

A simple comparison of the extrinsic noise in gene expression between native and foreign regulations in *Escherichia coli*

Bei-Wen Ying^{1,*}, Shigeto Seno², Hideo Matsuda² and Tetsuya Yomo^{3,*}

¹ Faculty of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

² Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

³ Institute of Biology and Information Science, East China Normal University, 3663 Zhong Shan Road (N), Shanghai 200062, China

*Corresponding authors

BWY ying.beiwen.gf@u.tsukuba.ac.jp

TY yomo@sei.cc.ecnu.cn

Abstract

Living cells reorganize their gene expression through regulatory machineries in response to external perturbations. The contribution of the regulation to the noise in gene expression is of great interest. In this study, we evaluate the contribution of both native and foreign regulations to the extrinsic noise in gene expression. We analyzed the gene expression data of a mini-library containing 70 genetic constructs of 136 clones into which the *gfp* gene had been chromosomally incorporated under the control of either native or foreign regulation. We found that the substitution of native by foreign regulation, *i.e.*, the insertion of the P_{tet} promoter, triggered a decrease in the extrinsic noise, which was independent of the protein abundance. The reanalyses of varied genomic data sets verified that the noisy gene expression mediated by native regulations is a common feature, regardless of the diversity in the genetic approaches used. Disturbing native regulations by a synthetic promoter reduced the extrinsic noise in gene expression in *Escherichia coli*. It indicated that the extrinsic noise in gene expression caused by the native regulation could be further repressed. These results suggest a tendency of released regulation leading to reduced noise and a linkage between noise and plasticity in the regulation of gene expression.

Keywords: gene expression, extrinsic noise, gene regulation, synthetic promoter, genomics, transcriptome

Introduction

Noise in gene expression has been considered universal in living organisms [1,2,3,4]. This stochasticity in gene expression was proposed and demonstrated to be at least partly beneficial for the adaptation and sustenance of living cells [2,5-7]. Because the biological impact of the noise in gene expression is evident, there has been an increased effort to obtain a snapshot of the global overview of protein abundance and fluctuations, particularly in yeast and bacteria [8-10]. Genome-wide studies provided evidence of the fluctuating native regulations [8-10].

Regulation of gene expression in bacteria is generally known as the operon and regulon [11-13]. Such genetic architectures may provide efficiency in the gene regulation responsive to external perturbations [11,12,14,15]. The synchronized fluctuation of co-expressed genes is hypothesized to result in an equal amount of correlated transcripts (mRNA) and is potentially beneficial for the fitness and evolution of the cells [16-20].

The native regulatory process in living cells offers both fluctuation and efficiency to gene expression. An increased level of the complexity in regulatory processes may not only improve the regulatory efficiency but also amplify the stochasticity in regulatory process. Previous studies proposed a relationship between the noise that causes the fluctuation of expression level and the plasticity that reflects the capacity of the expression changes [8,21,22]. This is somehow linked to the cellular function of the genes [23]. Regulation complexity, such as promoter architecture, has been proposed to be one of the main contributors to the noise and plasticity in gene expression [24-26]. These reports explained to a great extent why the regulation requires noise. However, whether the fluctuation of native regulation is large or small and whether the replacement of native regulation by simple regulation would amplify or reduce the noise in gene expression is still uncertain.

As known, the noise in gene expression is divided into intrinsic and extrinsic derivatives [1]. Variation in the regulation and chromosomal locations will obviously lead to different magnitudes of noise in gene expression, particularly intrinsic noise. In this study, we focused on extrinsically derived noise, which is independent of protein abundance and represents a general feature of fluctuating regulations. As a pilot survey, we used a previously constructed mini-library of *E. coli* strains in which the target genes, which are nonessential but required under certain conditions for cell growth, were replaced by a reporter gene, *gfp*, either with or without a foreign promoter, P_{tet} [27]. The protein abundance and the noise in the native and P_{tet} regulated *gfp* constructs reflected the gene expression that is controlled by native and foreign regulations, respectively. A reduction in the extrinsic noise was observed when the native regulations were disturbed through the insertion of the P_{tet} promoter. This reduction indicated that fluctuating native regulations largely contribute to the extrinsic noise in gene expression

and that the evolved regulatory architecture in *E. coli* most likely balances between efficiency and plasticity.

Material and methods

FCM data analysis

The flow data sets, containing 70 genetic constructs of 136 clones [27], were used and analyzed. A narrowly gated FSC data set could be used instead of the normalised FITC/FSC values. A cross-comparison between the gated and total *E. coli* cells in FCM analyses had been performed previously [8,9,28-31]. The results showed that a reduced magnitude of extrinsic noise was observed in the gated cells compared to the total cells; this finding is supported by the work of Silander *et al.*'s [8,9,28-31]. Consequently, one per cent of the total number of cells (events) from both the highest and the lowest values of the green fluorescence intensity were removed to eliminate unreliable rare events, such as, the systematic error resulting from events that occurred at the extremes of the instrument's range [8,9,28-31]. After subtracting background fluorescence using the native strain (MDS42), the mean and standard deviation of the GFP concentrations were calculated using the data with positive values. The *CV* was defined as the standard deviation divided by the mean [1,8,10,30,32]. For each genetic construct, the statistical values were averaged over multiple data sets. The native and foreign regulations were analyzed separately, and data below the detection limit were removed.

