity in the derivation of the model used in the main text.

Second, we examined the timescale it takes for the variation within a single cell lineage to reach that of the population. In other words, we aimed to probe the timescale for noise to be ergodic (population average = temporal average). We calculated the noise from temporal fluctuations of a single cell using time windows of varying sizes (Fig. S7A). If gene expression fluctuations are ergodic, the temporal fluctuations of a single cell over a very large time window should match the cell-to-cell variation across the population. We observe that several cell cycles is insufficient for the noise of a single cell to match the population noise.

To estimate the timescale required for ergodicity, we approximated the temporal noise as asymptotically approaching the population noise in an exponential manner as we increase the size of the time window. This approximation is motivated by the following analogy: For a particle diffusing within a harmonic potential from an initial position, the accumulated variance of position exponentially approaches the steady state value when the time interval approaches infinity.

Fitting an exponential function to the temporal noise versus time window gives as estimate of 8-21 cell cycles for the timescale of long time fluctuation. Because we are extrapolating longer timescales from short measurements, this is only an estimate.

We also found that applying an autocorrelation analysis provided a decay time of 1-2 cell cycles, as previously observed (*S18*). However, we conclude that this almost linear decay in our data and Rosenfeld's data is the result of an insufficient time window, since our direct comparison of temporal and population noise indicates that 1-2 cell cycles are insufficient to average out extrinsic noise (Fig. S7B). We note fluctuations occur at many different time scales. Our time traces only have 4-5 cell cycles, i.e. 500 min. During this time, autocorrelation function is sufficient to determine the short timescale of fluctuation. However, fluctuations at longer time scales can be estimated by the method described above.

How is such a long timescale generated? Our extrinsic noise is likely determined by global factors such as ribosomes and RNA polymerases, not specific transcription factors participating in a positive feedback loop for each gene. We know that ribosomal and RNAP activities have very complex regulation and are highly sensitive to growth conditions, but how absolute transcription or translation rates are regulated on a single cell basis is still largely unknown. It is very probable that there are feedback loops in the regulation of global factors and cell metabolism. The value of 8-21 cell cycles is an estimate, but there are other examples of epigenetic effects in bacteria with ~10 cell cycles.

9. Formulation of intrinsic and extrinsic noise

We start with the model of stochastic gene expression:

$$\mathrm{DNA} \xrightarrow{k_1} \mathrm{mRNA} \xrightarrow{k_2} \mathrm{Protein}$$
$$\gamma_1 \downarrow \qquad \gamma_2 \downarrow$$
$$\emptyset \qquad \emptyset$$

In this scheme, protein production randomly occurs in bursts, in which each mRNA molecule is translated into an exponentially-distributed number of protein molecules before its degradation, as supported by previous observations (*S9, S12*).

We note here that mRNA life time is substantially short compared to the protein dilution time ($\gamma_1 >> \gamma_2$) in *E. coli*. In *E. coli*, proteins are generally degraded on timescales of significantly longer than the cell cycle. Koch and Levy (*34*) found that <0.1% of protein mass is degraded per hour in *E. coli* growing in exponential phase. Therefore, most of the protein lifetime is controlled by a dilution due to cell division, whose timescale equals with the cell cycle (~150 minutes). In parallel, we have determined by RNAseq that the mRNA of most genes has lifetime between 2 to 10 minutes (Section 13). These observations support the model of an exponential protein burst.

We define the mRNA burst frequency and size as a (= k_1/γ_2) and b (= k_2/γ_1), considering that the mRNA life time is short compared to the protein life time ($\gamma_1 >> \gamma_2$). Assuming a and b are uniform among cells in a population, the steady-state distribution can be solved as a gamma distribution by solving a chemical master equation based on this scheme (*S13*). For low copy proteins, a discrete master equation gives a more accurate solution as a negative binomial distribution (*S14*). In the argument below, we use the negative binomial distribution for an accurate description.

In this section, we extend it to the case that the a and b values have static cell-to-cell variation. We assumed that the variation of a and b are given by the stationary probability density functions of $p(a)$ and $p(b)$ and that the a and b are independent of each other. Under this assumption, the steady-state protein number distribution is given by:

$$p(n) = \int_0^\infty \int_0^\infty p(n \mid a,b)\, p(a)\, p(b) da db, \qquad \text{(S1)}$$

where n is the protein number. $p(n|a, b)$ is the conditional probability for a and b, equal to a negative binomial distribution:

$$p(n \mid a,b) = \frac{b^n}{(1+b)^{a+n}} \frac{\Gamma(a+n)}{\Gamma(a) n!}, \qquad \text{(S2)}$$

where $\Gamma(a)$ is the Gamma function. The first and second moments of this distribution are given by:

$$\sum_{n=0}^{\infty} np(n \mid a,b) = ab, \quad \sum_{n=0}^{\infty} n^2 p(n \mid a,b) = ab(1+b) + a^2 b^2. \qquad \text{(S3)}$$

Using (S1)-(S3), the mean and standard deviation of the protein number is calculated as:

$$\mu \equiv \langle n \rangle = \sum_{n=0}^{\infty} np(n) = \langle a \rangle \langle b \rangle, \tag{S4}$$

$$\sigma^2 \equiv \langle n^2 \rangle - \langle n \rangle^2 = \sum_{n=0}^{\infty} n^2 p(n) - \left(\sum_{n=0}^{\infty} np(n) \right)^2, \tag{S5}$$

$$= \langle a^2 \rangle \langle b^2 \rangle + \langle a \rangle \langle b^2 \rangle + \langle a \rangle \langle b \rangle$$

where $\langle a^n \rangle$ and $\langle b^n \rangle$ are the nth moments of a and b, defined by: $\langle a^n \rangle = \int a^n p(a) da$ and $\langle b^n \rangle = \int b^n p(b) db$. Therefore, the noise, $\eta^2 \equiv \sigma^2/\mu^2$, is given by:

$$\eta^2 = \frac{1 + \langle b \rangle}{\mu} + \frac{\langle b \rangle \eta_b^2}{\mu} + \eta_a^2 + \eta_a^2 \eta_b^2 + \eta_b^2 . \tag{S6}$$

η_a^2 and η_b^2 are the noise of a and b values defined by: $\eta_a^2 = \sigma_a^2/\mu_a^2$ and $\eta_b^2 = \sigma_b^2/\mu_b^2$, where $\mu_\xi = \langle \xi \rangle$ and $\sigma_\xi^2 = \langle \xi^2 \rangle - \langle \xi \rangle^2$ ($\xi = a$ or b). This noise equation can be separated into an intrinsic noise term independent of η_a and η_b, and an extrinsic noise term:

$$\eta_{\text{int}}^2 = \frac{1 + \langle b \rangle}{\mu}, \tag{S7}$$

$$\eta_{\text{ext}}^2 = \frac{\langle b \rangle \eta_b^2}{\mu} + \eta_a^2 + \eta_a^2 \eta_b^2 + \eta_b^2 . \tag{S8}$$

In the case of the gamma distribution for continuous variables, the noise values are the same except that the term $1/\mu$ is removed in equations (S6) and (S7).

The co-existence of μ^{-1} independent terms and μ^{-1} dependent terms explains the two scaling of noise (Fig. 2B). Without the extrinsic noise, the noise only has a μ^{-1} dependent term resulting in a single $1/\mu$ scaling, assuming that translational efficiencies for all genes are constant or the same order. This constant average translational efficiency would be natural because gene-specific gene regulation often happens in the transcriptional process in bacteria. The incorporation of extrinsic noise adds μ^{-1} independent terms to the total noise, resulting in another flat scaling that appears at high expression levels as the μ^{-1} terms vanish. Thus, the two noise scaling is due to the existence of extrinsic noise in addition to intrinsic noise.

This equation also explains the distributions of the noise plot in Fig. 2B. At low copy numbers, $1/\mu$ represents the lower limit of the noise, which is the case that no extrinsic noise ($\eta^2_{\text{a}} = 0$ and $\eta^2_{\text{b}} = 0$) and minimum protein burst size ($b = 1$) is assumed in Eqn. 4. The noise of each protein is set by adding the extrinsic noise and burst size contribution above this limit. This is why almost all noise plots are scattered above $1/\mu$, which we call as the $1/\mu$ scaling. A different lower noise limit is also found at high copy numbers, which is dominated by the extrinsic noise.

10. Real-time observation of low copy strain

The heterogeneous stochastic gene expression model (Section 9 in the supporting data) predicts that, in low copy proteins, the α and β values obtained by the noise profiling experiments equal the a and b values, that is the frequency and size of the mRNA transcriptional bursts, correspondingly. This has been shown to be true for a selected system (*S12*). To confirm this equality more generally, we directly observed the burst occurrence of 3 low-copy membrane-bound library strains by using the real-time method of visualizing individual protein localization

(*S9*). We found that the observed *a* and *b* values are proportional to the α and β values (Table S4), consistent with our prediction. The scaled difference in the *b* and β value would be caused by the difference in measurement environments between the real-time and steady-state experiments. For example, cells for the real-time experiment were grown on an agarose pad while cells for the steady-state experiments were grown in liquid culture.

11. Identification of global extrinsic noise from two-color experiments

There are two dominant classes of extrinsic noise factors affecting the noise scaling: (i) global extrinsic factors and (ii) gene-specific extrinsic factors. Whereas the global factors affect all the genes in a similar way (e.g. heterogeneity in RNA polymerase and ribosome numbers, or cell state), the gene-specific factors affect only specific genes (e.g. heterogeneity in transcription factor numbers). The dominance of the global factor noise causes a global correlation for all protein levels in a single cell, meaning that the extrinsic noise contributions can be determined by measuring correlations between multiple genes.

In this section, by modeling the global extrinsic noise factors and the gene-specific extrinsic noise factors, we formulate the relationship between the global extrinsic noise and the global protein level correlation. We assume that the *a* and *b* values are provided as the result of the first order reaction:

$$a = a_\zeta a_G \text{ and } b = b_G, \tag{S9}$$

where a_G and b_G are the global contributions to *a* and *b*, and a_ζ is the gene-specific contribution to the *a* value for protein ζ. For example, a_ζ may be proportional to the concentration of an activator and a_G may be proportional to the concentration of RNAP. Those factors, a_G, b_G and a_ζ, are assumed to have cell-to-cell variation described as probability densities, $p(a_G)$, $p(b_G)$ and $p(a_\zeta)$, respectively, and are assumed to be independent of each other, similar to Section 10. Under these assumptions, the distribution of the number of protein ζ, n_ζ, is expressed as:

$$p(n_\zeta) = \int_0^\infty \int_0^\infty \int_0^\infty p(n_\zeta \mid a_\zeta, a_G, b_G)\, p(a_\zeta)\, p(a_G)\, p(b_G)\, da_\zeta da_G db_G, \tag{S10}$$

where the conditional probability, $p(n_\zeta|a_\zeta,a_G,b_G)$, is given by the negative binomial distribution:

$$p(n_\zeta \mid a_\zeta, a_G, b_G) = \frac{b_G^n}{(1+b_G)^{a_\zeta a_G + n_\zeta}} \frac{\Gamma(a_\zeta a_G + n_\zeta)}{\Gamma(a_\zeta a_G) n_\zeta!}. \tag{S11}$$

Using (S3) and (S9) to (S11), the mean and standard deviation of protein number are given by:

$$\mu_\zeta \equiv \langle n_\zeta \rangle = \sum_{n_\zeta=0}^{\infty} n_\zeta\, p(n_\zeta) = \langle a_\zeta \rangle \langle a_G \rangle \langle b_G \rangle \quad , \tag{S12}$$

$$\begin{aligned} \sigma_\zeta^2 &\equiv \langle n_\zeta^2 \rangle - \langle n_\zeta \rangle^2 = \sum_{n_\zeta=0}^{\infty} n_\zeta^2\, p(n_\zeta) - \left(\sum_{n_\zeta=0}^{\infty} n_\zeta^2\, p(n_\zeta) \right)^2, \\ &= \langle a_\zeta^2 \rangle \langle a_G^2 \rangle \langle b_G^2 \rangle + \langle a_\zeta \rangle \langle a_G \rangle \langle b_G^2 \rangle + \langle a_\zeta \rangle \langle a_G \rangle \langle b_G \rangle \end{aligned} \tag{S13}$$

Therefore, the total noise is described by:

$$\eta_\zeta^2 = \eta_{\text{int}}^2 + \eta_{\text{ext-global}}^2 + \eta_{\text{ext-gene}}^2, \tag{S14}$$

and:

$$\eta_{\text{int}}^2 = \frac{1+\langle b_G \rangle}{\mu_\zeta}, \tag{S15}$$

$$\eta_{\text{ext-global}}^2 = \eta_{aG}^2 + \eta_{aG}^2\eta_{bG}^2 + \eta_{bG}^2 \equiv \eta_G^2, \tag{S16}$$

$$\eta_{\text{ext-gene}}^2 = \frac{\langle b_G \rangle \eta_{bG}^2}{\mu_\zeta} + \eta_{a\zeta}^2 + \eta_{a\zeta}^2\eta_G^2, \tag{S17}$$

where $\mu_\xi \equiv \langle \xi \rangle$, $\sigma_\xi^2 \equiv \langle \xi^2 \rangle - \langle \xi \rangle^2$, $\eta_\xi^2 \equiv \sigma_\xi^2/\mu_\xi^2$ ($\xi = a_\zeta, a_G, b_G$). η^2_{int} is the intrinsic noise caused by a stochastic production and degradation of RNA and protein molecules. $\eta^2_{\text{ext-global}}$ is the global extrinsic noise, probably due to RNA polymerase or ribosome number fluctuations. $\eta^2_{\text{ext-gene}}$ is the gene-specific extrinsic noise scaled by gene-specific regulation such as transcription factors.