Noise calculation and theoretical fitting

The CV^2 of the intrinsic and extrinsic noise (CV_{int}^2 and CV_{ext}^2 , respectively) were analytically obtained based on the total noise (CV_{tot}^2) by fitting the data to the following equation.

$$CV_{tot}^2 = CV_{int}^2 + CV_{ext}^2 = a/\langle\text{GFP}\rangle + b$$

where a represents the coefficient for the intrinsic noise, b represents the magnitude of the extrinsic noise, and $\langle\text{GFP}\rangle$ represents the mean GFP concentration. This relationship has been widely demonstrated and has been commonly used in the studies on biological noise [8-10,32,33]. Good regressions, which were accompanied by a 95% confidence level, were achieved for the three collections. The details on the theoretical background have been largely described elsewhere [8,9,28-31]. In short, the coefficient for the intrinsic noise, a , is based on the Poisson statistics of the chemical reactions in gene expression, whereas the magnitude of the extrinsic noise, b , is based on a constant independent of the protein expression levels. Thus, the total noise decreases with decreasing expression level until a certain level of intrinsic noise is reached; subsequently, the total noise becomes constant due to the extrinsic noise. This estimation agrees well with standard dual-color experiments [8]. We used custom software with nonlinear regression functions (nls) written in R for the curve fitting. We used the mean GFP

concentration, the mean GFP expression and the mean GFP abundance as the values of <GFP> in our data sets, the data set published by Silander *et al.* and the data set published by Taniguchi *et al.*, respectively.

Transcriptome and proteome data sets

The transcriptome data sets for MDS42 and MG1655 growing in minimal medium were acquired from the raw data deposited in NCBI's Gene Expression Omnibus under the GEO Series accession number GSE33212. The proteome data sets were obtained from other publications: the mean values and noise (or SD) of the protein abundance were obtained from Table S6 (1,018 genes) of the report by Taniguchi *et al.* [10] and the online data file (1,832 genes) from the report by Silander *et al.* [34].

Genome information and gene categories

The gene names and chromosomal positions (distance from *oriC*) were based on the genome information for MDS42 that has been deposited in the DNA Data Bank of Japan under the DDBJ accession number AP012306. The genes in common (overlaps) among the different data sets were determined based on the genome information of the strains W3110 and MG1655 strains in GenoBase, Japan. The functional essentiality gene categories were determined according to the open access data banks and the related publications, as follows. First, the essential genes were identified through single gene deletion experiments [35] and the PEC data bank (Profiling of *E. coli* Chromosome). Second, the conditionally required genes were selected according to the information on selectable phenotypes from the National BioResource Project of the National Institute of Genetics, Japan (NBRP *E. coli* Strain). A total of 770 genes were deposited as selection markers to determine when mutations occurred. Genes for which phenotypic changes were unknown or non-existent and genes showing phenotypic changes in colony formation were excluded. Only the mutations or deletions of genes that directly influenced the cell growth rate were selected and defined as conditionally required genes. As a result, 339 genes with locus tags in MG1655 (equivalent to 343 locus tags in W3110), which largely contained the genes responsive to the nutritional state or environmental temperature, were identified (Table S1). The rest of the genes were categorized as belonging to the "others" category.

Results and discussion

A summary of genetic constructs reporting native and foreign regulations

A previously constructed mini-library that targeted 37 genes was used (Fig. S1). Most of these genes are related to amino acid biosynthesis and are located at various distances from *oriC*.

These genes were replaced with a reporter gene (*i.e.*, *gfp*) that either remained under the control of the native regulation or was inserted with a foreign promoter, P_{tet} , as previously described [27]. The *galK::P_{tet}-gfp* construct was used as a control because this chromosomal site has been widely studied [29,31,36-39]. Since there is no repressor, *i.e.*, *tetR*, in the genome, the expression of the P_{tet} controlled constructs represented the steady expression capacity of the corresponding chromosomal location. A comparison between these two genetic formats allowed us to determine whether and how the noise in gene expression changed when the native regulation was replaced by a foreign promoter with steady expression. Evaluation of the noise in gene expression was based on green fluorescence intensity.

Decreased noise dependent on the chromosomal location

A chromosomal position bias was observed in the gene expression noise. The gene expression noise (CV^2) was calculated using the FCM measurements acquired previously [27]. The CV was defined as the standard deviation divided by the mean and the noise was evaluated as the square of the CV . In the P_{tet} -oriented constructs, the genes located closer to *oriC* tended to exhibit lower noise (Fig. 1A, left). However, no obvious trend was observed in the native constructs (Fig. 1A, right). The results were consistent to our previous finding that the chromosomal position bias was highly significant in foreign regulations but undetectable under the native regulation [27]. In addition, the results showed that the noise levels (the average CV^2) were different between the two types of regulations, indicating the highly repressed gene expression linked to the larger noise in gene expression under native regulation. Therefore, not only an increase in protein abundance, reported previously [27], but also a decrease in total noise clearly exhibited a chromosomal location dependency, when the native regulations were disturbed through the insertion of a foreign promoter for constitutive expression.

Decreased noise due to a reduction in the extrinsic derivative

Since the total noise in gene expression can be divided into the intrinsic and extrinsic derivatives, in the dependent and independent modes of protein abundance, respectively, which source contributed to more to the decrease in total noise was further analyzed. Dual color methods are commonly used to directly measure the intrinsic and extrinsic noise in a specific promoter [8-10,32,33]. However, in the case of single-color data sets that contain an assortment of regulations, the noise can be estimated through theoretical fitting, according to a general equation that has been previously reported and widely used [8-10,32,33,40]. We found that both the native and the foreign regulations showed an equivalent magnitude of intrinsic noise (Fig. 1B, chain lines), whereas the extrinsic source of noise was quite different (Fig. 1B, broken lines). The substitution of the native regulation with the P_{tet} promoter largely reduced the extrinsic

noise from 0.14 to 0.04. Thus, despite of a similar GFP protein abundance, the noise accumulated with the native regulations was larger than the noise obtained with the foreign regulation. The results demonstrated that the disturbance of the native regulations through the insertion of a foreign promoter with steady expression significantly reduced the extrinsic noise in gene expression. This finding implies that the native regulations that are mediated by the evolved molecular machineries maintain a relatively large magnitude of noise in living cells.

A constant magnitude of extrinsic noise in native regulations

Due to the limited data from our mini-library, we reanalyzed other global data sets from previous reports [8-10] to verify whether the magnitude of extrinsic noise contributed under the native regulation was dependent on the genetic backgrounds or the measuring approaches used in the studies. Various genetic engineering and measuring techniques have been employed to study the noisy gene expression reported by fluorescent proteins (Table 1). Taniguchi *et al.* used the carboxy-terminal fusion of a reporter gene under the control of the native regulation at the natural chromosomal position [8-10], and Silander *et al.* used a very low-copy plasmid reporting the control of the endogenous promoter [34]. The data sets were obtained from the supplementary information that accompanied the corresponding reports. A comparison among all three studies showed that the extrinsic source of the noisy gene expression (CV^2_{ext}) that is mediated by native regulations was nearly constant (approximately 0.10 – 0.15), regardless of the genetic background and method. Thus, the results verified that our limited data set reflected the true noise level.

Because our data set mostly comprised the regulation of starvation responsive genes, we additionally evaluated whether the extrinsic noise varied in accordance with the gene categories (*i.e.*, essential, conditionally required, and others). The essential genes exhibited the lowest level of extrinsic noise, which suggests that the essentiality of a gene function could lead to low noise caused by gene regulation. A correlation between the essentiality and the noise level was observed in the analysis of Silander *et al.*'s data (Fig. 3B). The largest noise was detected in the conditionally essential genes included in Taniguchi *et al.*'s data (Fig. 3A). This inherent magnitude of noise in the expression of these responsive genes could be reduced (approximately 70%) through regulation replacement. Since the conditionally required genes are highly responsive to external perturbations, the noise in gene expression may be amplified during their multi-step regulatory processes. Taken together, the results indicated that the native regulation contributed a constantly high magnitude of extrinsic noise in gene expression and suggest a potential relationship between high plasticity and high expression noise.

Discussion

Native regulation was responsible for a relatively high magnitude of the extrinsic noise, which could be further repressed. Although the data set was quite small, this finding was partially supported by a reanalysis of other data sets. The reason why the extrinsic noise in the gene expression was reduced through the insertion of a foreign promoter is unclear. We assumed that replacing the native regulators with P_{tet} released the endogenous signal transduction mechanisms that affected the transcription or translation process of the target gene, which resulted in a decrease in the extrinsic noise. In addition, the insertion of the P_{tet} promoter also led to changes in the ribosome binding sites, which affected the translation efficiency. The reduction in the extrinsic contribution to the total noise may be caused by both transcription and translation.

We previously reported that the restricted gene expression and the chromosomal location independency mediated by the native regulation [27] reflect highly stringent control. These native regulation-mediated processes may be valuable for energy saving, which suggests a predominant theory of “not too much” in living cells. As the low-noise regulation appeared to be favourable for the achievement of efficient response to external changes, we investigated the reasons behind the significant magnitude of noise that remained. There may be a balance between noise and plasticity in gene expression [41], which links to gene category.

Comparisons using our transcriptome data and the reanalyzed proteome data revealed that the gene expression presented a common feature at both the mRNA and the protein level; the order of the average gene expression level was linked to the order of the gene category, which was determined by the essentiality of the gene (Fig. 3). This order of the protein/mRNA abundance correlated with the magnitude of the extrinsic noise that is shown in Fig. 3B. Because the extrinsic noise is independent of the expression level, this linkage strongly suggests that the essentiality in the gene function is not only associated with the power of the steady expression level but is also linked to the noise level mediated by the regulation. In Fig. 3A, the genes of moderate essentiality (*i.e.*, conditionally required genes) exhibited the largest noise. The results based on our mini-library, which largely contained the genes responsible for the starvation response, verified that this high magnitude of noise could be further repressed. A high sensitivity accompanied by a relatively large noise was interpreted as plasticity in gene expression [21,22]. Thus, these results may reflect a potential relationship among the plasticity in gene expression, the essentiality in gene function and the responsivity to external perturbations.

In summary, noise in gene expression has been studied to illustrate a global view of the protein abundance and fluctuations *in vivo* and to highlight a particular type of gene regulation. Further systematic studies are required to illustrate a solid overview of the noise in the

expression of the genes in different functional categories and to clarify whether our conclusion that “released regulation leads to reduced noise” is universal across all genes.

Acknowledgements

This work was partially supported by a Grant-in-Aid for Scientific Research on Innovative Areas no. 26119702 (to BWY) from the Ministry of Education, Culture, Sports, Science and Technology, Japan. We thank Natsuko Yamawaki for the technical assistance and Saburo Tsuru for analytical assistance.