In addition, we assumed that the parameters, a_G, b_G and a_ξ, are the only determinants of correlation among the protein levels of different genes. Under this assumption, the joint probability of the numbers of protein *x* and *y* is given by:

$$p(n_x, n_y) = \int_0^\infty \int_0^\infty \int_0^\infty p(n_x, n_y \mid a_x, a_y, a_G, b_G)\, p(n_x)\, p(n_y)\, p(a_G)\, p(b_G)\, dn_x\, dn_y\, da_G\, db_G, \tag{S18}$$

and the conditional probability, $p(n_x,n_y|a_x,a_y,a_G,b_G)$, has the relationship:

$$p(n_x, n_y \mid a_x, a_y, a_G, b_G) = p(n_x \mid a_x, a_G, b_G)\, p(n_y \mid a_y, a_G, b_G). \tag{S19}$$

From (S3), (S9), (S11), (S18) and (S19), the first moment of the covariance of the numbers of protein *x* and *y* is given by:

$$\langle n_x n_y \rangle \equiv \sum_{n_x=0}^{\infty} \sum_{n_y=0}^{\infty} n_x n_y\, p(n_x, n_y) = \langle a_x \rangle \langle a_y \rangle \langle a_G^2 \rangle \langle b_G^2 \rangle \tag{S20}$$

By dividing (S20) by (S12), we obtained the relationship:

$$\frac{\langle n_x n_y \rangle}{\langle n_x \rangle \langle n_y \rangle} = 1 + \eta_G^2, \tag{S21}$$

where η_G^2 is a sum of the global factor noise defined in (S16). The derived equation shows that the global noise factor is related with the normalized correlation factor, $\langle n_x n_y \rangle / \langle n_x \rangle \langle n_y \rangle$.

To determine the global factor noise, we randomly selected 13 combinations of doubly-labeled high expression gene pairs, where one gene is probed by Venus and the other is probed by mCherry. We observed positive correlations (r = 0.2-0.8) for all 13 two-protein combinations (Fig. 2D), confirming the existence of a global noise factor. We found that the normalized correlation factors, $\langle n_x n_y \rangle / \langle n_x \rangle \langle n_y \rangle$, was very uniform for all measured strains ($\langle n_x n_y \rangle / \langle n_x \rangle \langle n_y \rangle$ = 1.09 ± 0.03, mean ± SD, 13 strains), supporting the gene-independent property of the extrinsic noise. Thus, the global noise factor was determined to be 0.09. This means that the robustness of any gene function must take into account the unavoidable 30% variation in expression levels. The determined extrinsic noise limit is consistent with the limiting value of protein noise.

The determined extrinsic noise limit is consistent with the limiting value of noise obtained in our system-wide noise measurements (Fig. 2B). At high mean expression levels, and in the absence of other gene-specific noise, we note that $\eta^2 \approx \eta_{\text{ext-global}}^2$. Thus, the combined contributions of

this global noise and the μ^{-1}-scaled intrinsic noise set the lower fundamental limit of total noise. Contributions from gene-specific noise, $\eta_{\text{ext-gene}}^2$, will increase the noise value for particular genes above this limit.

12. Real-time dynamics of correlation between two gene expressions.

We studied the dynamic properties of global extrinsic noise by performing a real-time observation of the two-color strains. Figure S8 shows the real-time two-color traces. The two protein levels are correlated, suggesting the contribution of global extrinsic noise. We analyzed a correlation coefficient for different time window sizes as in Section 8. For very short observation time windows (<20 minutes), correlation is small because intrinsic fluctuations within each gene (Fig. S9). However, on longer timescales, the two colors become correlated because of the common extrinsic noise. Thus, one would expect a fast rise in the correlation from intrinsic noise, and then an eventual plateau at the level of correlation from extrinsic noise. The 20 minute timescale would correspond to the intrinsic noise fluctuations.

13. Consistency with other bulk measurement

In this work, we have shown that average protein abundances span five orders of magnitude. We note that this is consistent with previous observations. Previously a mass spec and western blot analysis (*S4*) has shown an abundance data for high copy genes in an E. coli strain (K-12, MOPS media), and has indicated that the average abundance spans 10^3-10^4 magnitude, which is largely consistent with our result. We confirmed that their data sets, despite the different growth condition and strain background, are largely correlated with our data (r = 0.58 and 0.48, respectively), especially for high copy numbers, as shown in Figure S21. In addition, for a specific gene (LacZ), we have already confirmed that the YFP fluorescence is linearly proportional to its enzymatic activity at different expression levels on 2-3 orders of magnitude (Section S4 and Fig. S4).

14. Validation of single molecule detection in FISH

The detection of single hybridized probes was validated by the observation of single-step photobleaching (Fig. S17). When multiple mRNA molecules are localized within a diffraction-limited volume, multi-step photobleaching was also observed (Fig S18). Prior to the last photobleaching step, the average signal strength is the same as that of a single fluorophore with >95% confidence. The distribution of the signal intensities from single fluorophores (Atto 594) was Gaussian-like, with ~27% variation (Fig. S17).

15. Determination of false-positive and false-negative in single molecule FISH

To determine the false-positive rate (including nonspecific hybridization) in our assay, the same FISH method and analysis were applied to the E. coli strain BW25993 (*S15*) which contains no YFP coding sequence. One hybridization event was detected in an average of ten cells, indicating a false-positive rate of 0.1 per cell. Similarly, same false-positive rate was detected in an *E. coli* strain that contains the YFP coding sequence and in which the expression of the *yfp* mRNA was suppressed under the *lac* promoter (strain PC2a). The mRNA expression level in this strain was determined to be < 0.04/cell (*S9*), which is below our false-positive rate. The fact that we observed the same false-positive rate in the wild type strain and in the strain that contain the YFP coding sequence but no *yfp* mRNA indicates that there is minimal hybridization between the FISH probe and the genomic DNA under the experimental conditions. The intensity distribution

of the false-positive signals resembles that of a single fluororphore, suggesting that the signal arose from actual probes, rather than noise in the imaging system, that are nonspecifically bound in the cell.

To determine the false-negative rate for our assay, we compared the mRNA copy number per cell measured by FISH verses that measured by quantitative PCR in bulk. *E. coli* strain PC2a, in which YFP is induced under the *lac* promoter, was grown in M9 glycerol minimal medium supplemented with 1 mM IPTG, amino acids, and vitamins, and was harvested in log phase for the proceeding FISH, RNA extraction, or cell counting. The average *yfp* mRNA copy number per cell measured by quantitative PCR is 4.1 ± 0.7 (SEM, N = 6), whereas that measured by FISH is 3.77 ± 0.07 (SEM, N = 6,528). Therefore, the detection efficiency of our FISH assay is ~92%.

16. Comparison of YFP fluorescence in live cells and in fixed cells

Figure S19 shows that the YFP fluorescence after FISH is proportional to that in live cells. The fluorescence intensity after fixation is at ~70% of the original level in live cells. Comparing the fluorescence before and after FISH for each gene, the correlation is r = 0.88.

17. Stochastic model for mRNA-protein correlation under Poissonian or non-Poissonian transcription

Here we describe five basic models to quantitatively understand the lack of mRNA-protein correlations (Fig. 4). We show that while the difference in mRNA and protein lifetimes predicts a small correlation coefficient, it is not small enough to account for the near zero correlation observed for the majority of genes. We found that the extrinsic noise originated from translational heterogeneity or autorepression can further reduce the mRNA-protein correlation coefficient, whereas the noise from transcriptional heterogeneity can increase the correlation coefficient. Based on these models, the lack of correlation is primarily due to the lifetime differences between mRNA and protein. However, translational heterogeneity most likely further reduces the correlation to zero.

Case 1: Constitutive promoter

For the simplest case, we consider stochastic gene expression from a constitutive promoter:

$$\mathrm{DNA} \xrightarrow{k_1} \mathrm{mRNA} \xrightarrow{k_2} \mathrm{Protein}$$
$$\gamma_1 \downarrow \qquad \gamma_2 \downarrow$$
$$\varnothing \qquad \varnothing$$

In a deterministic model, every cell has the same mRNA level, $\mu_m = k_1/\gamma_1$, and the same protein level, $\mu_p = k_1 k_2/\gamma_1\gamma_2$.

In a stochastic model, every cell has different mRNA and protein levels. To understand the correlation of mRNA and protein within a single cell, we analyzed the stochastic reactions using the approach devised by Paulsson (*S16*). The steady-state solution provides the standard deviation of mRNA and protein distribution, σ_m and σ_p, and their covariance, σ_{mp}, to be:

$$\sigma_m^{\,2} = \frac{k_1}{\gamma_1}$$

$$\sigma_p^{\,2} = \frac{k_1 k_2}{\gamma_1 \gamma_2}\left(1 + \frac{k_2}{\gamma_1 + \gamma_2}\right)$$
$$\sigma_{mp}^{\,2} = \frac{k_1 k_2}{\gamma_1(\gamma_1 + \gamma_2)} \quad . \tag{S25}$$

And the Pearson correlation, r, is given by:

$$r_{constitutive} = \frac{\sigma_{mp}^2}{\sigma_m \sigma_p} = \sqrt{\frac{\gamma_2}{\gamma_1 + \gamma_2}\frac{1}{1 + \frac{\gamma_1 + \gamma_2}{k_2}}} \quad . \tag{S26}$$

From the FISH measurement for high copy transcripts, we obtained the translation rate to be 0.6-60/min. With mRNA lifetime of ~5 min and protein lifetime of ~180 min, the correlation would be r = 0.13-0.16. This is not quite small enough to explain the almost zero correlation observed for the majority of genes.

Case2: Two-state promoter fluctuation

Here we consider an extended model in which the transcription rate is allowed to fluctuate between zero and a finite value.

$$\text{Gene off} \underset{k_-}{\overset{k_+}{\leftrightarrow}} \text{on} \xrightarrow{k_1} \text{mRNA} \xrightarrow{k_2} \text{Protein}$$
$$\text{mRNA} \xrightarrow{\gamma_1} \emptyset \qquad \text{Protein} \xrightarrow{\gamma_2} \emptyset$$

This so-called "two-state" model has been extensively explored theoretically (*S16*), and can be used to describe many possible molecular mechanisms that causes non-Poissonian mRNA distribution, such as RNA polymerase poising or stalling, transcription factor binding and unbinding, and chromatin remodeling. It produces non-Poissonian mRNA distribution when the transcription fluctuation rates (k_+ and k_-) are comparable to or slower than the mRNA degradation rate (γ_1).

Define $P \equiv k_+/(k_+ + k_-)$ as the probability to beat the "on" state, and $k_0 \equiv k_+ + k_-$ as the rate of promoter fluctuation. The steady-state solution ($d\sigma/dt = 0$) for the variance and covariance is:

$$\sigma_m^{\,2} = P\frac{k_1}{\gamma_1}\left(1 + (1-P)\frac{k_1}{k_0 + \gamma_1}\right)$$
$$\sigma_p^{\,2} = P\frac{k_1 k_2}{\gamma_1 \gamma_2}\left(1 + \frac{k_2}{\gamma_1 + \gamma_2}\left(1 + (1-P)\frac{k_1(k_0 + \gamma_1 + \gamma_2)}{(k_0 + \gamma_1)(k_0 + \gamma_2)}\right)\right)$$
$$\sigma_{mp}^{\,2} = P\frac{k_1 k_2}{\gamma_1 \gamma_2}\frac{\gamma_2}{\gamma_1 + \gamma_2}\left(1 + (1-P)\frac{k_1(k_0 + \gamma_1 + \gamma_2)}{(k_0 + \gamma_1)(k_0 + \gamma_2)}\right) \quad . \tag{S27}$$

And the Pearson correlation between mRNA and protein is:

$$r_{2state} = \sqrt{\frac{\gamma_2}{\gamma_1 + \gamma_2}\frac{1 + \frac{k_1\gamma_1(1-P)}{(1-P)k_1(k_0+\gamma_2) + (k_0+\gamma_1)(k_0+\gamma_2)}}{1 + \frac{\gamma_1+\gamma_2}{k_2}\left(1 + (1-P)\frac{k_1(k_0+\gamma_1+\gamma_2)}{(k_0+\gamma_1)(k_0+\gamma_2)}\right)^{-1}}} \tag{S28}$$

The correlation of the two-state model is always greater than the correlation of a constitutive promoter, i.e.

$$r_{2state} > \sqrt{\frac{\gamma_2}{\gamma_1+\gamma_2}\frac{1}{1+\frac{\gamma_1+\gamma_2}{k_2}}} = r_{constitutive} \quad . \tag{S29}$$

This suggests that the non-Poissonian, two-state transcription would make mRNA and protein more correlated, and thus cannot explain the vanishing correlation observed.

Case 3: Extrinsic noise at transcription level

Next we take into account the extrinsic noise that dominates protein level variation. As the real-time experiments suggested, the extrinsic noise fluctuates much slower than any other time scale, and can be viewed as static heterogeneity among a population. The heterogeneity can be at either the transcription level or the translational level. In this section we first consider the former case, and show that the transcriptional heterogeneity does not reduce the mRNA-protein correlation.

Let's consider again the constitutive promoter, but this time assume that the rates of transcription, k_1 and γ_1, stably varies from cell to cell, and are independent each other. Using the law of total variance and covariance and equations S25, we can write down the components of the covariance matrix:

$$\begin{aligned}
\sigma_m^2 &= \langle k_1 \rangle \left\langle \frac{1}{\gamma_1} \right\rangle + V_1 \\
\sigma_p^2 &= \frac{k_2}{\gamma_2}\langle k_1 \rangle \left\langle \frac{1}{\gamma_1} \right\rangle + \frac{k_2^2}{\gamma_2}\langle k_1 \rangle \left\langle \frac{1}{\gamma_1(\gamma_1+\gamma_2)} \right\rangle + \frac{k_2^2}{\gamma_2^2} V_1 \\
\sigma_{mp}^2 &= k_2 \langle k_1 \rangle \left\langle \frac{1}{\gamma_1(\gamma_1+\gamma_2)} \right\rangle + \frac{k_2}{\gamma_2} V_1
\end{aligned} \tag{S30}$$

where $V_1 = Var[k_1/\gamma_1] = \langle k_1^2 \rangle \langle 1/\gamma_1^2 \rangle - \langle k_1 \rangle^2 \langle 1/\gamma_1 \rangle^2$. And the correlation coefficient is:

$$r = \sqrt{\frac{\gamma_2}{\frac{1}{k_2}\frac{\left\langle \frac{1}{\gamma_1} \right\rangle^2}{\left\langle \frac{1}{\gamma_1(\gamma_1+\gamma_2)} \right\rangle^2} + \frac{\left\langle \frac{1}{\gamma_1} \right\rangle}{\left\langle \frac{1}{\gamma_1(\gamma_1+\gamma_2)} \right\rangle}}} \sqrt{\frac{1 + 2V_1 \frac{1}{\gamma_2 \langle k_1 \rangle \left\langle \frac{1}{\gamma_1(\gamma_1+\gamma_2)} \right\rangle} + V_1^2 \left[\frac{1}{\gamma_2 \langle k_1 \rangle \left\langle \frac{1}{\gamma_1(\gamma_1+\gamma_2)} \right\rangle} \right]^2}{1 + V_1 \frac{1}{\langle k_1 \rangle} \left[\frac{1}{\left\langle \frac{1}{\gamma_1} \right\rangle} + \frac{1}{r_2 \left(\frac{1}{k_2} \left\langle \frac{1}{\gamma_1} \right\rangle + \left\langle \frac{1}{\gamma_1(\gamma_1+\gamma_2)} \right\rangle \right)} \right] + V_1^2 \frac{1}{\langle k_1 \rangle^2 \left\langle \frac{1}{\gamma_1} \right\rangle r_2 \left(\frac{1}{k_2} \left\langle \frac{1}{\gamma_1} \right\rangle + \left\langle \frac{1}{\gamma_1(\gamma_1+\gamma_2)} \right\rangle \right)}}} \tag{S31}$$

where $V_1 = Var[k_1/\gamma_1] = \langle k_1^2 \rangle \langle 1/\gamma_1^2 \rangle - \langle k_1 \rangle^2 \langle 1/\gamma_1 \rangle^2$. The second square-root factor is greater than one, because the coefficients from the numerator are greater than the coefficients of the same order in V_1 from the denominator. This can be understood by using the relationship: $\langle 1/(\gamma_1(\gamma_1 + \gamma_2)) \rangle = (1/\gamma_2)(\langle 1/\gamma_1 \rangle - \langle 1/(\gamma_1 + \gamma_2) \rangle) < (1/\gamma_2)\langle 1/\gamma_1 \rangle$. From the Cauchy-Schwarz-Buniakowsky

inequality, we obtain the relationships: $\langle 1/\gamma_1\rangle^2 \le \langle 1/(\gamma_1(\gamma_1+\gamma_2))\rangle\langle(\gamma_1+\gamma_2)/\gamma_1\rangle$, and $1 \le \langle\gamma_1\rangle\langle 1/\gamma_1\rangle$. Using these relationships, the first square-root becomes:

$$r \ge \sqrt{\frac{\gamma_2}{\frac{1}{k_2}\frac{\left\langle\frac{1}{\gamma_1}\right\rangle^2}{\left\langle\frac{1}{\gamma_1(\gamma_1+\gamma_2)}\right\rangle^2}+\frac{\left\langle\frac{1}{\gamma_1}\right\rangle}{\left\langle\frac{1}{\gamma_1(\gamma_1+\gamma_2)}\right\rangle}}} \ge \sqrt{\frac{\gamma_2}{\langle\gamma_1\rangle+\gamma_2}\frac{1}{1+\frac{\langle\gamma_1\rangle+\gamma_2}{k_2}}} = r_{constitutive} \qquad \text{(S32)}$$

Therefore, any heterogeneity at the transcripton level (k_1 and γ_1) would only result in higher mRNA-protein correlation, contrary to our experimental observation.

Case 4: Autorepression

Here we consider the case that the transcription is affected by negative or positive feedback, and the transcriptional rate is given by multiplying k_1 and the feedback factor, $c(p)$, where p is the protein number. Using relationships in ref. S16, the variance and covariance are given by:

$$\sigma_m^2 = \langle m\rangle\left(1-\frac{H}{1+H}\frac{1}{\langle p\rangle}\frac{\langle p\rangle r_2-\langle m\rangle r_1 H}{\gamma_1+\gamma_2}\right)$$

$$\sigma_p^2 = \langle p\rangle\left(1+\frac{1}{1+H}\frac{1}{\langle m\rangle}\frac{\langle p\rangle r_2-\langle m\rangle r_1 H}{\gamma_1+\gamma_2}\right)$$

$$\sigma_{mp}^2 = \frac{1}{1+H}\frac{\langle p\rangle r_2-\langle m\rangle r_1 H}{\gamma_1+\gamma_2}$$

where

$$H = -\frac{\partial \ln c(p)}{\partial \ln p} \qquad \text{(S36)}$$

Assuming that $c(p)$ is a monotonous function of p, $H > 0$ corresponds to negative feedback, whereas $H < 0$ corresponds to positive feedback. $\langle m\rangle$ and $\langle p\rangle$ is the expectation value of the mRNA and protein number, respectively. The correlation coefficient is given by:

$$r = r_{constitutive}\frac{\frac{|1+H|}{1+H}\left(1-\frac{\langle m\rangle}{\langle p\rangle}\frac{r_1}{r_2}H\right)}{\sqrt{1+r_1 H\frac{1+\frac{\langle m\rangle}{\langle p\rangle}\frac{r_1}{r_2}H}{r_1+r_2}}\sqrt{1+r_2 H\frac{1}{r_1+r_2+\frac{\langle p\rangle}{\langle m\rangle}r_2}}} \qquad \text{(S37)}$$

Therefore, in the case of autorepression, r can be negative (when $H > \langle p\rangle r_2/(\langle m\rangle r_1)$).

Case 5: Extrinsic noise at translation level

Suppose the extrinsic noise arises from the heterogeneity at the translation level. Define k_2 and γ_2 as the characteristic translation rate of each cell, and assume that they are independent of each

other. Again using the law of total variance and covariance and equations S25, the covariance matrix components can be written as:

$$\sigma_m^2 = \frac{k_1}{\gamma_1}$$

$$\sigma_p^2 = \frac{k_1}{\gamma_1}\langle k_2\rangle\left\langle\frac{1}{\gamma_2}\right\rangle + \frac{k_1}{\gamma_1}\langle k_2^2\rangle\left\langle\frac{1}{\gamma_2(\gamma_1+\gamma_2)}\right\rangle + \frac{k_1^2}{\gamma_1^2}V_2$$

$$\sigma_{mp}^2 = \frac{k_1}{\gamma_1}\langle k_2\rangle\left\langle\frac{1}{\gamma_1+\gamma_2}\right\rangle \qquad \text{(S33)}$$

where $V_2 = Var[k_2/\gamma_2] = \langle k_2^2\rangle\langle 1/\gamma_2^2\rangle - \langle k_2\rangle^2\langle 1/\gamma_2\rangle^2$. And the correlation coefficient is:

$$r = \sqrt{\frac{1}{\frac{1}{\langle k_2\rangle}\frac{\left\langle\frac{1}{\gamma_2}\right\rangle}{\left\langle\frac{1}{\gamma_1+\gamma_2}\right\rangle^2} + \frac{\langle k_2^2\rangle}{\langle k_2\rangle^2}\frac{\left\langle\frac{1}{\gamma_2(\gamma_1+\gamma_2)}\right\rangle}{\left\langle\frac{1}{\gamma_1+\gamma_2}\right\rangle^2}}}\sqrt{\frac{1}{1+V_2\frac{k_1}{\gamma_1}\frac{1}{\langle k_2\rangle\left\langle\frac{1}{\gamma_2}\right\rangle + \langle k_2^2\rangle\left\langle\frac{1}{\gamma_2(\gamma_1+\gamma_2)}\right\rangle}}} \qquad \text{(S34)}$$

It is straightforward that the second square-root is less than one. In the case of $\gamma_1 >> \gamma_2$, using a derivation of the Cauchy-Schwarz-Buniakowsky inquality: $\langle k_2^2\rangle > \langle k_2\rangle^2$ and $1 \le \langle\gamma_2\rangle\langle 1/\gamma_2\rangle$, the first square-root becomes:

$$r \le \sqrt{\frac{1}{\frac{1}{\langle k_2\rangle}\frac{\left\langle\frac{1}{\gamma_2}\right\rangle}{\left(\frac{1}{\gamma_1}\right)^2} + \frac{\left(\frac{1}{\gamma_1}\right)\left\langle\frac{1}{\gamma_2}\right\rangle}{\left(\frac{1}{\gamma_1}\right)^2}}} \le \sqrt{\frac{\langle\gamma_2\rangle}{\gamma_1}\frac{1}{1+\frac{\gamma_1}{\langle k_2\rangle}}} = r_{constitutive} \qquad \text{(S35)}$$

Therefore, unlike two-state transcription (Case 2) and heterogeneous transcription (Case 3), heterogeneity in translation rates can further reduce the mRNA-protein correlation from that of a constitutive promoter. This offers a plausible cause for the near zero correlation that we have experimentally observed.

18. Goodness of fit of gamma distribution to protein number distributions

Here we show that the gamma distribution provides the best description of protein number distributions over the proteome, compared to several other fitting functions, including the Gaussian distribution, the Poisson distribution, the lognormal distribution, the negative binomial distribution, and the gamma distribution. As seen in Figures 1C-E, the protein distributions are often asymmetric, so that Gaussian distributions do not provide a global fit. The Poisson distribution is also not appropriate because the ratio of variance to mean of single-cell expression levels (Fano factor) is much more than one at high expression levels (Fig. S5, upper middle). Lognormal distribution provides similar fitting shapes to the gamma distribution, but we found

that gamma distributions can provide comparable fitting even at high copy proteins (Fig. S20A, χ^2 = 8.2 for a gamma distribution, whereas χ^2 = 14 for a lognormal distribution in the case of TufA). However, the lognormal distributions fit poorly for low copy protein distributions (Fig. S20B, χ^2 = 2.3 for a gamma distribution, whereas χ^2 = 441 for a lognormal distribution in the case of FadB).

The negative binomial is the mathematically accurate solution to Scheme 1. However, the gamma distribution works better for practical and experimental reasons. We have carried out a systematic and statistical comparison between negative binomial and gamma functions. We found that the gamma distribution can be used to fit 1,009 of 1,018 genes with $p > 0.05$, whereas only 823 proteins can be fit equally well by the negative binomial distribution. Practically, this is because the negative binomial distribution is sensitive to the calibration of fluorescence per protein, i.e. small changes to the calibration result in very different values of a and b in the fit, unlike the gamma distribution.

The two-state promoter model (*14*, *S17*) has been proposed in order to consider transcriptional bursts due to dynamic changes such as chromatin remodeling. While this might be a possible model, previous work has shown that a Poisson description of transcription is sufficient to describe the observed protein production for low expression levels (*S12*). In addition, the two-state model leads to a distribution with four adjustable parameters (*S17*). The cellular fluorescence histograms can be fit better with the resulting distribution function because of the higher degrees of freedom, but we found that the errors of fitting parameters became significantly higher (Fig. S20C), which means that the quality of our data is not enough to determine the four parameters. Thus, we didn't focus on the two-state model, or any other more complicated model, but looked only for the best phenomenological fit.

As the result of the above considerations, we concluded the gamma distribution is the most appropriate fitting function for our data.

To show that the protein number distribution follows gamma distribution in the existence of extrinsic noise, we simulated the distribution with a fluctuating a and b value with a Gaussian, gamma, uniform and lognormal noise. We found that the resulting distributions still can be largely fitted to gamma distributions ($p > 0.05$), if the fluctuations of a and b values are less than 30%. This confirms the validity of gamma distribution at high copy proteins.

19. Number of annotated genes

Over 60% of low expression genes had functional characterizations and full names, as opposed to y genes, when the nomenclature was first developed (*S21*). The addition of new annotations continues, which would further increase this percentage.