References

- [1] M.B. Elowitz, A.J. Levine, E.D. Siggia, P.S. Swain, Stochastic gene expression in a single cell, *Science*, 297 (2002) 1183-1186.
- [2] E. Kussell, S. Leibler, Phenotypic diversity, population growth, and information in fluctuating environments, *Science*, 309 (2005) 2075-2078.
- [3] E.M. Ozbudak, M. Thattai, I. Kurtser, A.D. Grossman, A. van Oudenaarden, Regulation of noise in the expression of a single gene, *Nat Genet*, 31 (2002) 69-73.
- [4] W.J. Blake, K.A. M, C.R. Cantor, J.J. Collins, Noise in eukaryotic gene expression, *Nature*, 422 (2003) 633-637.
- [5] M. Acar, J.T. Mettetal, A. van Oudenaarden, Stochastic switching as a survival strategy in fluctuating environments, *Nat Genet*, 40 (2008) 471-475.
- [6] N.Q. Balaban, J. Merrin, R. Chait, L. Kowalik, S. Leibler, Bacterial persistence as a phenotypic switch, *Science*, 305 (2004) 1622-1625.
- [7] G.M. Suel, R.P. Kulkarni, J. Dworkin, J. Garcia-Ojalvo, M.B. Elowitz, Tunability and noise dependence in differentiation dynamics, *Science*, 315 (2007) 1716-1719.
- [8] J.R. Newman, S. Ghaemmaghami, J. Ihmels, D.K. Breslow, M. Noble, J.L. DeRisi, J.S. Weissman, Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise, *Nature*, 441 (2006) 840-846.
- [9] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, N. Barkai, Noise in protein expression scales with natural protein abundance, *Nat Genet*, 38 (2006) 636-643.
- [10] Y. Taniguchi, P.J. Choi, G.W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, X.S. Xie, Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells, *Science*, 329 (2010) 533-538.
- [11] F. Jacob, J. Monod, Genetic regulatory mechanisms in the synthesis of proteins, *J Mol Biol*, 3 (1961) 318-356.

- [12] C. Yanofsky, T. Platt, I.P. Crawford, B.P. Nichols, G.E. Christie, H. Horowitz, M. VanCleemput, A.M. Wu, The complete nucleotide sequence of the tryptophan operon of *Escherichia coli*, *Nucleic Acids Res*, 9 (1981) 6647-6668.
- [13] E.B. Keller, J.M. Calvo, Alternative secondary structures of leader RNAs and the regulation of the *trp*, *phe*, *his*, *thr*, and *leu* operons, *Proc Natl Acad Sci U S A*, 76 (1979) 6186-6190.
- [14] R.M. Gemmill, S.R. Wessler, E.B. Keller, J.M. Calvo, *leu* operon of *Salmonella typhimurium* is controlled by an attenuation mechanism, *Proc Natl Acad Sci U S A*, 76 (1979) 4941-4945.
- [15] T.M. Henkin, C. Yanofsky, Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions, *Bioessays*, 24 (2002) 700-707.
- [16] J.J. Tabor, T.S. Bayer, Z.B. Simpson, M. Levy, A.D. Ellington, Engineering stochasticity in gene expression, *Mol Biosyst*, 4 (2008) 754-761.
- [17] J.G. Lawrence, Shared strategies in gene organization among prokaryotes and eukaryotes, *Cell*, 110 (2002) 407-413.
- [18] P.S. Swain, Efficient attenuation of stochasticity in gene expression through post-transcriptional control, *J Mol Biol*, 344 (2004) 965-976.
- [19] K. Sneppen, S. Pedersen, S. Krishna, I. Dodd, S. Semsey, Economy of operon formation: cotranscription minimizes shortfall in protein complexes, *MBio*, 1 (2010).
- [20] K. Kovacs, L.D. Hurst, B. Papp, Stochasticity in protein levels drives colinearity of gene order in metabolic operons of *Escherichia coli*, *PLoS Biol*, 7 (2009) e1000115.
- [21] J.K. Choi, Y.J. Kim, Intrinsic variability of gene expression encoded in nucleosome positioning sequences, *Nat Genet*, 41 (2009) 498-503.
- [22] I. Tirosh, N. Barkai, K.J. Verstrepen, Promoter architecture and the evolvability of gene expression, *J Biol*, 8 (2009) 95.
- [23] B. Lehner, Selection to minimise noise in living systems and its implications for the evolution of gene expression, *Mol Syst Biol*, 4 (2008) 170.
- [24] B. Lehner, Conflict between noise and plasticity in yeast, *PLoS Genet*, 6 (2010) e1001185.
- [25] R.S. Cox, 3rd, M.G. Surette, M.B. Elowitz, Programming gene expression with combinatorial promoters, *Mol Syst Biol*, 3 (2007) 145.
- [26] V. Sasson, I. Shachrai, A. Bren, E. Dekel, U. Alon, Mode of regulation and the insulation of bacterial gene expression, *Mol Cell*, 46 (2012) 399-407.
- [27] B.W. Ying, S. Tsuru, S. Seno, H. Matsuda, T. Yomo, Gene expression scaled by distance to the genome replication site, *Mol Biosyst*, 10 (2014) 375-379.
- [28] J.M. Pedraza, A. van Oudenaarden, Noise propagation in gene networks, *Science*, 307 (2005) 1965-1969.

- [29] Y. Shimizu, S. Tsuru, Y. Ito, B.W. Ying, T. Yomo, Stochastic switching induced adaptation in a starved *Escherichia coli* population, *PLoS One*, 6 (2011) e23953.
- [30] S. Tsuru, J. Ichinose, A. Kashiwagi, B.W. Ying, K. Kaneko, T. Yomo, Noisy cell growth rate leads to fluctuating protein concentration in bacteria, *Phys Biol*, 6 (2009) 036015.
- [31] S. Tsuru, N. Yasuda, Y. Murakami, J. Ushioda, A. Kashiwagi, S. Suzuki, K. Mori, B.W. Ying, T. Yomo, Adaptation by stochastic switching of a monostable genetic circuit in *Escherichia coli*, *Mol Syst Biol*, 7 (2011) 493.
- [32] J. Paulsson, Summing up the noise in gene networks, *Nature*, 427 (2004) 415-418.
- [33] P.S. Swain, M.B. Elowitz, E.D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, *Proc Natl Acad Sci U S A*, 99 (2002) 12795-12800.
- [34] O.K. Silander, N. Nikolic, A. Zaslaver, A. Bren, I. Kikoin, U. Alon, M. Ackermann, A genome-wide analysis of promoter-mediated phenotypic noise in *Escherichia coli*, *PLoS Genet*, 8 (2012) e1002443.
- [35] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner, H. Mori, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection, *Molecular systems biology*, 2 (2006) 2006 0008.
- [36] E.M. Ozbudak, M. Thattai, H.N. Lim, B.I. Shraiman, A. Van Oudenaarden, Multistability in the lactose utilization network of *Escherichia coli*, *Nature*, 427 (2004) 737-740.
- [37] A. Kashiwagi, I. Urabe, K. Kaneko, T. Yomo, Adaptive response of a gene network to environmental changes by fitness-induced attractor selection, *PLoS One*, 1 (2006) e49.
- [38] E. Stolovicki, T. Dror, N. Brenner, E. Braun, Synthetic gene recruitment reveals adaptive reprogramming of gene regulation in yeast, *Genetics*, 173 (2006) 75-85.
- [39] S. Stern, T. Dror, E. Stolovicki, N. Brenner, E. Braun, Genome-wide transcriptional plasticity underlies cellular adaptation to novel challenge, *Mol Syst Biol*, 3 (2007) 106.
- [40] A. Hilfinger, J. Paulsson, Separating intrinsic from extrinsic fluctuations in dynamic biological systems, *Proc Natl Acad Sci U S A*, 108 (2011) 12167-12172.
- [41] D. Bajic, J.F. Poyatos, Balancing noise and plasticity in eukaryotic gene expression, *BMC Genomics*, 13 (2012) 343.