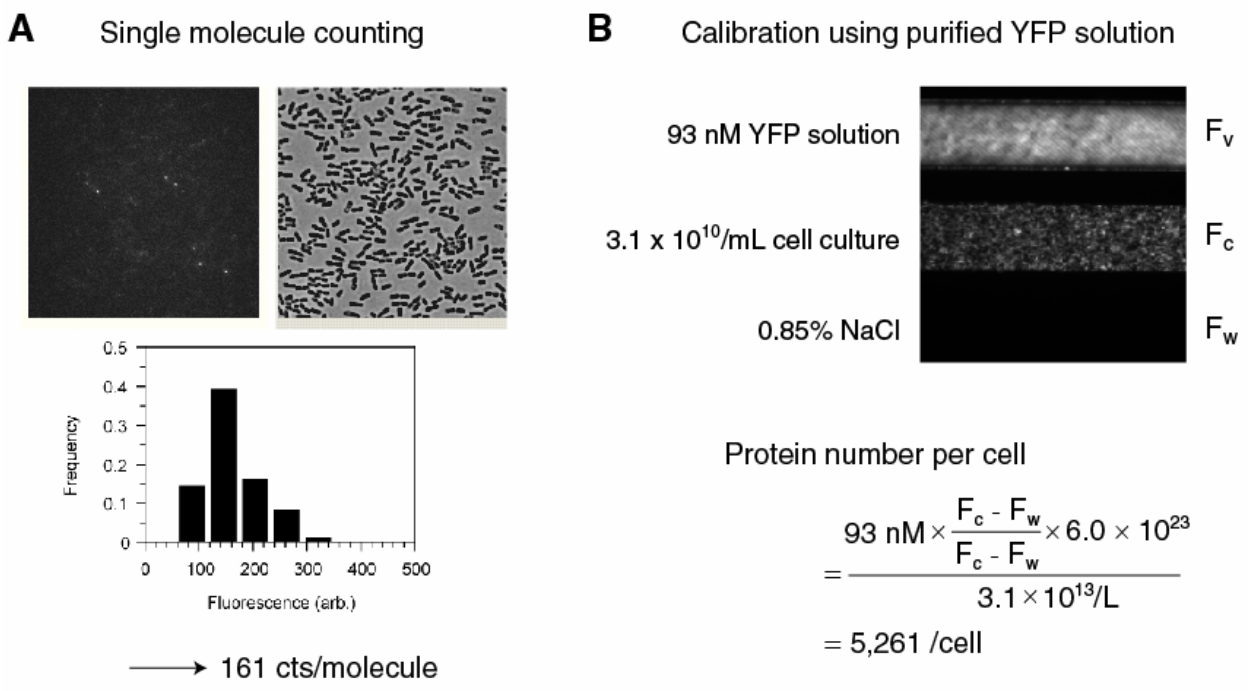

Figure S1. Calibration of single molecule fluorescence.
(A) Fluorescence (left) and phase contrast (right) images of membrane-localized YFP in strain SX4 imaged under high magnification and high laser intensities. The fluorescence from each YFP molecule is recorded in a histogram. The YFP molecule produces 161 ± 105 counts/molecule/average cell volume (mean ± SD, N = 134). This number is used to calibrate the absolute molecule numbers for the entire data set. **(B)** Purified YFP solution with known concentration determined from absorbance (top), AcpP-YFP library strain (middle), and 0.85% sodium chloride solution (bottom) imaged under low magnification and low laser intensities. The average number of YFP molecules per cell for this library strain is estimated to be ~5,261 using a comparison of fluorescence intensity with a solution of known YFP concentration. The number of YFP molecules per cell for AcpP-YFP calibrated by single-molecule imaging (as in (A)) is ~6,432 molecules per cell. The two independent methods to calibrate absolute molecule numbers agree to within a factor of 1.2.

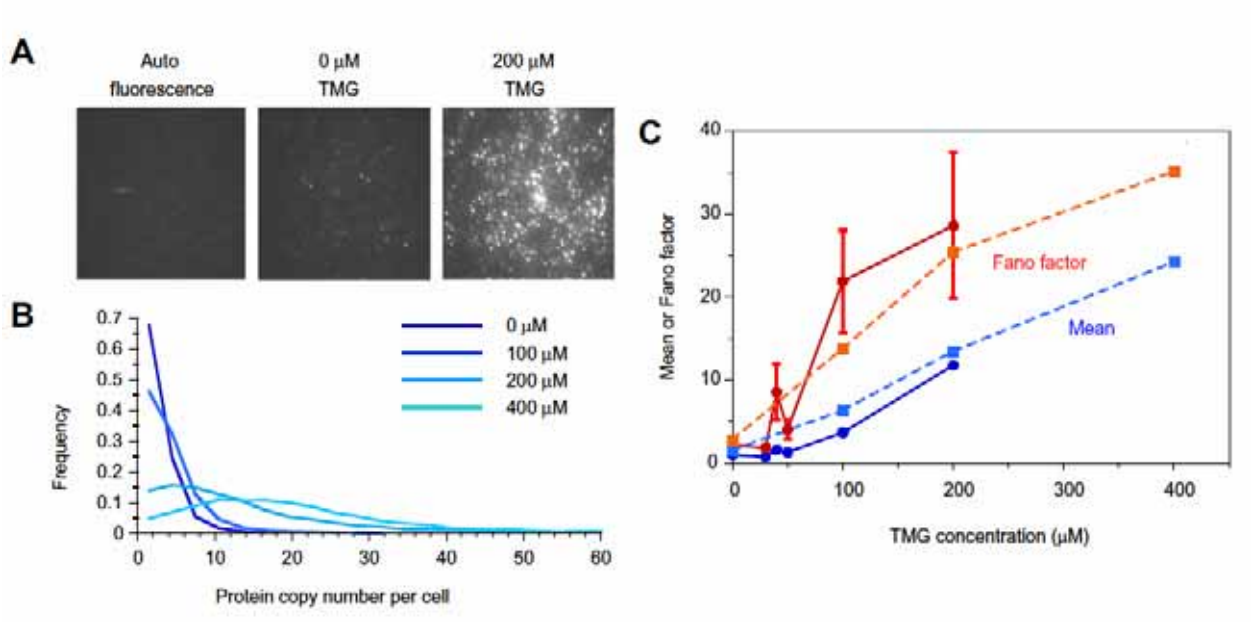

Figure S2. Consistency of the deconvolution method with the single molecule localization method. (A) Fluorescence images of SX701 cells grown with varying levels of TMG. At 0 μM of TMG, there is basal expression of Tsr-YFP. At 200 μM TMG, Tsr-YFP is highly induced. A non-fluorescent strain, BW25993, was used to control for autofluorescence. **(B)** Histogram of expression levels at various TMG concentrations. **(C)** Average count and Fano factor of Tsr-YFP by detection by localization (circle) and by deconvolution (square). Error bars represent SEM. *N* = 24,021 (0 μM), 20,317 (100 μM), 23,605 (200 μM) and 11,723 (400 μM).

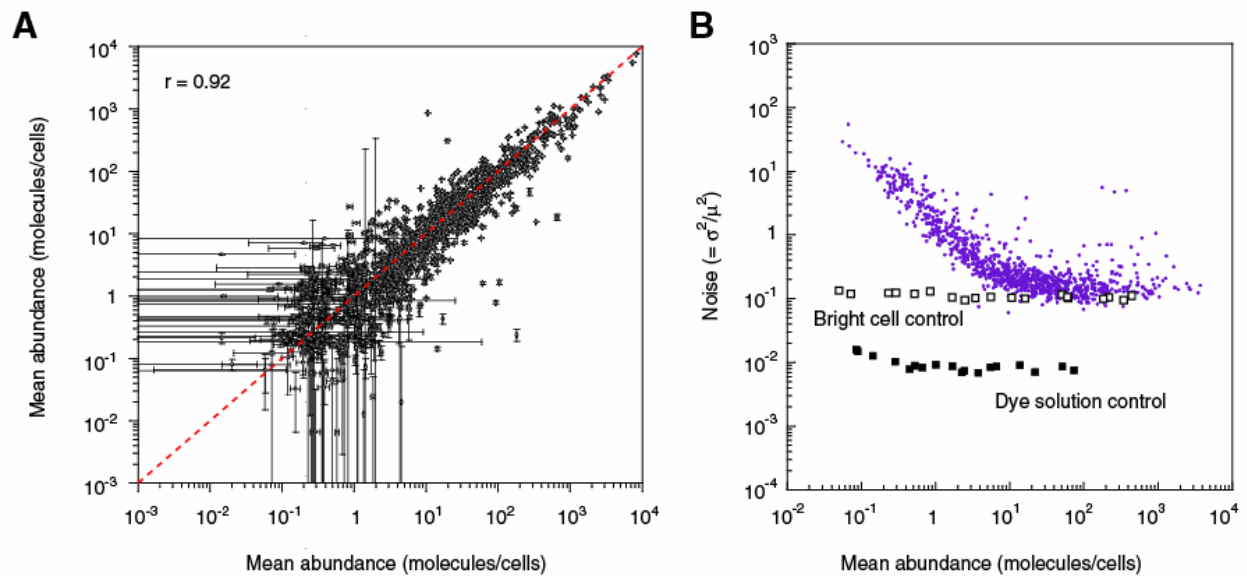

Figure S3. Data reproducibility and instrument noise. (A) Reproducibility of noise measurements of the same strains taken on two separate occasions. Abundance measurements of all strains grown and measured on separate days agree with each other (r = 0.92). Error bars represent SEM, which are calculated by bootstrapping for the data analysis procedure. **(B)** Detection limit of noise measurement. Data from two controls, a bright cell strain (open square) and a fluorescent dye solution (closed square), were taken at various laser intensities to simulate various mean abundance levels. The noise of both systems stayed constant throughout the range of simulated abundance, indicating that our microscope is capable of measuring noise for a wide range of signal levels. This indicates that instrument noise does not limit our measurement of biological noise.

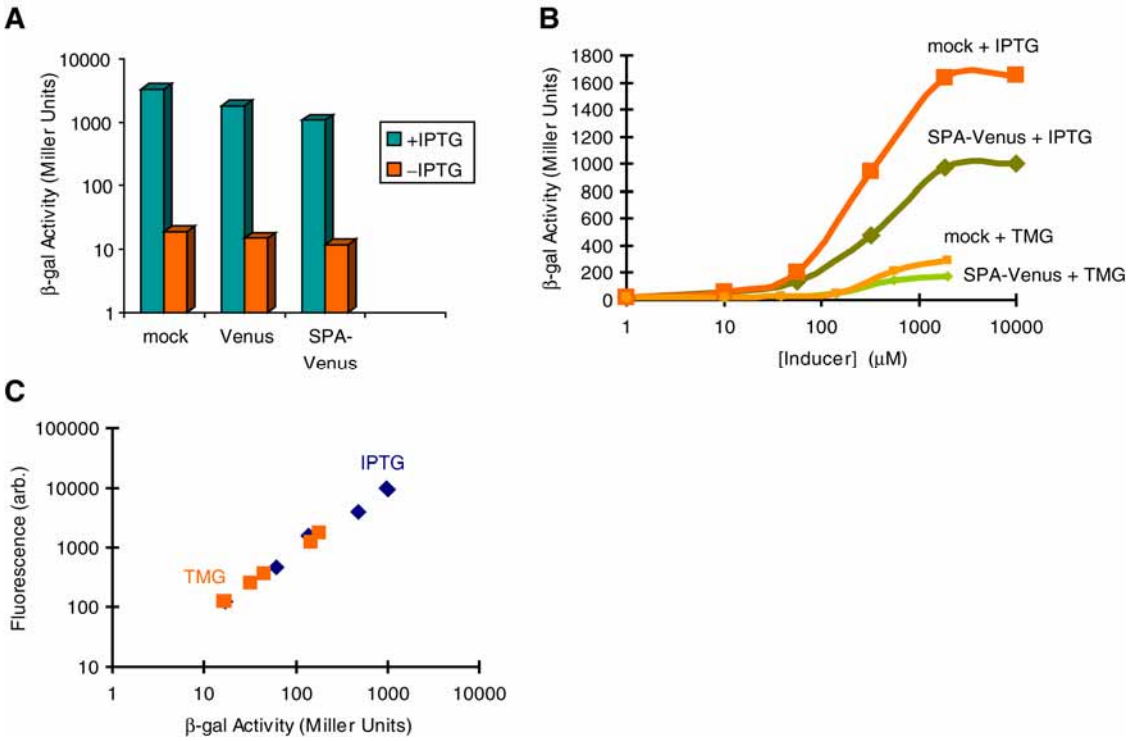

Figure S4. Comparison of the fluorescent reporter assay with the Miller assay. (A) The change in the β-galactosidase activity with and without 1mM IPTG. **(B)** The dependence of β-galactosidase activity as a function of concentrations of inducers, IPTG and TMG. **(C)** The correlation between the β-galactosidase activity and the fluorescence level in the LacZ-SPA-Venus construct.

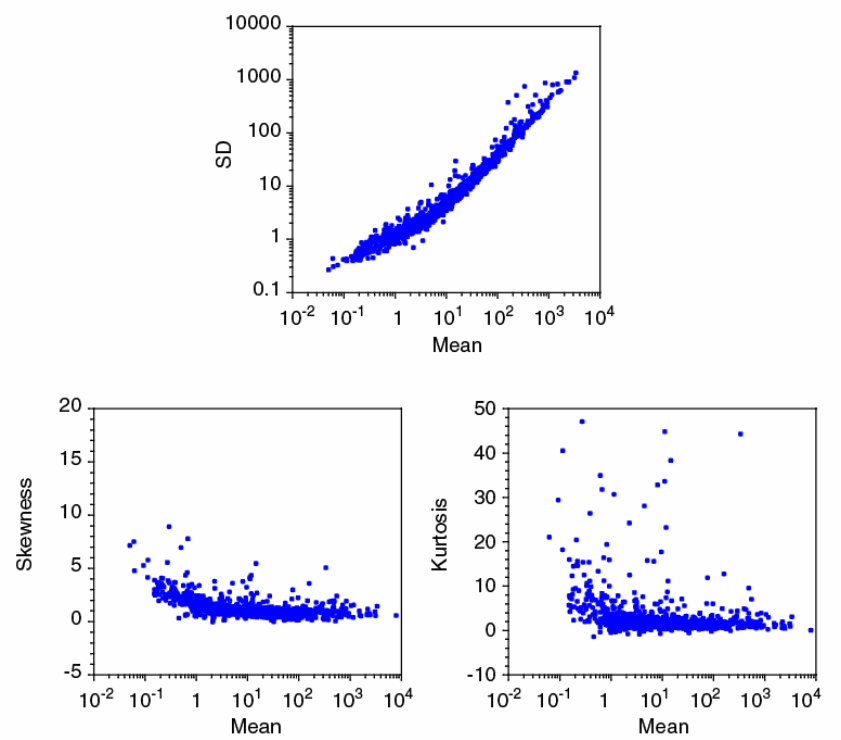

Figure S5. Global relationships of determined parameters.
The relationship between various metrics of expression level and properties of gene expression are plotted, where each point represents the values for one gene. *SD* is the standard deviation. Skewness and kurtosis are calculated for the copy number distribution of each gene.

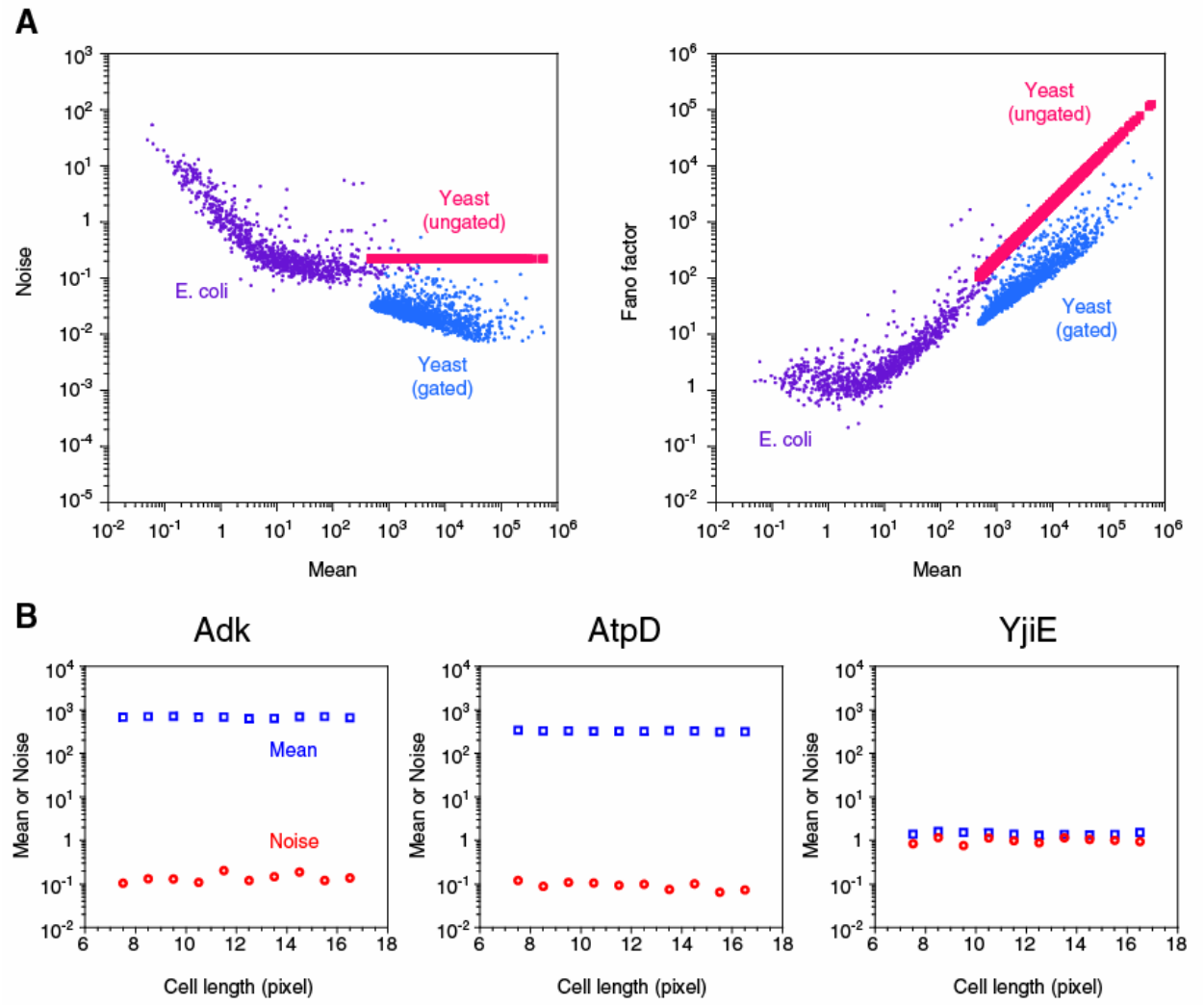

Figure S6. Comparison with yeast. (A) Noise and Fano factor in yeast cell. The protein copy numbers in yeast were roughly estimated by multiplying the flow cytometry counts with a scaling factor (= 10) that is estimated from apparent copy numbers indicated in the western blotting result reported by Newman *et al*. (*S11*). The gated data in yeast was calculated from a data table in their work. The ungated data was estimated from an apparent average of a noise plot in their work. **(B)** Cell length dependence. The mean and noise values of three genes, Adk, AtpD and YjiE were plotted as a function of cell length. For analysis, we gated as a function of cell length and determined the noise parameters.

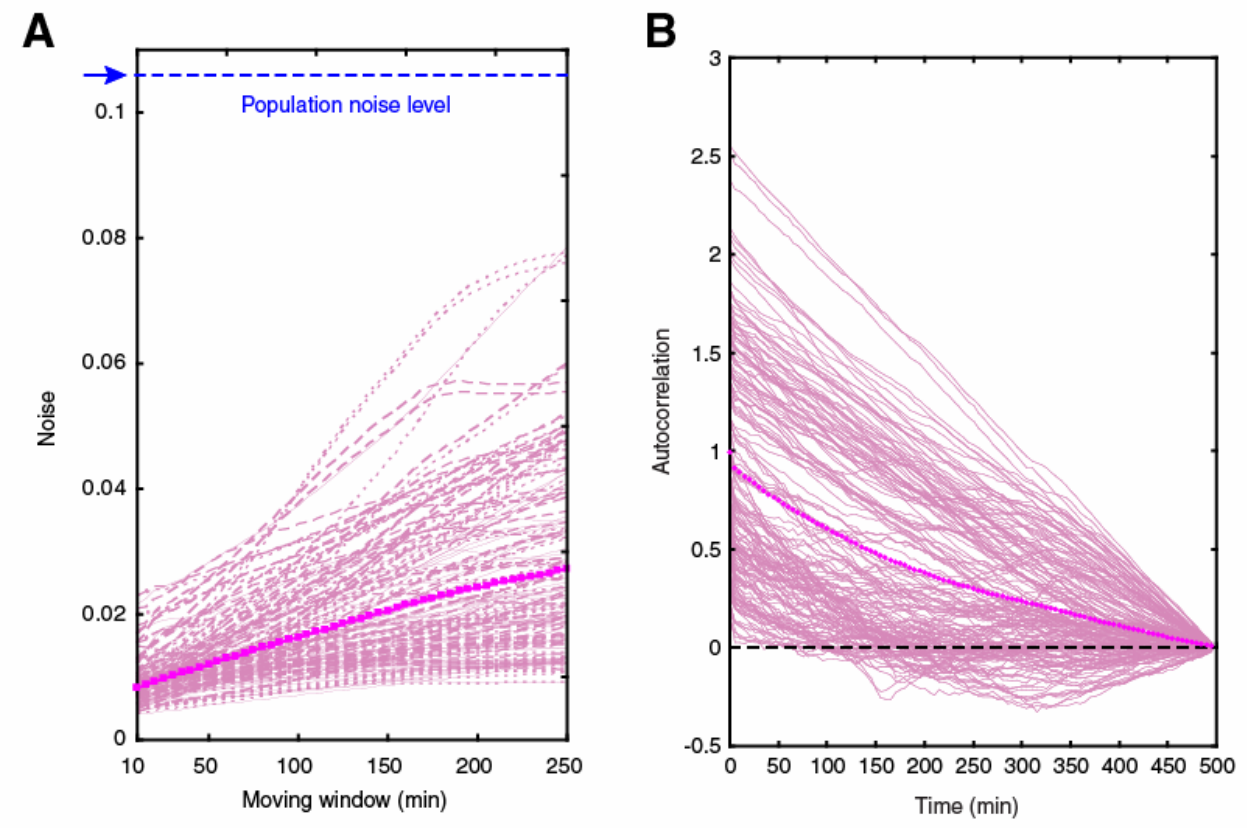

Figure S7. Time propagation of the protein level fluctuation of AcpP. (A) The noise at a certain window size was calculated as the average of noises in a time trace for all the possible window positions. The averaged noise from all the traces, indicated by the thick dashed line, was fitted to an exponential curve: $\eta^2 = \eta^2_{\text{cell-to-cell}} - (\eta^2_{\text{cell-to-cell}} - \eta^2_0)\exp(-t/\tau)$ with a time decay $\tau =$ 1350 minutes and an intercept $\eta^2_0 = 0.008$. The intercept should result from the intrinsic noise (*S18*). **(B)** Autocorrelation function of the time traces in single cells. The averages for all the time traces are indicated by thick dashed lines. Most of traces showed linear decay as oppose to exponential decay, suggesting that the correlation time is slower than the acquisition time of the time traces (= 500 min).

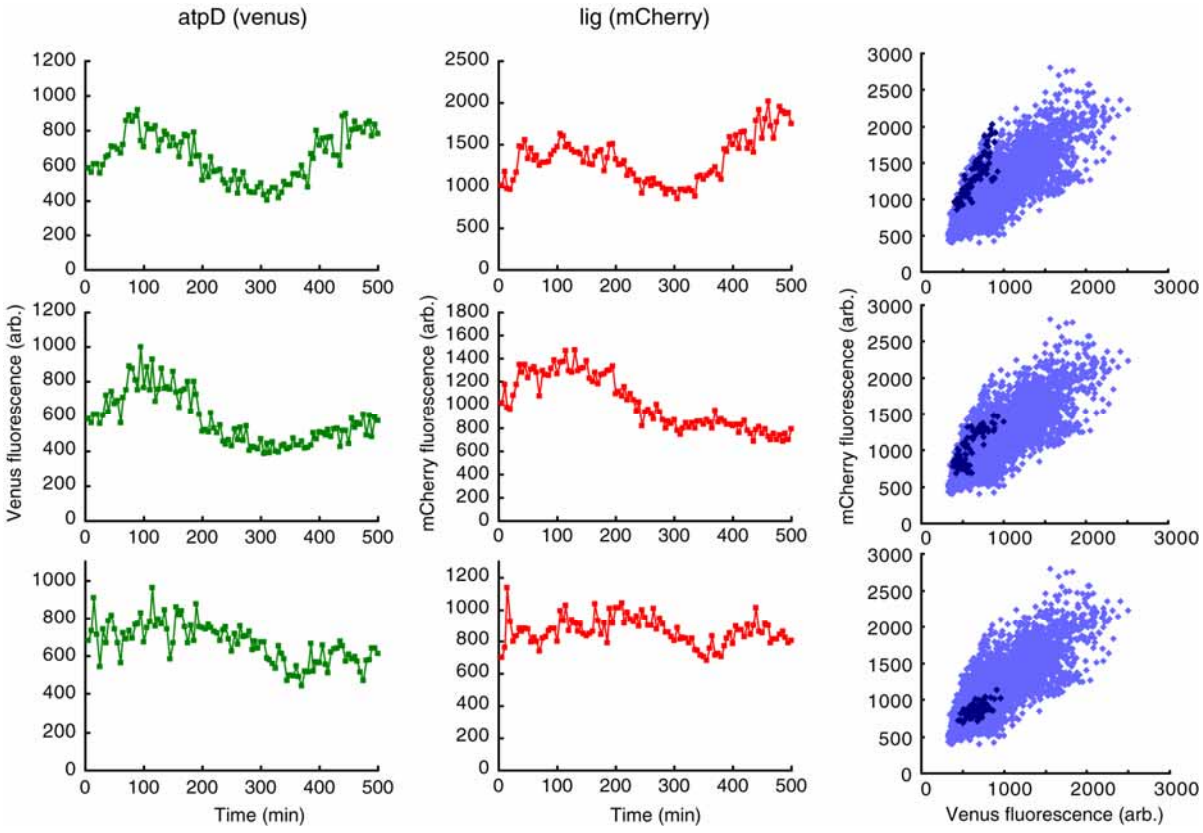

Figure S8. Time traces of different two protein levels in a single cell. The levels of two different proteins (AtpD and Lig) in a single cell were measured (first two columns). Three typical sample traces are shown in the display (top, middle and bottom), all indicating that the two different protein levels are correlated. The figures in the right column show correlation plots for these two protein levels over a timescale of 500 min. The data collected within a single cell lineage are shown in dark blue dots, whereas the data from the entire population are shown in light blue dots. These plots further demonstrate the difference between temporal noise in a single cell and population noise.