Table 1. A constant magnitude of extrinsic noise. The genetic backgrounds and measuring methods used in the various studies are summarized. Method: the measuring approaches used for the detection and analyses (FCM, flow cytometry); Genetic design: the structure and technique applied during the genetic construction; CV^2_{ext} : the extrinsic source of the total noise; Data sets: references for the data source. ¹The values were modified taking into account the special gating process used in the data processing reported by Silander *et al.* (estimated to exhibit approximately 20% lower noise [34]), the values were modified.

Method	Genetic design	CV^2_{ext}	Data sets
microscope	gene fusion genome-integrated native promoter	0.13 – 0.15	Taniguchi <i>et al</i> [10]
FCM gated cell	gene replacement plasmid-oriented native promoter	0.08 – 0.11 (0.10 – 0.13) ¹	Silander <i>et al</i> [34]
FCM total cell	gene replacement genome-integrated native promoter	0.14	this study

Figure Legends

Figure 1. Gene expression noise in the dual format. The relationship between the gene expression noise (CV^2) and the chromosomal location (**A**) and the average protein concentrations of all constructs (**B**) are shown. Left and right panels represent the foreign ($gene::P_{tet}-gfp$) and native ($gene::gfp$) regulation constructs, respectively. A close-up view of the P_{tet} -oriented expression is supplied in the inset. The standard deviations (errors) of repeated tests are indicated. The native regulation constructs that exhibited low fluorescence intensity were removed from the theoretical fitting. The average protein concentrations are from previous report [27]. The theoretical fitting of the extrinsic (broken horizontal lines) and intrinsic (dotted and dashed lines) sources of total noise (black solid lines) are shown. The values of the estimated extrinsic noise are indicated. The grey areas represent the 95% confidence band of the fit of the total noise.

Figure 2. Analyses of the global data sets. The gene expression noise of additional data formats was evaluated. The data sets from the studies of Taniguchi *et al.* (**A**) and Silander *et al.* (**B**) were reanalyzed and subjected to theoretical fitting (red solid lines). The gene expression noise is plotted against the mean protein abundance using the corresponding units. The genes were divided into four categories: all (white), essential (blue), conditionally essential (green) and other (grey) genes. The extrinsic source of the total noise is indicated by the broken horizontal lines, and the estimated values are indicated around the fitting lines. The grey areas represent the 95% confidence band of the fit of the total noise.

Figure 3. Box plots of gene expression patterns. The gene expression at the mRNA level is shown for the transcriptome data of MG1655 and MDS42 (**A**). The gene expression at the protein level is shown from the proteome data adopted from Taniguchi *et al.* (**C**) and Silander *et al.* (**B**). The genes are divided into four categories: all (white), essential (blue), conditionally required (green) and other (grey) genes. The respective numbers of genes in the all, essential, conditionally essential and others categories are as follows: 4,428, 302, 275 and 3,851 in MG1655 (**A**, left panel); 3,732, 302, 260 and 3,170 in MDS42 (**A**, right panel); 1,832, 111, 116 and 1,605 in Silander *et al.* (**B**); and 1,018, 115, 89 and 814 in Taniguchi *et al.* (**C**), respectively.

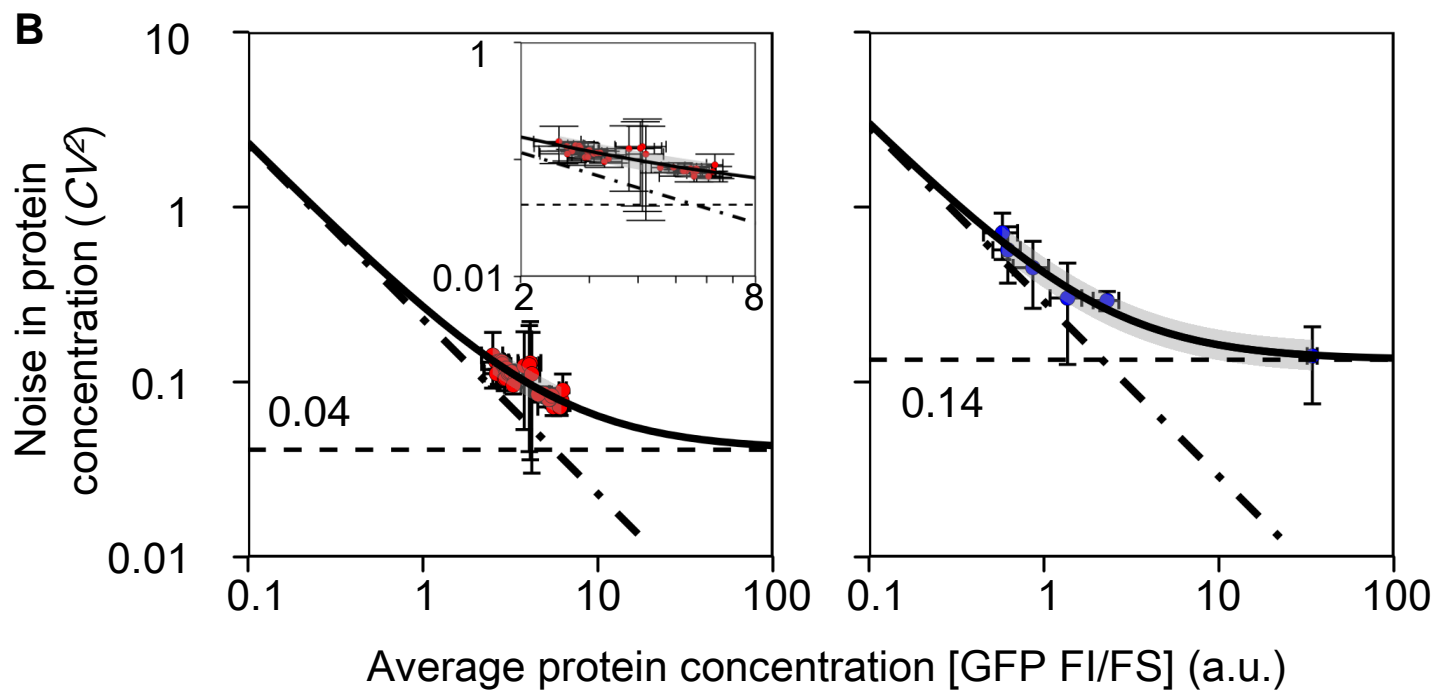
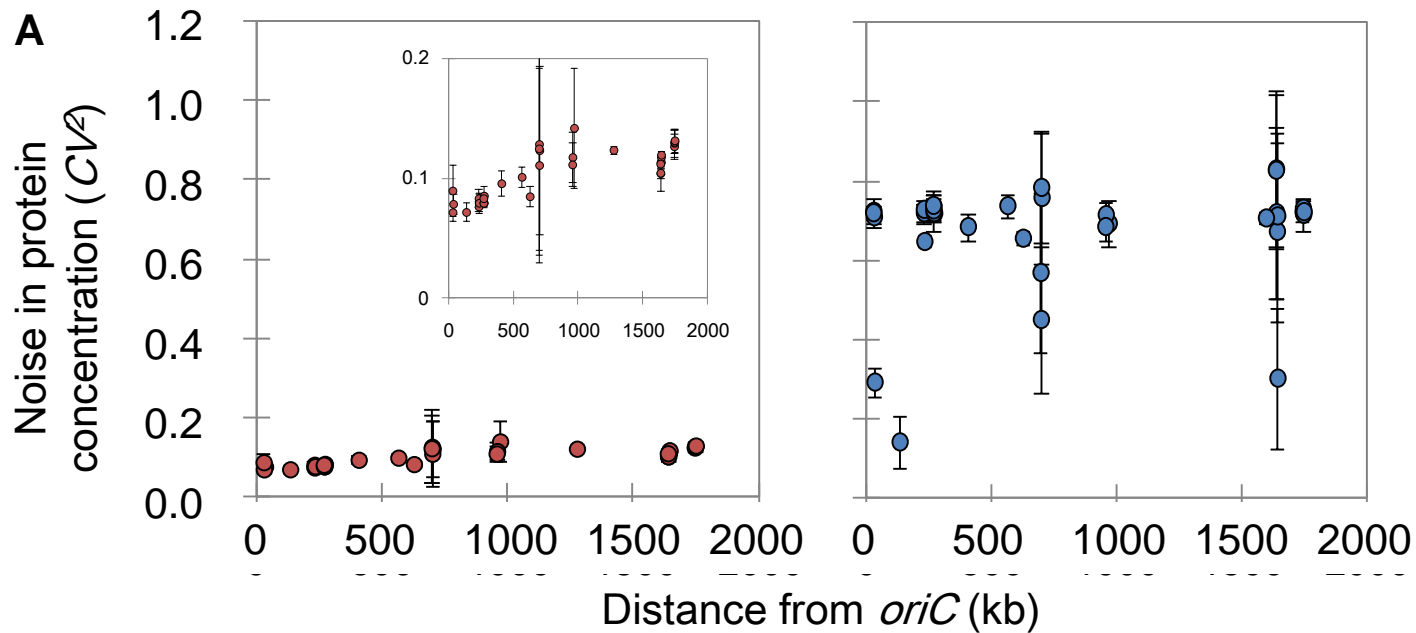