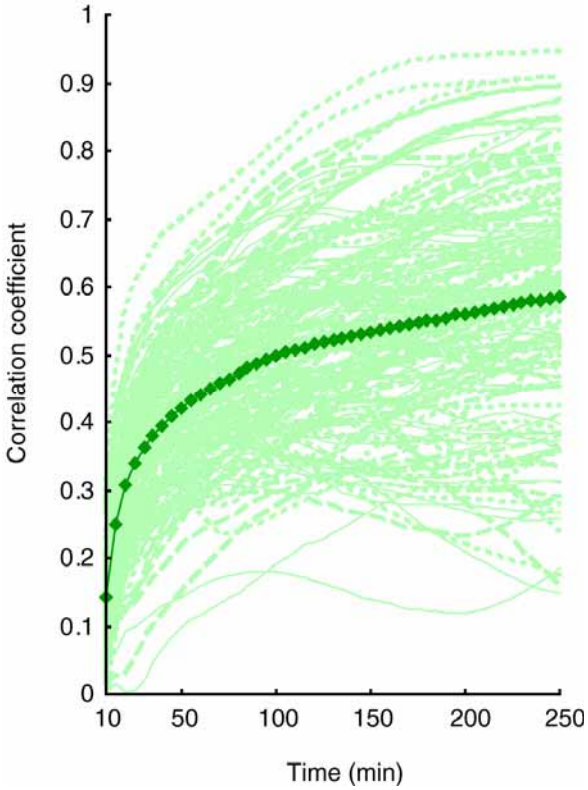

Figure S9. Time propagation of correlation coefficient of two protein levels. The correlation coefficient of two protein levels (AtpD and Lig) was calculated with different time window sizes. The thin dotted lines are correlation traces from all the cells in a population, and the thick dotted lines is their average.

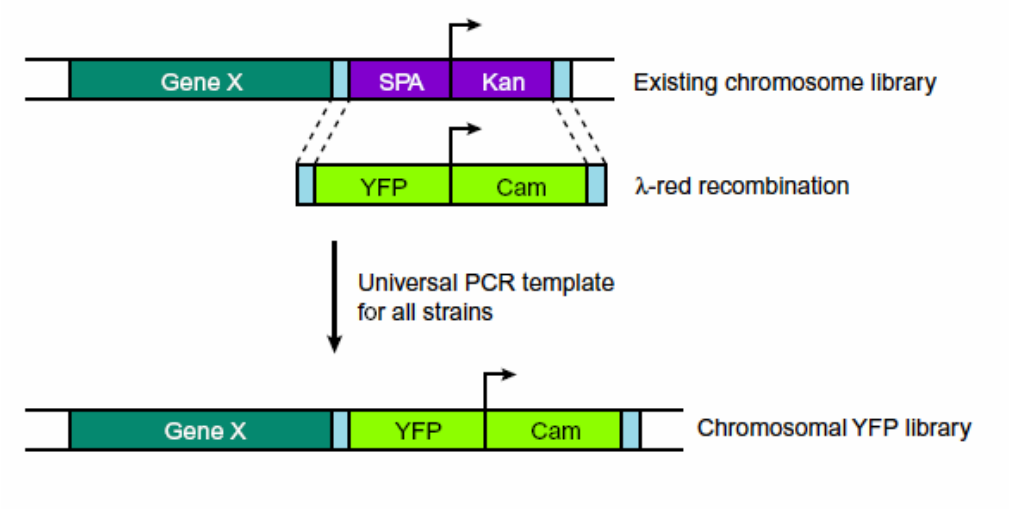

Figure S10. Library construction. The chromosomal YFP-fusion protein library was created via λ-RED recombination (*S19*) using a universal primer targeting the sequential peptide affinity (SPA) tag-kanamycin resistance sequence of an existing library (*S1*)**.**

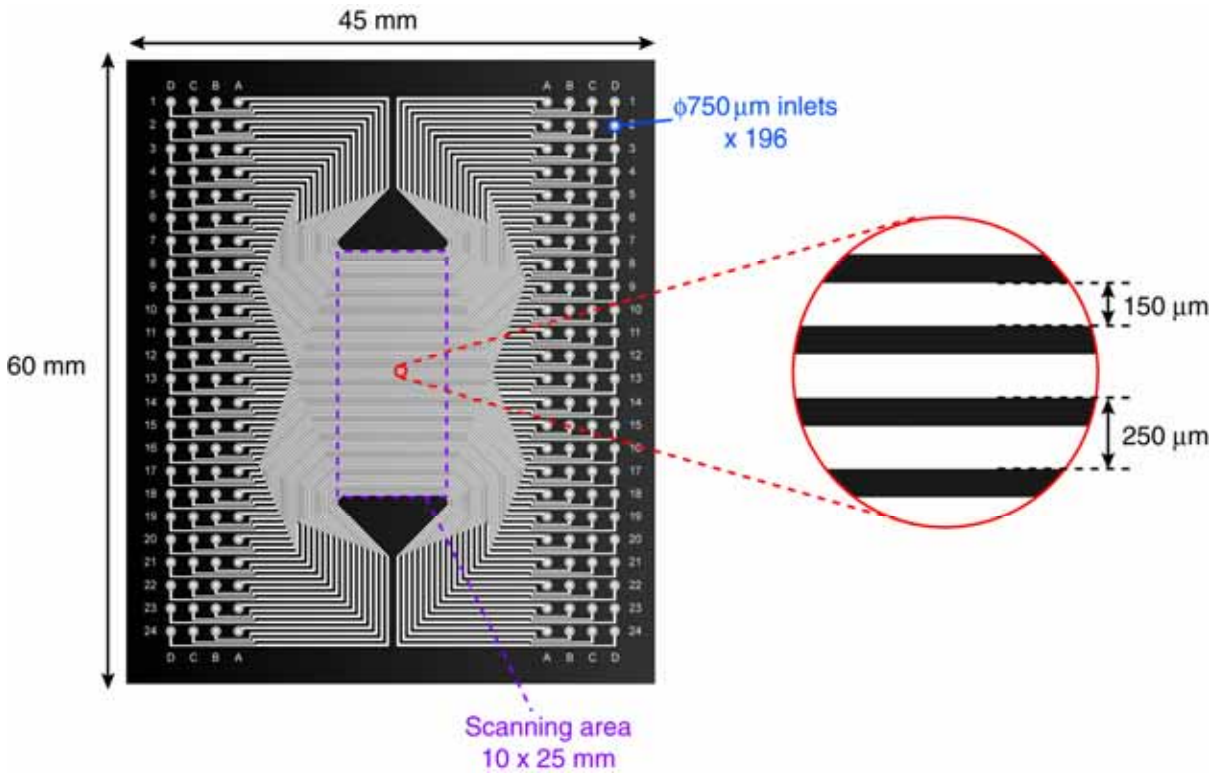

Figure S11. Microfluidic chip design. The micropattern integrates 96 independent microfluidic channels within a 45 mm × 60 mm area. The circles on the right or left side are the inlet/outlet of the channel and ϕ0.75 mm holes are punched through PDMS replicas at these locations. The height of the channel is 25 µm.

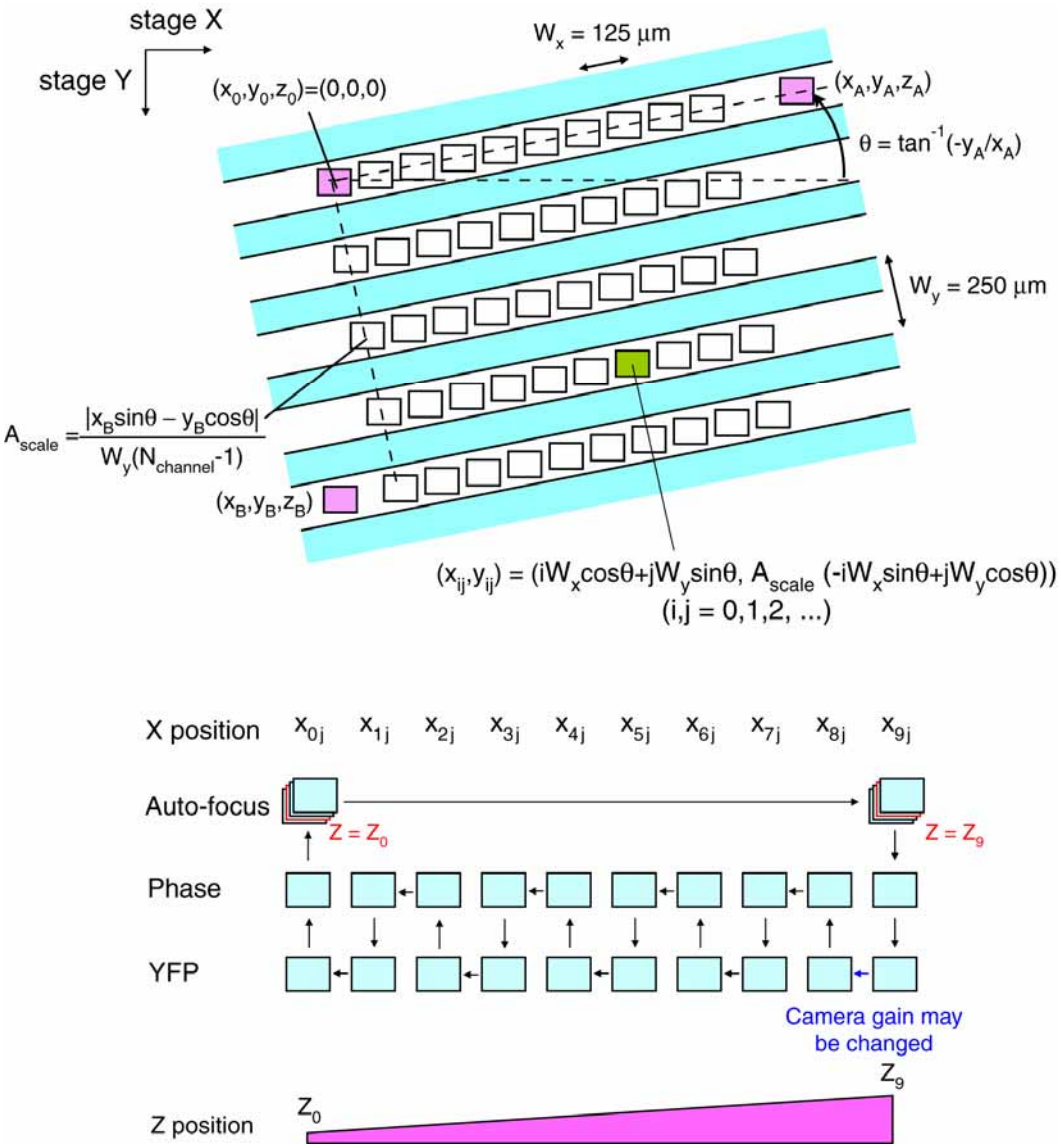

Figure S12. Scanning algorithm. The tilt and orientation of the microfluidic chip against the stage XYZ coordinates are calibrated by three-point calibration (pink squares in the upper figure). The XY positions for scanning were calculated from the parameters obtained by the calibration. At the beginning and end point of the channels, the image was auto-focused under a phase-contrast illumination to provide the Z position profile along the channel assuming a linear landscape (see the lower figure). Subsequently, phase contrast and fluorescent images are taken in turns at the given XYZ positions along the channel.