Figure 1

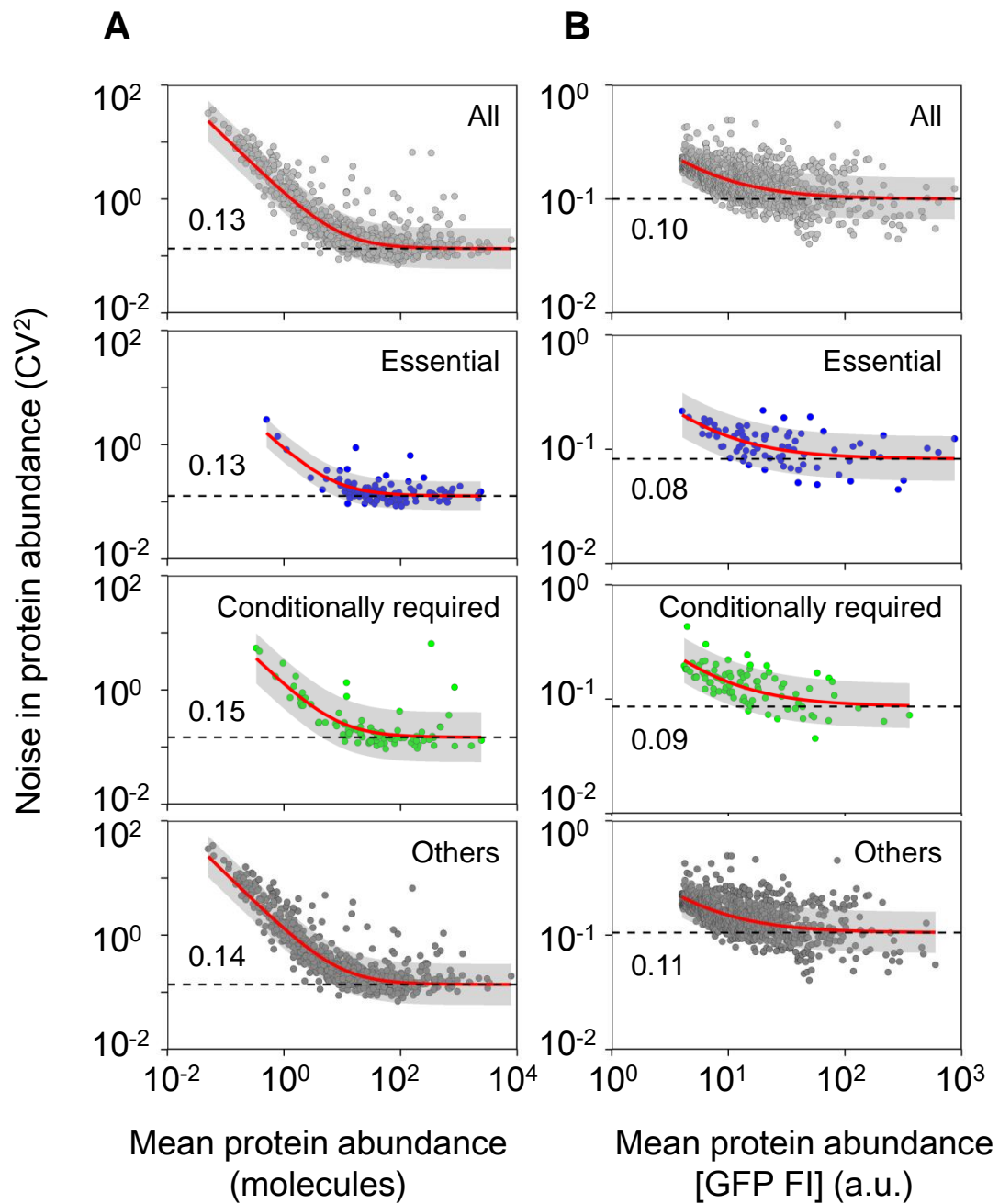


Figure 2

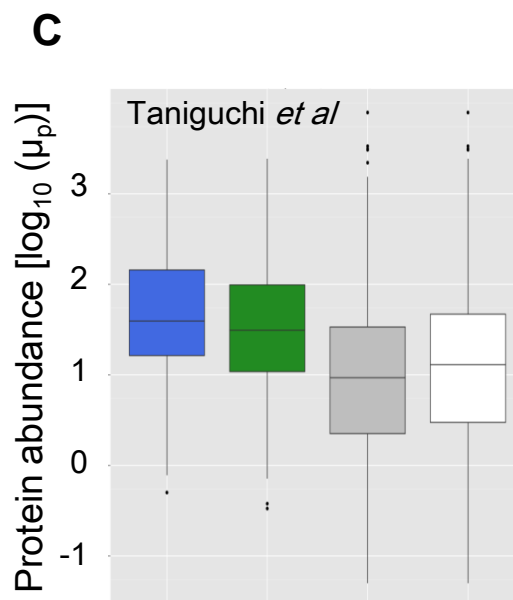
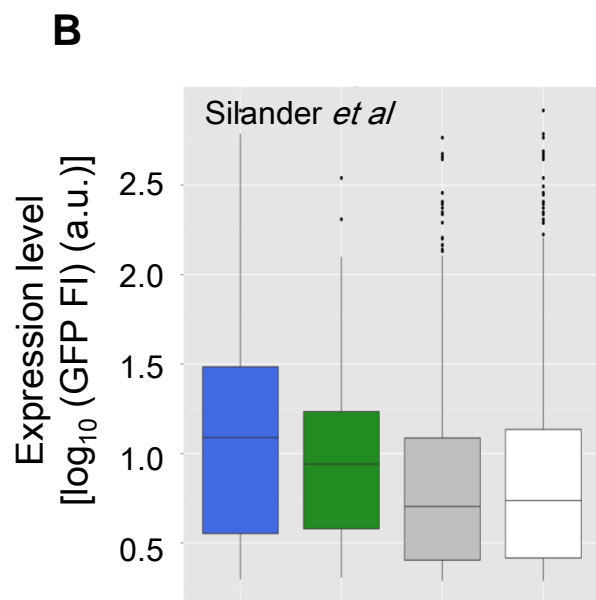
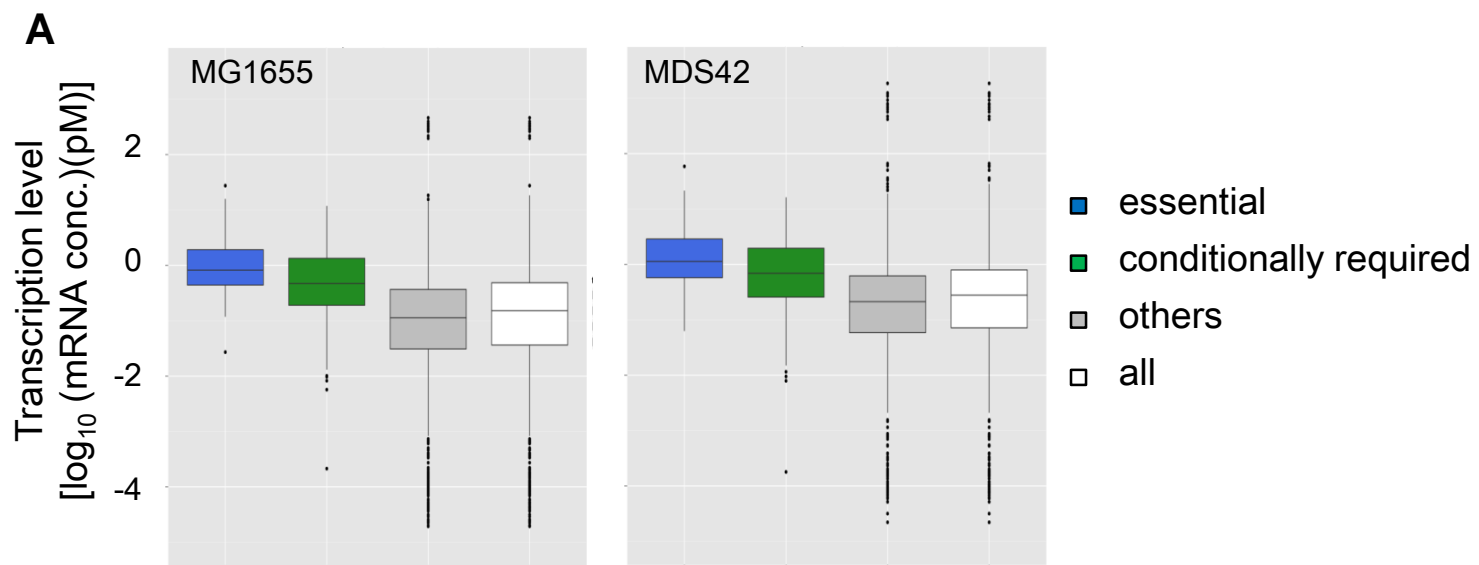