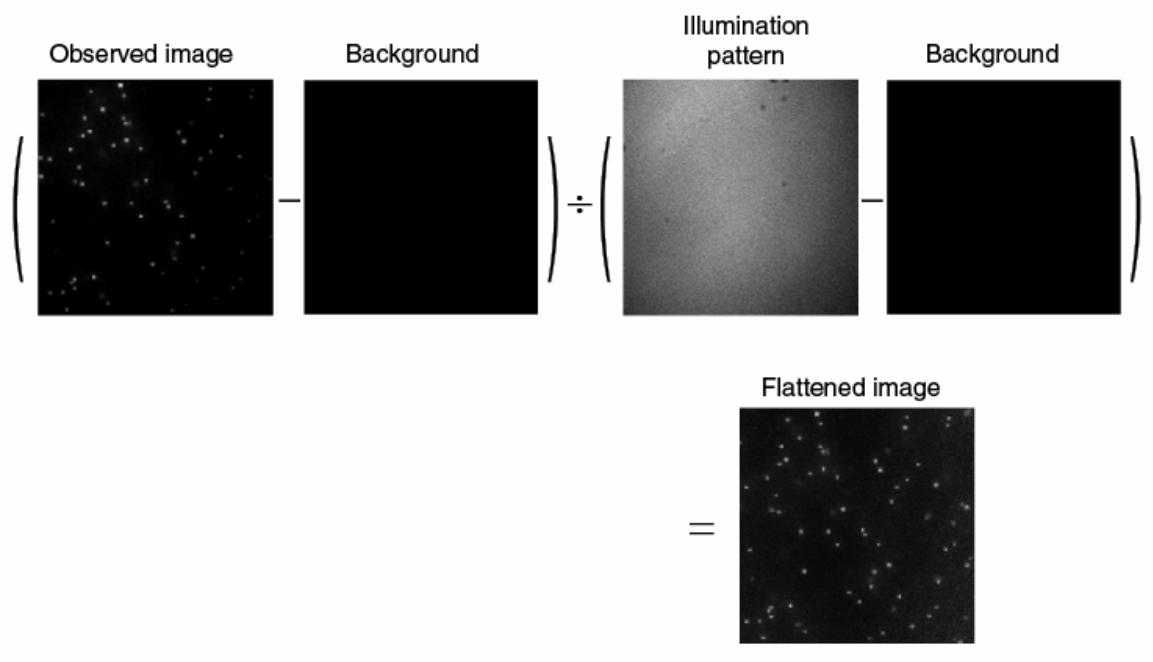

Figure S13. Compensation of laser illumination pattern. We compensated for the heterogeneity of laser distribution in the image field. We obtained the laser illumination pattern by imaging fluorescein dye solution that was injected into the channel. Flattened images were obtained by dividing the background-subtracted observed images by the background-subtracted laser intensity distribution.

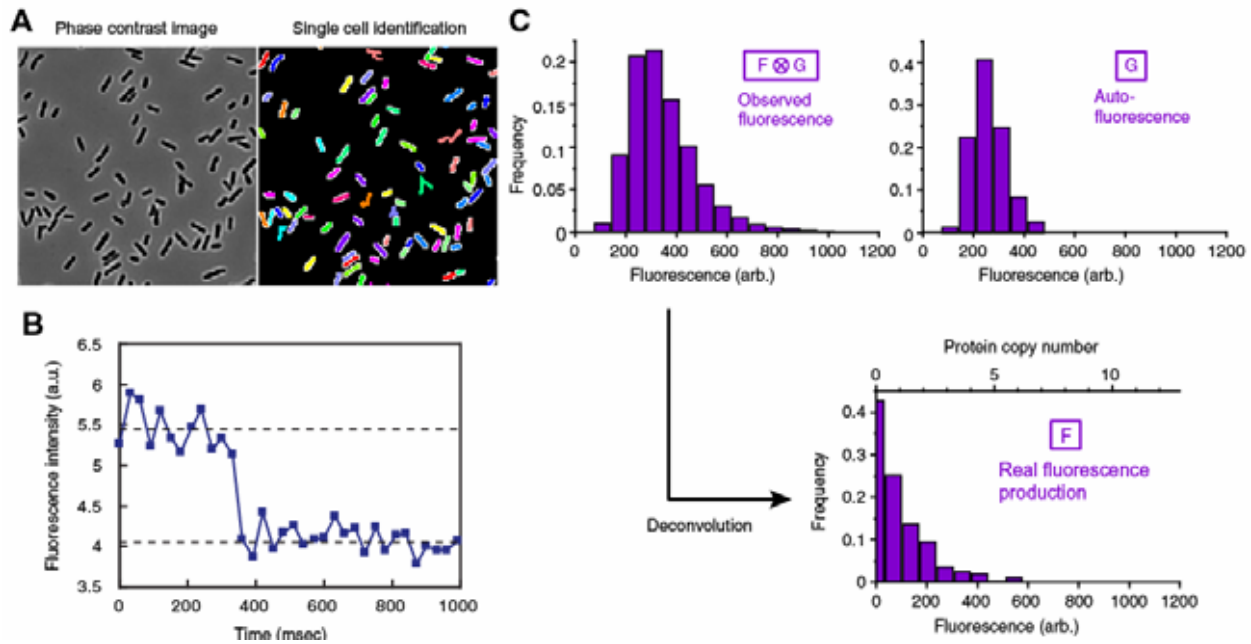

Figure S14. Single cell identification and autofluorescence deconvolution. (A) Automated image analysis to obtain single-cell distributions from the obtained phase-contrast and fluorescence image. The phase contrast images were reduced to binary images through image filters to find distinct particles that correspond to cells (the different colored objects in the right image). The particles were then filtered by particle size, area and shape to identify single cell boundaries (highlighted in white). The fluorescence count was integrated for the entire area for each cell. **(B)** Fluorescence time trace of a single YbdG-Venus molecule in an *E. coli* cell, showing abrupt photobleaching. **(C)** Deconvolution of cell auto-fluorescence background allows expression profiling with single molecule sensitivity. The net protein copy number distributions (F) were calculated by deconvoluting the measured fluorescence histogram (F⊗G) with cell auto-fluorescence histogram (G). The protein copy number per average cell volume, or the concentration, was determined as described in the main text and in the supplementary methods.

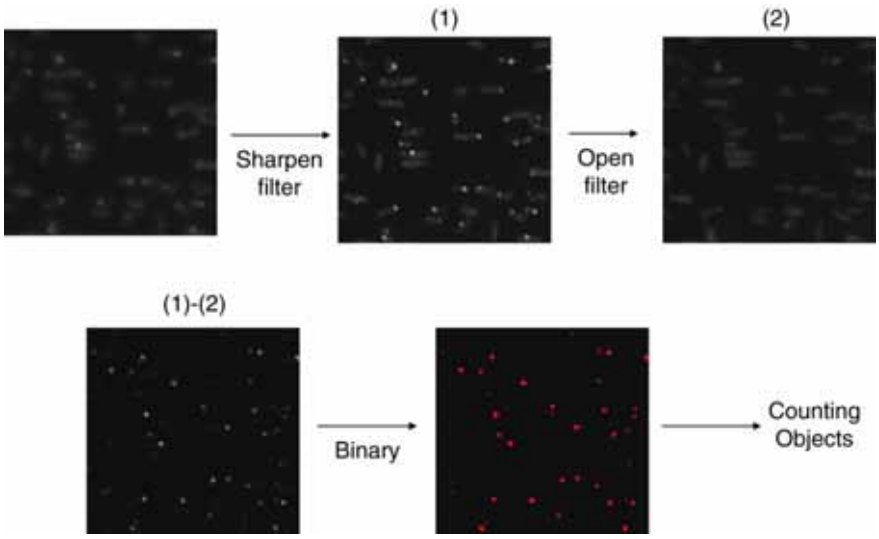

Figure S15. Characterization of point-like localization. To characterize point-like localization, we analyzed a differential image from the recorded image before and after an open filtering ((1) and (2)). The spot frequency was quantified by counting spots in a binary image.

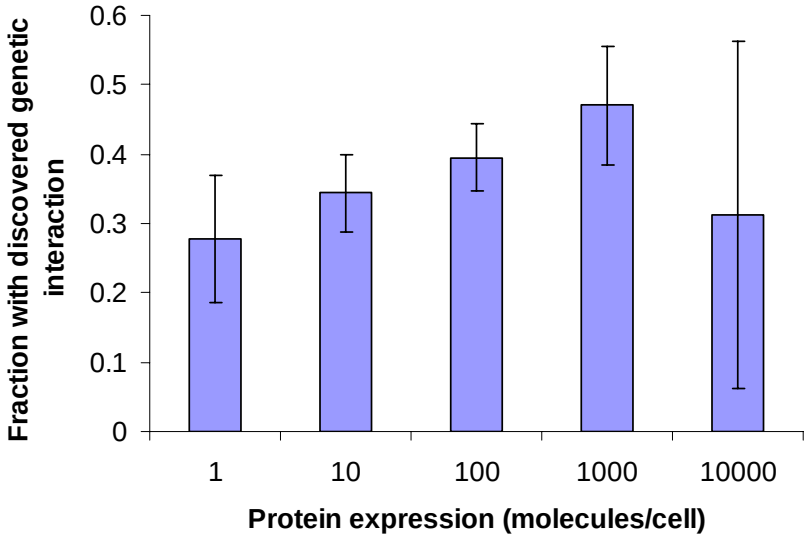

Figure S16. Dependence of genetic interactions on expression level
The fraction of measured proteins with at least one genetic interaction discovered in a limited screen by Butland, *et al* (*S20*) is shown as a function of protein expression level. The probability of finding a genetic interaction is weakly dependent on protein expression. Error bars are the inverse square root of sample size.

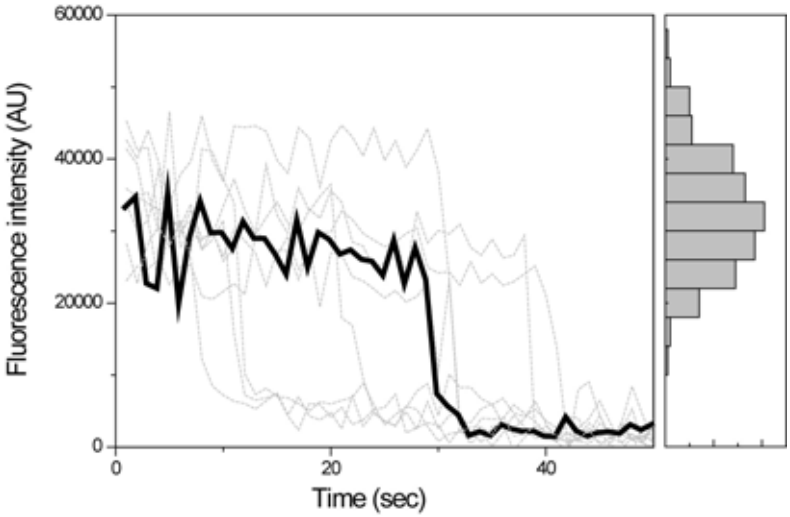

Figure S17. Single-step photobleaching traces.
Photobleaching traces for ten diffraction-limited spots in the FISH samples are shown here (grey). One representative trace is shown in the black thick line. Each trace shows abrupt single-step photobleaching, which is a signature of a single fluorophore. Photobleaching is a stochastic process and occurs random time points. The distribution of fluorescence intensities from a single fluorophore is shown on the right panel (N = 210). The variation in intensity is 27%, which is low enough to allow digital counting of mRNA (See also Fig. S18).

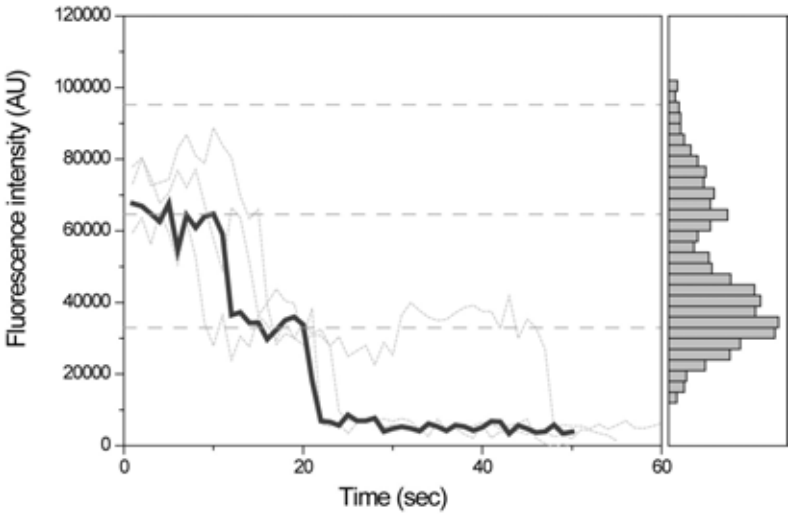

Figure S18. Double-step photobleaching traces.
Photobleaching traces for four diffraction-limited spots, each consisting of two hybridized probes, are shown here (grey). One representative trace is shown in the black thick line. The distribution of fluorescence intensities from a mixture of single fluorophores and two fluorophores is shown on the right panel (N = 1,700).