Figure 3

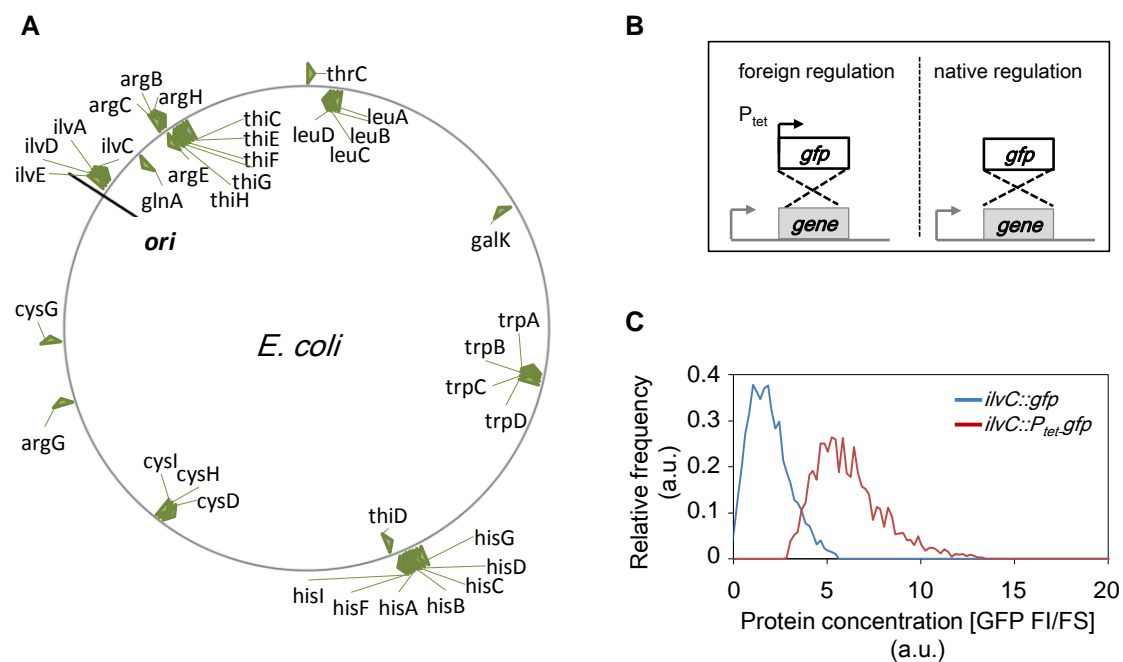


Figure S1. A summary of genetic constructs and FCM analysis. **A.** A genome map of the candidate genes. The chromosomal locations of the genes (native regulations) are marked on the *E. coli* genome of MDS42. The 37 genes employed for the genome replacement are indicated. The position of *oriC* is the initiation site of genome replication. **B.** A schematic drawing of the dual format. The target genes (indicated in **A**) that were replaced with the native (right) or P_{tet} -oriented (left) reporter gene, *gfp*, are shown. **C.** An example of FCM data set. The exponentially growing cell populations with the *ilvC::gfp* (red) and *ilvC::P_{tet}-gfp* (blue) genetic constructs were measured using flow cytometry. The distributions of the cellular protein concentrations, as reported by the genome-integrated *gfp*, are shown. GFP FI and FS (forward scattering) represent the fluorescence of the green emission from GFP and the relative cell size, respectively.