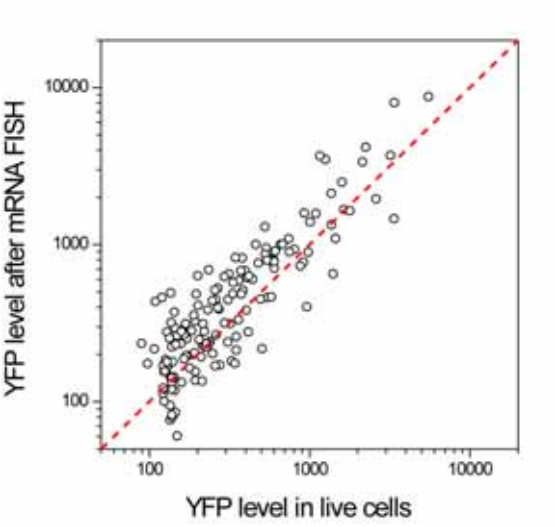

Figure S19. YFP level before and after FISH.
The mean YFP fluorescence level for each strain is plotted as a circle. The YFP fluorescence measured in live cells is preserved after the mRNA FISH procedure

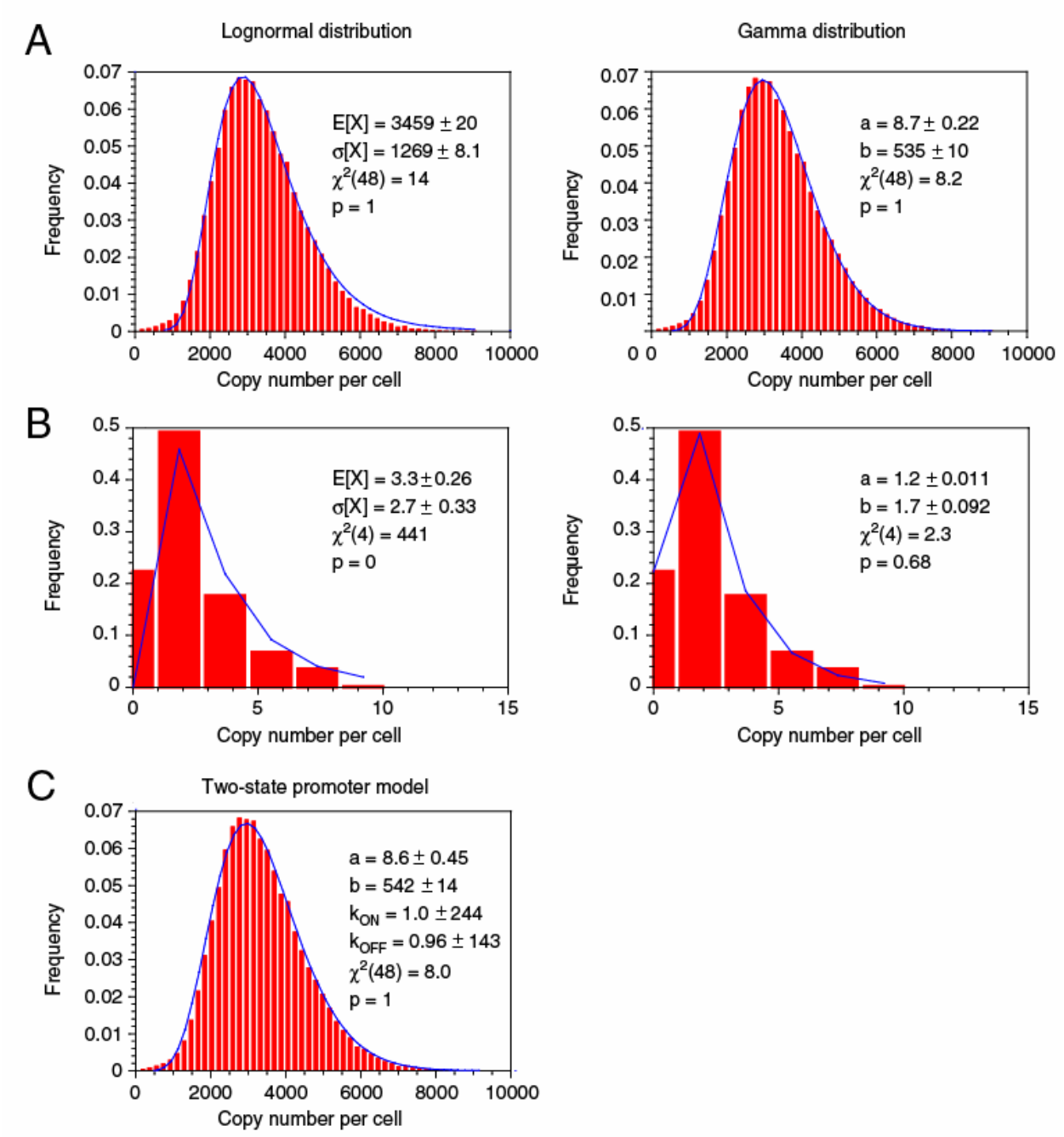

Figure S20. Gamma distribution fitting and lognormal distribution fitting to protein number distribution.

(A-B) The histogram of protein numbers of TufA (**A**; N = 63,446 cells) and FadA (**B**; N = 9,124 cells) are fitted with lognormal (left) and gamma (right) distribution, respectively. Gamma distribution fits well to the histogram over the entire shape including tail regions. The lognormal distribution is given by: $p(x) = (1/(x\sigma(2\pi)^{1/2}))\exp[-(ln(x) - \mu)^2/2\sigma^2]$. **(C)** The histogram of protein numbers of TufA fitted with a two-state promoter model (*S17*). The derived distribution has four adjustable parameters, where a and b is k_1/γ_2 and k_2/γ_1, respectively. k_{ON} and k_{OFF} is the rate

constants of the transition between the active and inactive promoter state, divided by γ_2. k_{ON} and k_{OFF} have substantial error/uncertainty.

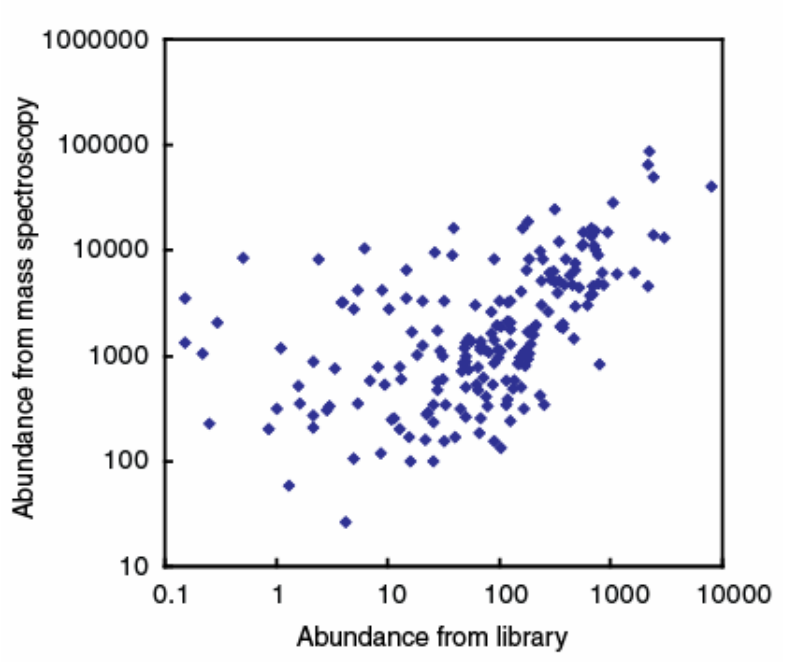

Figure S21. Comparison of our library data with a mass spectrometry data.
The correlation plot between protein abundance obtained by mass spectroscopy (*S4*) and mean abundance determined by our study is shown. The correlation coefficient between them is 0.58.

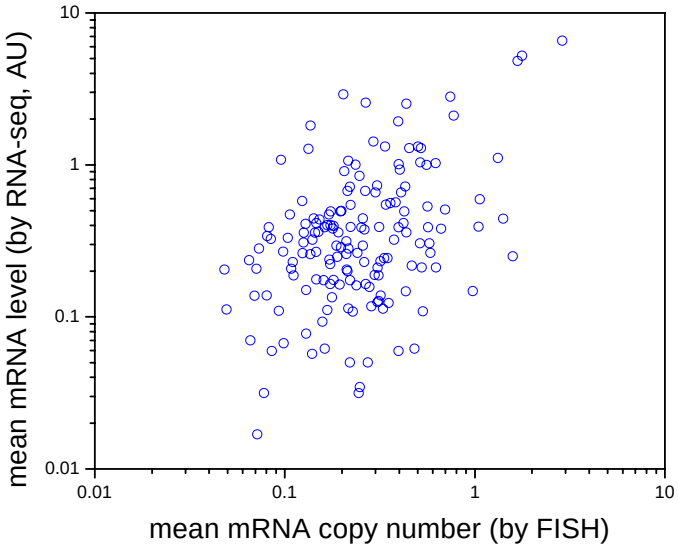

Figure S22. Comparison of mRNA copy number measured by FISH and by RNA-seq. The mean mRNA copy number measured by FISH is correlated with that measured by RNA-seq. The correlation coefficient is $r = 0.51$. We do not expect to see perfect correlation, however, because the strains used for these experiments are different. For FISH experiment, each mRNA is fused to the YFP coding sequence. For RNA-seq, each mRNA is in its native form and has no YFP attached to the 3' end. Since the 3' UTR can affect transcript stability in bacteria, the absolute mRNA copy number may differ in these strains. This does not affect our measurement of protein-mRNA correlation, because the same strains are being used for comparison.

Table S1. The list of strains measured.

Table S2. Correlation coefficients (*r*) and Z-scores (*Z*) between the noise parameters. *Z* scores of more than 3 (indicated by red) represent a significantly larger quantity compared with the whole genome distribution with >99.9% confidence, and Z scores of less than -3 (indicated by blue) represent a significantly smaller quantity.

Table S3. Preference of noise parameters for a subset of strains that relate to a specific biological function. The "essentiality" column represents the fraction of essential genes in the category. The "# of conserved in different organisms" represents the number of organisms that have similar DNA sequence for each gene to E. coli. 12 organisms (Helicobacter pylori 26695, Pseudomonas aeruginosa PAO1, Chlamydia trachomatis A/HAR-13, Haemophilus influenzae 86 028NP, Neisseria meningitidis MC58, Rickettsia prowazekii Madrid E, Borrelia burgdorferi B31, Bacillus subtilis 168, Staphylococcus aureus Mu50, Streptococcus pneumoniae R6, Enterococcus faecalis V583, Mycoplasma genitalium G-37) were examined to determine the conserved number using a homology search (http://cmr.jcvi.org/). The table has three parts. The first shows the Z score of the average parameter ranks within a category. The second shows the actual median parameter values. The third shows the Z score of the standard deviation of the parameter ranks within a category.

Table S4. Burst frequency (*a*) and size (*b*) of library strains compared to the steady-state noise (α and β)

Strain	Observed *a* (measured α)	Observed *b* (measured β)
CorA	11.4 (5.3 – 13.9)	3.7 (5.9 – 6.5)
YbdG	5.9 (2.5 – 5.5)	2.1 (4.7)
YcjF	3.75 (1.77 – 5.58)	2.3 (5.2 – 5.58)

Table S5. List of previously identified genetic interactions compared with expression level
Statistically significant genetic interactions from a screen of 39 × 4,000 double deletion mutants in reference (*S20*) and corresponding expression levels are listed. Of 448 genes we measured to have expression below 10 molecules/cell, 143 were found to have a genetic interaction with a |Z score| greater than 3.

Table S6. List of *a* and *b* values and other parameters determined in this work.
a and *b* values are calculated as $a = \mu^2/\sigma^2$ and $b = \sigma^2/\mu$, respectively, using the mean, μ, and standard deviation, σ, of the histograms.

References

S1. G. Butland *et al.*, *Nature* **433**, 531 (Feb 3, 2005).
S2. J. C. McDonald, G. M. Whitesides, *Acc Chem Res* **35**, 491 (Jul, 2002).
S3. M. Bon, S. J. McGowan, P. R. Cook, *FASEB J* **20**, 1721 (Aug, 2006).
S4. P. Lu, C. Vogel, R. Wang, X. Yao, E. M. Marcotte, *Nature biotechnology* **25**, 117 (Jan, 2007).
S5. P. M. Sharp, W. H. Li, *Nucleic acids research* **15**, 1281 (Feb 11, 1987).
S6. P. J. Choi, L. Cai, K. Frieda, X. S. Xie, *Science* **322**, 442 (Oct 17, 2008).
S7. A. M. Femino, F. S. Fay, K. Fogarty, R. H. Singer, *Science* **280**, 585 (Apr 24, 1998).
S8. H. Maamar, A. Raj, D. Dubnau, *Science* **317**, 526 (Jul 27, 2007).
S9. J. Yu, J. Xiao, X. Ren, K. Lao, X. S. Xie, *Science* **311**, 1600 (Mar 17, 2006).
S10. J. Elf, G. W. Li, X. S. Xie, *Science* **316**, 1191 (May 25, 2007).
S11. J. R. Newman *et al.*, *Nature* **441**, 840 (Jun 15, 2006).
S12. L. Cai, N. Friedman, X. S. Xie, *Nature* **440**, 358 (Mar 16, 2006).
S13. N. Friedman, L. Cai, X. S. Xie, *Physical review letters* **97**, 168302 (Oct 20, 2006).
S14. J. Paulsson, M. Ehrenberg, *Physical review letters* **84**, 5447 (Jun 5, 2000).
S15. K. A. Datsenko, B. L. Wanner, *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640 (Jun 6, 2000).
S16. J. Paulsson, *Nature* **427**, 415 (Jan 29, 2004).
S17. V. Shahrezaei, P. S. Swain, *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17256 (Nov 11, 2008).
S18. N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, M. B. Elowitz, *Science* **307**, 1962 (Mar 25, 2005).
S19. D. Yu *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5978 (May 23, 2000).
S20. G. Butland *et al.*, *Nature methods* **5**, 789 (Sep, 2008).
S21. K. E. Rudd, *Microbiology and Molecular Biology Reviews* **62**, 985 (Sept, 1998).
S22. M. Acar, A. Becskei, A. van Oudenaarden, *Nature* **435**, 228 (March, 2005).